Università di Pisa

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
Corso di Laurea Magistrale in Ingegneria Elettronica

Tesi di laurea magistrale

Design of a CMOS chopper instrumentation amplifier with rail-to-rail input and output ranges

Candidato:
Giovanni Gabrielli

Relatore:
Prof. Paolo Bruschi

Correlatore:
Ing. Massimo Piotto

Anno Accademico 2014–2015

Abstract

This thesis deals with the design of a current feedback instrumentation amplifier, optimized for the readout of thermal sensors. This topology stands out for its excellent CMRR and the predisposition to feature low frequency error reduction techniques. Versatility is a main target for this work: 1 kHz bandwidth and Rail-To-Rail input common mode range allow the interfacing of a wide variety of sensors.

Chopper modulation is used to reduce offset and flicker noise, achieving a 19 nV/$\sqrt{\text{Hz}}$ RTI noise density and a flicker corner frequency of 6 mHz. A low total output noise power is achieved as well, reaching an ENOB of 12 bits with less than 350 µA current consumption.

The peculiar issue for this architecture, that is gain error, is solved by means of Port Swapping technique, together with an input Common Mode Equalization, that reduce untrimmed gain error to 0.5%. Chopped offset and Port Swapping ripple are completely filtered away by a third order Butterworth State Variable low pass filter, implemented with Gm-C integrators.

Contents

List of Figures

List of Tables

Introduction

Nowadays, while digital electronics is striving to reach higher and higher computational power and speed, analog electronics has found a renewed, fundamental role in the field of sensors readout. Sensors are used for in a wide variety of applications, from safety to environmental awareness, and can measure quantities such as temperature, pressure, radiation intensity, acceleration and angular velocity, strain and many others. The most important boost to the developement of sensors was the introduction of MEMS in measurement systems: indeed, these extremely compact, low power, precise and low cost devices allow the integration of sensors in almost every context of everyday life, and smartphones are only one of the several examples of this trend.

In precision applications, it is compulsory to design a readout interface, called Analog Front End, that allows an accurate conversion of the measured quantity to a digital code. One of the most diffused architecture for an AFE is the instrumentation amplifier (InAmp), that is basically a differential amplifier with precise gain and high input impedance, thus able to interface a great variety of sensors. Among the several topologies, current feedback instumentation amplifiers are often used, due to their high CMRR, effeciency and flexibility to be easily adapted to met different specification. The most relevant issue is gain error, caused by mismatch between the input stages, that can be critical when the circuit is implemented in a CMOS technology.

This technology is widely used thanks to its reduced cost, but CMOS analog circuits are affected, as known, by high offset, caused again by matching errors, and flicker noise. These problems are usually mitigated with the introduction of dynamic techniques such as Auto-Zero, Correlated Double Sampling and Chopper Modulation.

The purpose of this thesis is to propose an original solution for an instrumentation amplifier, with advanced strategies to reduce errors. Chopper modulation is used to reject offset and flicker noise, while the consequent products of demodulation (chopped offset) are strongly attenuated by a

state variable third order Butterworth low pass filter. Port swapping is used to reduce the effects of the mismatch between the input stages, and thus to increase the gain accuracy, while it has a benefical side effect in increasing the input impedance, that with a simple chopper modulation would be too low for many application. On one hand, compactness and low power consumptions were goals as well; on the other hand, flexibility was the main keywork for this work. Indeed, while this amplifier is optimized to interface thermal sensors, it can work with a great variety of sensors thanks to its Rail-To-Rail common mode input range, that makes it suitable also for other application, such as current sensing.

Chapter 1 will present an overview of the most diffused sensors, to give an idea of the importance of analog design even in the digital era of these days and age. Then, sensors will be classified according to their output quantity, and the instumentation amplifier will be presented, together with a brief description of all its characteristics, as the best achitecture to interface these sensors.

Chapter 2 will present the state of the art in the design of instrumentation amplifiers, highlighting pros and cons of the different topologies. From the most relevant references available in the literature, examples will be given for each one of them.

Chapter 3, instead, will focus on the reference design for this work, underlining the unsolved issues that were the starting point for this work. Then, the new solution is proposed, and the high level design is described: at this point, ideal stages will be considered, that however will allow the first expectations on the final performances of the proposed architecture.

A detailed transistor level description of each single block will be given in chapter 4, with an accurate mathematical analysis of the relevant equations that led to the sizing of the single devices. Noise, ranges, current consumption, compactness and feedback stability will be the main costraints to lead the transistor level design.

Finally, in chapter 5 we will present the most relevant simulations performed on the complete system, in order to evaluate the results of the accurate design and to present the most significant electrical specifications of the proposed solutions.

Chapter 1

Sensors and instrumentation amplifiers

1.1 Diffusion of sensors

The impressive developement of electronics of the last few decades has deeply changed our lifestyle in several ways. On one hand, digital electronics has reached incredible levels of computational power and the telecommunication field has given us the chance to share information with the whole world in a stunningly simple and user-friendly way. On the other hand, the opportunity to acquire information about the physical world that surronds us was given by the development and the spread of sensors. Hundreds kinds of sensors are widely used in different context of our lives, and are more and more deeply embedded in our environment, so that common people usually don't even notice their presence nor appreciate their importance. But indeed they are often irreplaceable in many applications, whether they are just a means of monitoring physical quantities or they belong to a control loop together with a computational core and actuators.

Safety related applications There are many examples of critical contexts in which safety goals are achieved through a wide use of sensors. Let us think, for instance, of really complex architectures such as airplanes. The huge amount of flight and environmental variables, such as external air temperature and relative velocity, pressure and altitude, would make impossible for pilots alone to safely fly a plane, steadily facing all odds. Thus, these variables are constantly monitored by sensors literally spread all over the airplane, so letting pilots always make the better decision. In addition, redundant measurements (by means of multiple sensors of each

kind) make the monitoring so accurate that is possible to let an automatic control system grab the reins of the flight. The importance of these sensors is testified by rare examples in which they did not work properly. One of the most famous is without any doubt the accident of flight Air France 447, in which on June, 1 2009 the 228 passengers of an Airbus A330-200 lost their lives: the final report showed that the accident was most probably caused by an incoherent air speed data measured by the Pitot tubes, that were obstucted by ice crystals. This situation led to a sudden disabling of the automatic pilot and the switching to *alternate law* flight program, an emergency program in which many other safety systems are disabled. The reaction of the pilots to the consequent stall was too slow, leading to the crash. [1]

Sensors of various types are of course embedded also in less complex systems, like trains or cars, in which they can achieve safety-related goals or simply can assist the user in common situations. One of the most popular cars' safety system is, for example, the Anti-locking Brake System (ABS): this feature prevents the wheels from blocking due to a too strong braking, thus avoiding unexpected skidding and keeping the braking distance short. The system continuously monitors each wheel's rotational speed through dedicated sensors, in order to release the brake and let the wheel regain static friction with the ground in case one of them turns significantly slower than the others. Another useful safety system in cars is the so called Electronic Stability Control (ESC). During normal driving, ESC works in background and continuously monitors steering and vehicle direction. It compares the driver's intended direction (determined through the measured steering wheel angle) to the vehicle's actual direction (determined through measured lateral acceleration, vehicle rotation (yaw), and individual road wheel speeds). ESC intervenes only when it detects a probable loss of steering control: this may happen, for example, when skidding during emergency evasive swerves, understeer or oversteer during poorly judged turns on slippery roads, or hydroplaning. In this case, ESC controls independently each wheel's brake in order to create torque about the vehicle's vertical axis, opposing the skid and bringing the vehicle back in line with the driver's commanded direction. Other safety related sensors can be the pressure sensors embedded in the newest cars' tyres, or the ones that detects if the driver is sitting and driving without the seat belt fasten. In addition, a relatively old safety system is the airbag: this balloon must inflate within a few milliseconds after a crash, that is detected by a fast accelerometer.

Another important and well-known sensor in modern cars is the so called

[1] www.airfrance447.com

lambda sensor (figure 1.1). This sensor measures the quantity of uncombusted oxigen in the exaust stream, from which a CPU can calculate the air-fuel ratio in the engine and optimize the combustion process through a control loop.

Figure 1.1: A lambda sensor.

Another one is the speedometer, that is usually implemented as a set of magnets mounted on the output shaft or (in transaxles) differential crownwheel, or a toothed metal disk positioned between a magnet and a magnetic field sensor. Thus, the sensor produces a pulse wave at a frequency proportional to the average speed of the car. The speedometer, that usually is used just as an indicator for the driver, can be part of a control loop, like many other senors, to form a cruise control: this feature, present in almost every new car, allows the driver to set a speed and then controls the fuel stream in order to mantain that speed within a certain range. This is just one of the many applications created to make the driving more comfortable: let us think, for instance, to assisted parking systems, in which proximity sensors inform the driver about the distance between the car and obstacles (other cars), or the automatic air conditioner, that controls the room temperature of the car with a closed loop that needs temperature sensors.

These are just a few examples of the most popular sensors: an average car can embed between 60 and 100 sensors, while this numbers are projected to reach as many as 200 sensors per car within few years.

Security related sensors are obviously present also in other contexts, like in factories where workers deal with dangerous machines (presses, saws or grindstones), or for monitoring the temperature of ovens in concrete production, or in the chemical industry where the concentration of toxic or dangerous gases must be constantly kept under control. Also the houses often feature safety systems such as anti-theft or antifire devices that use dedicated sensors.

Sensors and environmental awareness Safety is not the only reason for such a wide diffusion of sensors. In fact, one of the newest and most influencing trends is the aim to monitor and thus to protect the environment against pollution, especially in big cities. This field, known as *environmental awareness*, has grown strong in the latest years and aims to create a network of sensors, especially within the biggest urban areas, to monitor all

the parameters that can affect the citizens' quality of life, such as air pollution but also traffic condition and availability of the major public services. This high level and complete monitoring, made possible also by the strong development of portable and low-cost digital devices with great communication capability, plays a key role in the creation of the idea of *Smart Cities*. One of the many definitions of this concept is given by Gerhard P. Hancke *et al.* (The Role of Advanced Sensing in Smart Cities,): "A smart city is a city which functions in a sustainable and intelligent way, by integrating all its infrastructures and services into a cohesive whole and using intelligent devices for monitoring and control, to ensure sustainability and efficiency." One of the most advanced and effective example of the idea of Smart City is represented by the project called *SmartSantander*. In the spanish city of Santander, researchers from several universities have built a wide network of sensor nodes in order to constantly monitor the major environmental parameters and share them with all the citizen. The project exploits the rising idea of the Internet of Things, that is a network of thousands of intelligent sensing nodes communicating through an internet based protocol. The collected data can be of help to the common citizen to get information about traffic, available parking spots, and many other public services. On the other hand, data can be monitored by experts in order to control the air pollution level, the wind speed or the light intensity. From the sensors' point of view, the major innovation is the wide use of sensor nodes, that are tiny embedded systems with strong connection between sensors and computational/communication core: only the city of Santander counts as many as 12000 sensor nodes, monitoring parameters such as temperature, CO, noise, light, car presence in parking spots, etc. This is a great boost for the development of small and low power sensors of different kinds, the improvement of sensor fusion technologies and, with the advancing of process technologies, the diffusion of mixed signal architectures within miniaturized Systems on a Chip.

MEMS sensors In fact, the major boost to the wide diffusion of sensors was due to the development of small, low power, integrated sensors called MEMS (Micro Electro-Mechanical Systems). With this technologies it is possible to miniaturized even complex mechanical structures such as multi axes accelerometers and gyroscopes, advanced chemical and optical sensors, temperature and pressure sensors, as well as compasses, microphones, microfluidic devices and many others, all within a Silicon die (thus on a millimetric surface). The possibility to build such mechanical structures on the Silicon surface with a common CMOS process has significantly lowered

these products' price, while usually really small dimensions bring low power consumption: these reasons make MEMS sensors particularly suitable for battery powered, portable devices like, most of all, smartphones.

In such devices, sensors are used to create a strong interaction with the user, even on a physical level: let us think, for example, of a multi axes accelerometer that informs the device about its orientation referred to the ground, thus, turning the screen in panorama mode when the phone's long edge is turned parallel to the ground (to better enjoy videos, for instance). Of course, the major benefits from the integration of accelerometers and gyros are earned by videogames, nowadays fully diffused in smartphones. This technology for videogames, actually, was first implemented by Nintendo when it launched on the market its Wii products: this console, first in its category, used interactive remotes to control the game directly with the user's movement. The accelerometer is nothing more than a mass-spring-damper system, with submillimetric characteristic dimensions, that translates an acceleration into a displacement: the body of the sensor is tied to the device, thus when an acceleration is imposed to it, inertial forces are induced on the suspended proof mass that shifts from its quiescent position proportionally (for small variations) to the acceleration. The displacement then can be measured by a system of capacitive plates, giving a voltage difference proportional to it as output. After an analog to digital conversion, the data can be processed by a CPU and used, for instance, to control the game parameters.

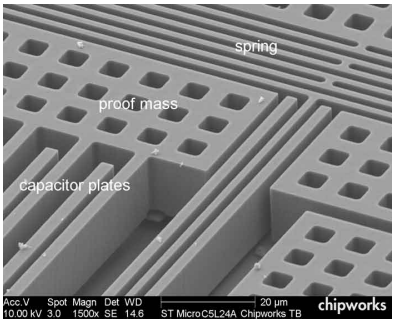


Figure 1.2: A particular of a MEMS accelerometer.

Obviously, movement sensors are not used only in videogames. For instance, gyroscopes have found an important application in another field of consumer electronics, like professional cameras. Almost all new lenses for reflex cameras, as a matter of fact, feature a Vibration Reduction (VR) system, that tends to avoid the motion blur: often, when light conditions are not critical but not excellent either, the exposure time is long enough to let the optical sensor be influenced by the natural vibrations of the photographer's hand, thus creating an unpleasant micro-blur effect. It has been proved that these movements consists primarily in rotations along a horizontal axis (pitching, that is a vertical movement) and a vertical axis (yawing, that is a horizontal movement). A system of two gyros detects each one of these effects: a dedicated CPU then samples the data around a thousand times per second and controls actuators, that compensate the

shift by moving the lens.[2]

Though, the most recent and advanced expample of miniaturization and integration between sensors and human life is probably the field of the *wearable* devices. This new concept exploits to a maximum level the concepts of miniaturization and energy efficiency in order to create and develop intelligent devices, thought to be constantly in touch with the human body, as an actual extension of it with a brand new set of possibility always at hand. One of the most important event in this field was the conference held in Santa Clara, California (US) on November 12-13, 2014, called "Wearable Sensors and Electronics" [3], where reseachers and CEOs from all over the world gathered together to discuss the development of wearable devices, boosting the interest in areas such as energy harvesting, printable electronics and, of course, all kinds of integrated, low power MEMS sensors. The research in this field is being currently led in several institutions, one for all the Center for Wearable Sensors of the UCSD Jacobs School of Engineering [4], Here they are working on the development of different kinds of sensors that should monitor several body parameters, through chemical, physical and electrophysiological techniques, sometimes even with polymer-based substrates that can be implanted inside the tissues for long term analysis. The aim of this project is, of course, healtcare and preventive medicine, through constant analysis of heart rate, hydration levels, blood sugar and more, or simply the user's comfort, like in smart clothes for personalized cooling and heathing.

1.2 Classification of sensors

As seen in the previous paragraph, electronic systems can deal with an enormous variety of sensors: what they have in common is the need to produce a value of the measured quantity to be elaborated by a processing unit, typically a digital CPU. Thus, a analog-to-digital converter (ADC) is needed to turn the analog output of the sensors into a properly coded binary number, so that the CPU can handle the data. On one hand, ADCs typically can convert an input voltage V_{in} into a coded binary number C_{out}, accordingly to an equation such as

$$C_{out} = \left\lfloor \frac{V_{in}}{V_{REF}} \cdot 2^n \right\rfloor \tag{1.1}$$

[2] http://www.nikon.com/about/technology/rd/core/software/vr_e/

[3] www.wearablesensors2014.com

[4] http://www.jacobsschool.ucsd.edu/wearablesensors

where V_{REF} is a reference voltage (often V_{DD}) and n is the number of bits. On the other hand, a sensor is by definition a transducer, that is a device that translate a physical quantity (acceleration, temperature, pressure, etc.) into an electrical one, but not necessarily a voltage; and even if it is a voltage, this is often not suitable for a direct conversion. Here comes the need for an analog block, called Analog Front End (AFE), whose goal is to transorm the sensor's output quantity into a suitable voltage for the ADC. Thus, from the point of view of an analog designer, one of the most relevant classifications of sensors is according to their output quantity.

Voltage The category of sensors that produce a voltage as output quantity is surely the most populated. It gathers different subcategories, the most important of which are listed below.

- **Thermoelectric sensors** can transform a temperature difference into a voltage thanks to the Seebeck effect. The most diffused example is the thermocouple, that is basically a set of two different metals joined toghether in two different points (junctions).

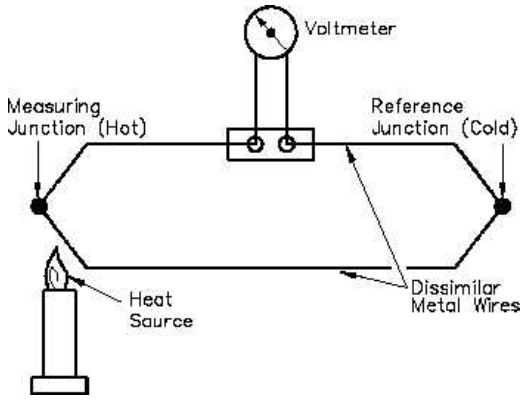


Figure 1.3: Electric schematic of a thermocouple.

If one of the metals is cut, and the two junctions (A and B) are at different temperatures (T_A and T_B), a voltage difference V_{out} is generated between the two ends of the cut metal. Usually one junction, called *measuring junction* (hot), is actually used to measure an absolute temperature T_x, and the other, called *reference junction* (cold), is kept at a reference temperature T_{REF}, like room temperature (that can be known or opportunely compensated) or at $0°C$ in a melting ice mixture. The output voltage is then a function of the temperature difference and can be approximated by a piecewise linear function.

For most application, the function can be considered linear:

$$V_{out} \approx k(T_x - T_{REF}) \tag{1.2}$$

where the constant k depends strongly on the materials involved. One of the reason of the huge diffusion of thermocouples is that they can be easily integrated in a silicon chip: usually the two conductive materials are n and p polysilicon wires, with a silicide metalization on the junction to avoid a rectifying behaviour.

Furthermore, temperature measurements can be used for many other indirect measurements: it can be, in fact, the basic element of fluximeters, bolometers (IR light level detectors) or hot plates (sensors of eso/endothermic chemical reactions).

- **Electrochemical sensors** measure ion concentration in a liquid or a gas. Similarly to galvanic cells, they produce an output voltage proportional to the logarithm of a ion concentration:

$$V_{out} = k \cdot \log[C] \tag{1.3}$$

They can be used, for instance, to measure pH in solution, or gas concentrations like the *lambda sensor* (see above) to monitor the O_2 level in the exaust stream.

- **Piezoelectric sensors** are crystals that, if compressed by an external force, produce on their faces an electric charge proportional to it (actually they are sensible only to force variations), thus a voltage difference can be measured. These are characterized by an extremely high sensitivity, that makes them perfect for integrated microphones (and, used as actuators, small speakers) or tactile sensors.

- Another widely used category of sensors with voltage as output quantity is that of **magnetic Hall sensors**. These are basically metal lines biased with a constant current I: if a magnetic field $\mathbf{B}$ is present, the current carriers travelling through the conductor undergo the Lorentz force, that pushes them in a direction orthogonal to both the magnetic field and their velocity. In the steady state condition, an electric field is generated in order to compensate the charge displacement, and the consequent voltage difference can be measured. Since the Lorentz force is $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ and is equal to the electrostatic force, then $E = vB_{\perp}$. We have then $V = Ew$, $I = j \cdot wh$ and $j = qnv$. Finally an expression for the measured voltage is:

$$V_H = \frac{1}{hnq} I B_\perp = k_H I B_\perp \tag{1.4}$$

being k_H a constant depending only on the sensor's electrical and geometrical parameters.

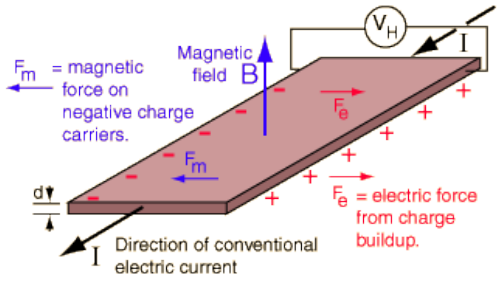


Figure 1.4: Hall effect in magnetic sensors.

A single Hall plate, like the one in figure 1.4, is obviously sensible to only one of the three components of the magnetic field **B**: in order to measure the direction of the field, three different plates are needed, and the three independent measurements must be opportunely elaborated by a CPU. This, for example, is how an integrated compass is achieved.

Even though these sensors produce a voltage as output quantity, in most cases it is not ready to be directly converted by an ADC (the reasons for this will be discussed later). Thus, a typical interface for this kind of sensors is the **instrumentation amplifier** (In-Amp): this device is mainly a differential amplifier with extremely precise gain, in order to adapt the input signal to the ADC input ranges (to increase resolution) without loss of accuracy in the measurement. All the characteristics of this kind of amplifier will be deeply discussed in the next paragraphs.

Current Among the sensors that produce a current as output quantity, one of the most relevant examples is the *photodiode*. These are basically *pn* junctions that, if hit by energetic particles like photons, produce electron-hole pairs, that is an inverse bias current proportional to the number of photons that hit it (thus to the light intensity). To transform this current into a voltage, a first solution could be to simply bias a test resistor with that current and measure the voltage drop across the resistor. Actually, the sensor is nothing like an ideal current generator (its output impedance is often too low), so this way the loading effect would not be negligible. Instead,

a Trans-Inpedance Amplifier (TIA) is more often used, that guarantees a low loading effect at least al low frequencies.

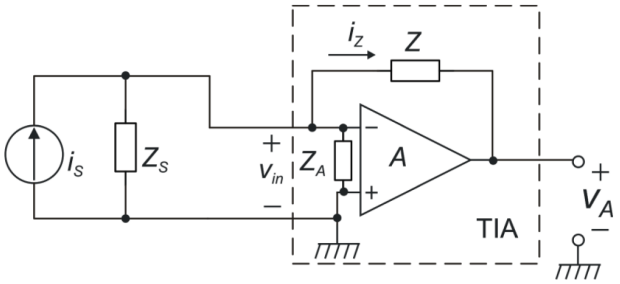


Figure 1.5: TIA for interfacing sensors.

Other sensors have not a current output, but more precisely a charge output, that is the output quantity proportional to the signal to be measured is an electric charge. Though, the only relevant example is the Charge Coupled Device (CCD), a light sensor used years ago for image aquisition (mainly digital cameras), now almost completely replaced by CMOS sensors, that are basically photodiodes.

Capacitance Sensors belonging to this category have a capacitive structure, whose value of capacitance can vary mainly for two reasons. In one case, a physical quantity can modify the value of the relative dielectric constant ϵ_r of the material between the conductive plates: in this case chemical substances can react with the dielectric material, obtaining thus a measure of its concentration. Far more frequent is the case in which the capacitance is changed because the geometry of the structure varies, that is the plates shift one on another (reducing the effective area of the capacitor) or simply they get closer or more distant. This, for example, is the case of inertial sensors, such as accelerometers and gyroscopes: as said in chapter 1, these consist in a proof mass that, thanks to inertial forces, shift from the quiescent position. Then, they usually feature one conductive plate per side while two others plates are tied to the substrate, as seen in figure 1.6.

So, being $C_0 = \epsilon A/d$ the quiescent value of the two capacitances, when the mass shifts towards one side by a distance x the two capacitances change their value as follows:

$$C_1 = \epsilon \frac{A}{d+x} = \frac{C_0}{1+\frac{x}{d}} \tag{1.5}$$

$$C_2 = \epsilon \frac{A}{d-x} = \frac{C_0}{1-\frac{x}{d}} \tag{1.6}$$

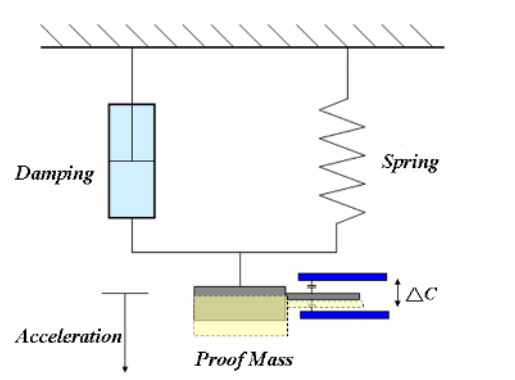


Figure 1.6: Capacitive structure of inertial sensors.

One simple way to extract a voltage from this structure is to supply a reference $\pm V$ across the series of the two capacitances (see figure 1.7). The voltage present in the central node will then be:

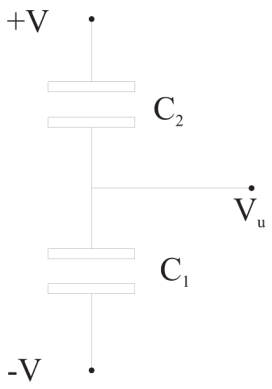


Figure 1.7: Capacitive interface for inertial sensors.

$$V_{out} = V\frac{C_2 - C_1}{C_2 + C_1} = V\frac{1 + \frac{x}{d} - 1 + \frac{x}{d}}{1 + \frac{x}{d} + 1 - \frac{x}{d}} = V\frac{x}{d} \tag{1.7}$$

Alternatively, the measure can be made with means of a TIA and a sinusoidal generator, or with charge amplifiers (not discussed here).

Resistance A good variety of sensors have an output quantity that is a resistance: this is the case, for instance, of temperature sensors like thermistors or RTDs (Resistance Temperature Detectors). The latter are frequently Platinum resistors (PT100), that features a particularly linear and accurate T-R transcharacteristic:

$$R(T) = R_0 \left[1 + \alpha(T - T_0)\right] \tag{1.8}$$

where R_0 is the resistance shown at T_0 (for PT100 these values are 100Ω at $0°C$). The coefficient α, also called Temperature Coefficient of Resistance (TCR) is crucial for the accuracy of the sensor: PT100 is often used because it has a particularly constant and precise value of TCR ($0.00385/K$ over a wide range of temperatures).

Another important sensor belonging to this category is the piezoresistor, a resistor that changes its value of resistance when strained: it is used widely as a *strain gauge* to measure the strain ($\epsilon = \Delta l/l$) of mechanical structures. Its resistance, at least for small variations, can be written as

$$R(\epsilon) = R_0(1 + G\epsilon) \tag{1.9}$$

where G is called gauge factor and depends strongly on the material. For example, for non-piezoresistive material this effect is only caused by geometric reasons: when a metal resistor is stretched, it gets longer and narrower, so its resistance gets higher, and vice versa. For piezoresistive materials, the effect is enhanced because the value of resistivity (independent on geometry) can change strongly with the strain.

In order to interface these kinds of sensors, an accurate bias current could be sufficient to produce a voltage, but in general it is preferred to use a configuration called *Wheatstone bridge* (figure 1.8). Since both RTDs and strain gauges have the same kind of mathematical expression for R, let's consider the latter, with the variable $x = G\epsilon$.

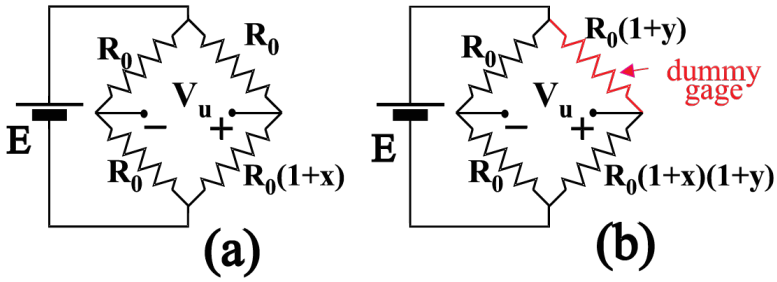


Figure 1.8: Wheatstone bridge configuration to interface resistive sensors.

The output differential voltage is then:

$$V_{out} = E\left(\frac{R_0}{R_0(2+x)} - \frac{R_0}{2R_0}\right) = -E\frac{x}{2(2+x)} \approx -\frac{E}{4}x \tag{1.10}$$

for small variations. The main advantage of using this configuration is that, if the R_0 resistors are dummy sensors (that is not used to measure anything, but shares the same environmental conditions) identical to the active one, all other dependancies and drifts are compensated. See again figure 1.8: on the right, the temperature dependance is expressed through the factor $(1 + y)$. It is evident that this way the output voltage is completely independent from thermal drifts. Furthermore, if strain gauges could be placed in antisymmetrical position on the object to be monitored, and properly connected in the Wheatstone bridge, the sensitivity could be increased by a factor of 2 (or as much as 4 if 4 gauges are used).

Thanks to the Wheatstone bridge, again the best solution to deal with this output voltage is an instrumentation amplifier, that will now be discussed in detail.

1.3 In-Amps for sensor interfacing

We have said that, even for sensors that already present a voltage as output quantity (or resistive sensors in a Wheatstone bridge configuration), it's impossible to interface them directly to an ADC: in fact, many reasons concurr to this conclusion, the most relevant of which is related to the concept of Dynamic Range (DR).

Dynamic Range can be defined as the ratio between the greatest measurable value (Full Scale voltage) and the smallest detectable value (Detection Limit voltage):

$$DR = \frac{\Delta V_{FS}}{V_{DL}} \tag{1.11}$$

Now, the term ΔV_{FS} is quite straightforward to understand: it is simply the sensor's output voltage swing correspondent to the difference between the highest and the lowest possible values of the measured quantity, and it depends only on the sensor's characteristics. As said above, it's quite common to have full scale voltages as high as a few mV: for example, a K type thermocouple (the most common type) has a sensitivity of $41\mu V/°C$, that is a voltage of only $4.1mV$ for a $100°C$ full scale. The term V_{DL}, on the other hand, is strongly dependent on the *noise voltage* that is summed to the signal. This noise, in a first approximation, can be considered gaussian white noise because originated by random thermal agitation of the carriers (thermal noise). Although it can spread among a non limited interval of values, we can indentify bands in which the noise voltage will *statistically* reside: that is, given a certain RMS value v_n, equal to the standard deviation

σ_n of the gaussian distribution, it is known that the noise will assume values in the range $\pm 2\sigma$, for example, in 95.4% of a given time interval. Then, we can consider a band wide $4\sigma_n$, that sets the measurement's detection limit: actually we have assimilated our noise to a periodic wave with a crest factor of 2, that is a $v_{n-pp} = 4v_n$. This value v_{n-pp} is then the width of the noise band of the signal: it is evident than that two different values of the signal can be distinguished reliably only if their noise bands don't overlap.

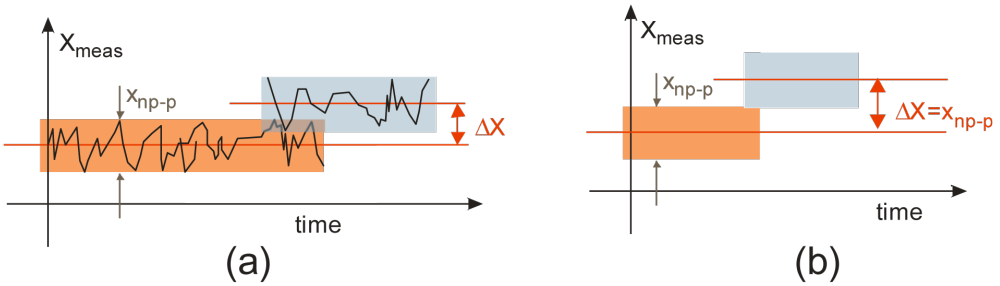


Figure 1.9: Noise bands determine the detection limit.

The relevance of Dynamic Range is that it is directly related to the Effective Number Of Bits (ENOB) that the following ADC will have. DR is in fact the number of different values of the input voltage that an ADC is able to distinguish: so an expression for the ENOB can be

$$ENOB = \lfloor \log_2(DR) \rfloor \tag{1.12}$$

Now, a crucial point is that DR cannot be raised during the signal elaboration: this is because each stage of the elaboration chain will surely present a Referred-To-Input (RTI) noise that will sum to the noise already present and will degrade the Dynamic Range. Our goal is to keep this degradation negligible, and the use of an In-Amp will be the best solution.

Let's suppose to place an ADC directly at the output of the sensor: the DR will then be (the notation v_{n-pp} will be omitted)

$$DR = \frac{\Delta V_{FS-s}}{v_{n-s} + v_{n-ADC}} \tag{1.13}$$

Since ADCs' RTI noise is usually extremely high and greater than the sensor's noise, such degradation of the dynamic range would compromise the measurement.

The use of a In-Amp as a *low noise preamplifier* solves this problem. In fact, the total RTI noise will now be

$$v_{nRTI} = v_{n-pre} + \frac{v_{n-ADC}}{A} \approx v_{n-pre} \tag{1.14}$$

if the preamplifier gain A is high enough. Then, if the preamplifier RTI noise v_{n-pre} is smaller enough than the sensor's noise, the degradation of the DR will be negligible:

$$DR = \frac{\Delta V_{FS-s}}{v_{n-s} + v_{n-pre}} \approx \frac{\Delta V_{FS-s}}{v_{n-s}} \tag{1.15}$$

Of course, the preamplifier gain cannot be made too high for range related matters: indeed, a condition to assert is that the preamplified full scale signal will fall in the ADC's input voltage ranges, otherwise the measurement would saturate before reaching the full scale values. This condition is expressed by the relation

$$A \cdot \Delta V_{FS-s} \leq V_{REF-ADC} \tag{1.16}$$

In order to exploit the full DR of the ADC (given by its ENOB), it is necessary that the output range of the amplifier matches as closely as possible with the ADC input range.

While one of the main reasons to use a In-Amp has been made clear, it is necessary to analyze in detail all its characteristics, the reasons that make them relevant and non idealities to be considered in order to keep errors low.

Gain First of all, gain must be precisely known, constant and uniform through all the input signal bandwidth. This condition is compulsory, given the need to guarantee the precision of the measurement. Since gain is a dimensionless quantity, usually the best way to make it precise is to design an architecture in which gain is function of only ratios of homogeneic quantities, like resistances or capacitances. This is done to prevent the gain to depend too much on process errors and temperature: while global errors are usually huge, they usually affect the components' parameters with similar relative errors, if proper precautions are taken during the layout design. So, matching errors, that is parameters variations between two components of the same kind, in the same chip, are extremely lower than global errors: furthermore, in most processes matching is best for passive components, and this justifies the use of resistors or capacitors to set the gain of In-Amps. Of course, if matching errors affect the gain precision more than what allowed by specifications, other countermeasures are needed: the most frequent one in such cases is laser trimming, that is a post-production process in which dimensions of components are individually modified by burning sections of them with a high power laser beam, in order to reach the precision needed. Of course, since this process must be made for each chip individually, it

implies a recurring cost that cannot be cut down with large scale production: laser trimmed chips are then far more expensive and are reserved for extremely high precision applications.

Differential input The use of differential approach has several advantages, often irreplaceable, that are making it more and more preferrable to a single-ended one in almost all fields of electronics. So, In-Amps have diffierential inputs not only to better interface all kinds of sensors, some of which have intrinsic differential output (like resistive sensors in a Wheatstone bridge configuration), but to take full advantages from the differential approach. In fact, while not requested by the definition of In-Amps, most of them in integrated applications are *fully differential*, meaning that they have both differential input and output.

A differential signal is the difference between two signals, both referred to ground. If these signals are v_p and v_n, we can define for the signal a differential mode and a common mode, as follows:

$$v_d = v_p - v_n \tag{1.17}$$

$$v_{cm} = \frac{v_p + v_n}{2} \tag{1.18}$$

In case of a fully differential amplifier (see fig 1.10), we can define the differential mode and common mode for both input and output ports.

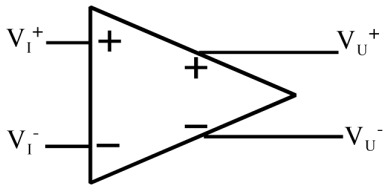


Figure 1.10: Input and output signals in a fully differential amplifier.

In such configuration, we can also define four different gains as ratios between different small signals:

$$A_{dd} = \frac{v_{d-out}}{v_{d-in}} \qquad A_{cd} = \frac{v_{d-out}}{v_{cm-in}} \tag{1.19}$$

$$A_{dc} = \frac{v_{cm-out}}{v_{d-in}} \qquad A_{cc} = \frac{v_{cm-out}}{v_{cm-in}} \tag{1.20}$$

Of course, the most relevant gain is A_{dd}, since it deals with the meaningful signals and it has to be extremely precise.

The strongest power of a differential approach is, beyond any doubt, disturbance rejection: if schematic and layout design is adequate, most interferences can be influencing only the common mode and thus be completely rejected by the differential mode, that is the meaningful signal. One of the most frequent interference that affect electronic circuits is caused by capacitive/inductive coupling of 50 Hz mains voltage: in biomedical applications - and it is just one of many examples - measurement instruments are tied through probes to the human body, that offers a huge parasitic capacitance with the building's electric net, thus cathcing a 50 Hz (60 Hz in America) large sine wave that is sensed by the measurement instruments. But if two differential probes are used, this disturbance would affect both of them almost in the same way, resulting in a common mode interference and leaving the differential mode unaltered.

Another problem, this time more frequent in PCBs, is non equipotential ground: when a cascade of components (or ICs) is given supply from one side, the supply current flowing through the parasitic resistance of ground wire causes a voltage drop between the ground pins of the different ICs, that thus work at different ground potentials. While this would be dangerous for a single ended signal, it is just a common mode disturbance for fully differential architectures.

Inside Silicon dies (but not only), in addition, a frequent issue is the so called *crosstalk*, that is capacitive coupling between metal lines close one to another: if a large voltage signal is in one line, it will pass into the other through the parasitic capacitance between them, resulting in a huge disturbance, especially if the second line is a high impedance node. With an accurate layout, a differential architecture can solve this problem trying to balance the parasitic capacitance between the first line and the other two: ideally, if they match (for example if the disturbing line is at the same distance from the other two), the disturbance will be the same in both lines and again will not affect the differential mode. All these advantages are achieved, of course, only if the common mode disturbance do not affect the output differential voltage: this is true if the following condition on the Common Mode Rejection Ratio (CMRR) applies.

$$CMRR = \left| \frac{A_{dd}}{A_{cd}} \right| \gg 1 \tag{1.21}$$

Furthermore, a full differential approach brings other relevant advantages that must be cited. First of all, it doubles the output swing: it is straitforward to understand that, if each signal could swing from V_{min} to V_{max}, the total output swing would be

$$\Delta V_d = V_{d-max} - V_{d-min} = (V_{max} - V_{min}) - (V_{min} - V_{max}) = 2(V_{max} - V_{min}) \tag{1.22}$$

This, keeping equal the noise, brings an increment by a factor 2 to the Dynamic Range, so the ENOB is increased by one.

Last but not least, a subtle improvement is related to distortion. Let for example be $v_{out-p} = f(v_{in-d})$ the transfer function between differential input and the positive output. The negative output will be characterized by an antisymmetric transfer function, that is $v_{out-n} = f(-v_{in-d})$. For a linear amplifier, it should be as linear as possible, but it will surely have second, third etc. order components, responsible for the degradation of the Total Harmonic Distortion (THD). If the output signal is differential, it will then be

$$v_{out-d} = f(v_{in-d}) - f(-v_{in-d}) \tag{1.23}$$

that is evidently an odd function, so it has no even order terms: since most of the THD is usually given by second order harmonics, this consists in a good improvement for linearity.

High CMRR We have said that, in order to exploit the benefits of a differential approach, it is necessary for an In-Amp to have a CMRR as high as possible. Usually, it is so high that it is expressed in dB, and it's quite common to find In-Amps with CMRR as high as 120 dB. It is a measure of how much the output depends on the differential input more than on common mode input. Having a high CMRR is not important only to reject disturbances: indeed some application require a high CMRR even if no disturbance affected the circuit. This is the case of applications in which the input common mode voltage changes during the measuring process: one example could be the *current sensing* use to monitor discharge of lithium batteries. The easiest way to monitor the charge level of a battery is, knowing the initial charge, to keep track of the charge erogated by measuring the current and integrating it with respect to time: this method is known as *Coulomb counting*. It is achieved by using a shunt resistor in series to the battery, making it possible to measure the voltage drop on it and so the current. Now, usually the shunt resistor is placed next to the hot pole rather than the cold one, to avoid the effect of non equipotential ground already mentioned. Since the output voltage of a battery gets lower as the battery discharges, the In-Amp will work with different common mode voltages and the measure should not be affected.

It must be noticed that, if a circuit is designed with perfect symmetry, there's no reason for a common mode input voltage to affect the differential output: in this case as well, are matching errors that cause a degradation of CMRR, and they can be kept low, as well as with a proper layout, with post-production solutions like laser trimming.

Input CM range Although some application set *a priori* the input CM that the In-Amp will be dealing with, it is often a good quality to have a wide input CM range, in order to keep flexibility. A general purpose In-Amp, for instance, should have a Rail-To- Rail (R2R) input CM range, so that it can work in the widest variety of applications. Current sensing, for example, requires the In-Amp to work with input CM voltages near the positive rail (V_{DD}); a Wheatstone bridge will give a CM around $V_{DD}/2$, while a microphone will usually give a $0V$ common mode.

Talking about flexibility - and not only about input CM range -, it must be noticed that even for non general purpose devices it could be a great quality. While flexible project can be sold in bigger quantities, with a decrease in the unitary price, they match with the concept of design reusability: in the context of a microelectronic firm, given the lenght and the cost of a complex analog design, a flexible device can be a solution for more than one project, thus decreasing the average cost and time-to-market of new products that can reuse old designs.

High input impedance As mentioned in the previous paragraphs, some sensors could have a really high output impedance, that makes them far from ideal voltage generators. So, an extremely precise gain is meaningless if the amplifier causes a huge loading effect on the sensor's output voltage: in order to avoid this, both differential input impedance and common mode input impedance (called isolation impedances) must be as high as possible.

High PSRR The Power Supply Rejection Ratio is a way to measure how much a variation on the supply voltage affects the output. It is defined as the ratio between a variation on the supply voltage ΔV_{DD} and the input voltade that would cause the same effect on the output, that is $\Delta V_{d-out}/A_{dd}$:

$$PSRR = \frac{\Delta V_{DD}}{\Delta V_{d-out}} A_{dd} \tag{1.24}$$

Obviously, this parameter must be as high as possible, in order to keep the output signal independent from supply disturbances or drifts: it is usual to express it in dB given the high values that it can reach.

The importance of this parameter is more evident in Mixed Signal circuits, where the presence of a clock signal causes spikes on the supply voltage. This is because, at every clock cycle, the digital part of the chip may require a huge amount of current from the supply rail, to charge all the gates. In these moments, the supply voltage can have a negative spike, due to the finite output resistance of the supply voltage generator, resulting in a disturbance that can be as high as hundreds of mV, in worst cases.

Low noise, offset and offset drift, bias currents The importance of having a low RTI noise has already been mentioned in this paragraph: it is crucial to mantain a good Dynamic Range on the signal. Togheter with noise, also offset and bias current can affect the measurement. See figure 1.11: an ideal In-Amp has been separeted from its non ideality generators.

V_{IO} is the offset voltage, that is the voltage that must be imposed to the input in order to have a zero output. I_{B1} and I_{B2} are the bias currents for each input. The difference $I_{IO} = I_{B1} - I_{B2}$ is called offset current.

In this configuration, the actual input voltage is:

$$V_{in} = (V_{S1} - V_{S2}) - V_n - R_S I_{IO} \tag{1.25}$$

if we suppose a balanced source (that is $R_{S1} = R_{S2} = R_S$).

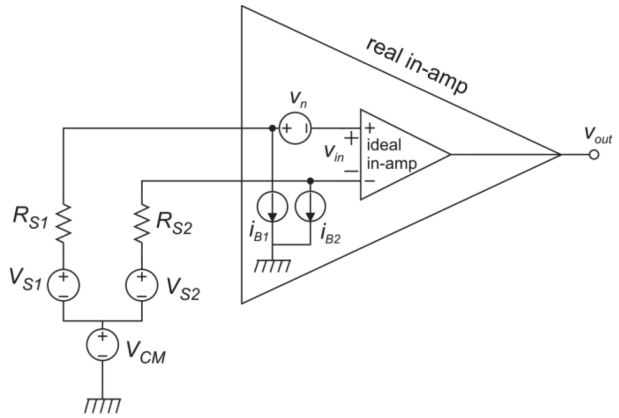


Figure 1.11: Real inamp interfacing a real voltage generator.

Bias currents are significant in JFET or BJT input amplifiers, while in MOSFET input devices they are often negligible (they are mainly leakage currents through the gate oxide); if input terminals are connected to the bonding pads, then there will be relevant bias currents caused by the protection diodes, but we they will be ignored for this purpose.

Noise voltage, as said, has a component given by thermal noise, a gaussian white noise, that is usually kept low by increasing the static current

consumption. There is another component, however, that is the so called Flicker noise, that has a peculiar Power Spectral Density proportional to 1/f (so it is dominant at low frequencies). This noise, caused by imperfections at the gate oxide-substrate border, is extremely high in MOSFET devices and can be reduced only making active areas large. This is one of the main reasons because analog circuits often occupy great amounts of area on Silicon dies. Then, offset voltage, again, is mainly caused by mismatch between internal components of the circuit. At first thought, since offset affects the measurement with a systematic offset, a calibration would be sufficient: however, offset strongly depends on temperature. So, even if the amplifier is calibrated, its drift caused by temperature will often counter the effect of calibration.

There are several architectural solution to reduce offset (and drift as well) and Flicker noise, without which such amplifiers would either be unusable for most applications or have impractical prices. The most diffused are Correlated Double Sampling (CDS, that works only for discrete time applications), Auto-Zero (AZ) and chopper modulation. The latter, briefly, ideally modulates (*chops*) the input signal with a square wave before the amplifier, that works on a high frequency signal to which offset and noise are summed. Then, the demodulator brings the signal back to baseband, while offset and low frequency noise are shifted to high frequencies, far from the signal bandwidth, where they can be filtered away. Chopper modulation will be discussed in detail in the following sections.

Low power consumption and NEF If on one hand a low noise design is needed for precision applications, on the other hand these systems are often destinated to mobile application, supplied by small batteries: this means that a particular attention must be paid to current consumption. As discussed in the following chapters, the quiescient current of an amplifier is directly linked to its thermal noise: in particular, if we can neglect flicker noise (rejected by dynamic techniques like chopper modulation), quiescent current is inversely proportional to the total noise of the amplifier.

The different topologies can have different *noise efficiency*, that is the ability certain values of noise with low currents. A figure of merit of an amplifier, then, can be defined as Noise Efficiency Factor (NEF), introduced in 1987 by Steyaert and Sansen [1]. Basically, this parameter is the ratio between the total input RMS (Root Main Square) noise of the amplifier and the equivalent input noise of a bipolar transistor with same bandwidth B and with a collector current equal to the quiescent supply current I_{DD} of the amplifier. Here is the original definition and expression, as proposed in

[1]:

$$NEF = \frac{v_{n,amp}}{v_{n,bjt}} = \frac{v_{n,amp}}{\sqrt{B\frac{\pi}{2}\frac{4kTV_T}{I_{DD}}}} \tag{1.26}$$

Filtering behaviour It is often important for a good Analog Front End to feature a filter, usualy a low pass or a band pass, depending on applications, to eliminate unwanted frequencies from the signal: they could be low frequencies up to DC components, like in the audio field, or frequencies higher than the signal band, where only undesired noise is present. This last case is necessary when interfacing an ADC: in fact, every time a signal is sampled, noise undergoes the so called *foldover* effect, that is its frequencies higher than sampling frequency are brought back into the Nyquist band where they accumulate, increasing the baseband noise floor. In order to avoid (or limit) this effect, an anti-alias filter is needed: it is basically a low pass filter with cutoff frequency at least 2 times lower than the sampling frequency (but in practical applications even 5 or more times is advised). For most cases, having a filter following the In-Amp is not a good choice, because of the high offset and noise of common filters, especially for fully integrated implementations. Sometimes, then, it is better to embed a filtering behaviour inside the In-Amp itself, without compromising the flatness of its frequency response within the signal band (required by gain precision specifications). Such architecture, for instance, is advantageous for chopper amplifiers, in which the so called *chopped offset* (that is offset modulated to high frequency) must be reduced as much as possible. In next paragraph, several architectures will be proposed to achieve this filtering behaviour of In-Amps.

Chapter 2

Instrumentation amplifiers: state of the art

In order to achieve the main characteristics of an In-Amp, introduced in paragraph 1.3, several topologies are possible and have been proposed in the literature. As said, given the importance of a fully differential approach, single-ended output solutions will be left out, unless they are particularly relevant from a historical point of view. In addition, we will consider implicitly CMOS implementations for the circuits, due to the diffusion and reduced costs of this technology. So, topology evaluation will always take into account the usual issues of CMOS technology, such as flicker noise, mismatch and precision component availability.

2.1 In-Amp architectures

Now, an overview of topologies is presented, trying to highlight pros and cons of the different solutions ([2], Chapter 1).

2.1.1 Three Op-Amp

This is the fully differential version of the well known single-ended Three Op-Amp topology, in which the third Op-Amp is used to convert a differential signal to a single-ended one. A simplified schematic is presented in figure 2.1.

This topology achieves a precise differential gain through voltage feedback, resulting in a ratio between resistors: that is $A_{dd} = 1 + 2R/R_G$, given that the Op-Amps gain is high enough. In this topology, the first two Op-Amps form the gain stage (resistor R_G can be external, so that gain is set

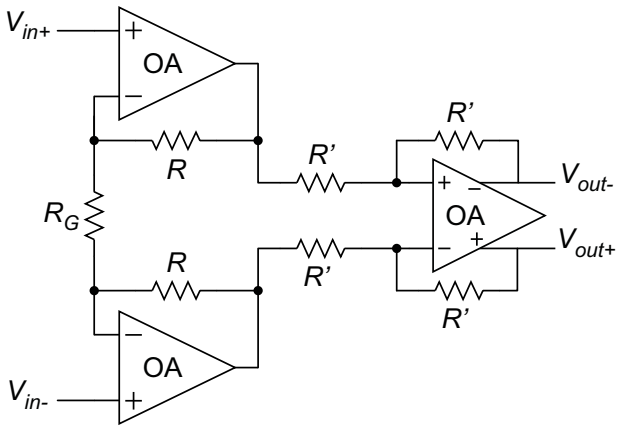


Figure 2.1: Three Op-Amp topology.

by the user), while the third Op-Amp buffers the output and can have a stabilized output common mode voltage.

Its main disadvantage is that the input common mode voltage is reproduced at the output of the gain stage: at this point, differential range may not be negligible (the signal has already been amplified), so if common mode voltage is near ground or the positive rail, the differential swing can be compromised. Furthermore, its CMRR depends strongly on matching between resistors. In addition, closed loop gain is affected by a relative error of $1/A$, so in order to keep it negligible three high gain (and, given a certain bandwidth, high Gain-Bandwidth product) Op-Amps are needed, thus degrading its power efficiency. However, it exhibits high input impedance and good linearity over a wide input and output range.

2.1.2 Switched Capacitor Topology

A different method to design an intrumentation amplifier is by means of the Switched Capacitor (SC) approach [3]. This kind of technique belongs to the Discrete Time (DT) domain: signals are sampled in precise moments, synchronous with a clock signal, and stored into capacitors. A set of switches (digital MOSFETs in most applications) then moves electric charge from one capacitor to another at every clock edge, in order to achieve amplification. In figure 2.2 is shown a simple example of SC In-Amp.

This circuit's operation is divided in two phases. During phase 1 (switches with Φ_1 closed, others open), input signal is sampled and stored into capacitors C_1, while the Op-Amp is in reset state. During phase 2 (Φ_2), charge flows from C_1 to C_2, setting the output voltage. Differential gain is then C_2/C_1, that is a simple ratio between capacitances: with accurate layout,

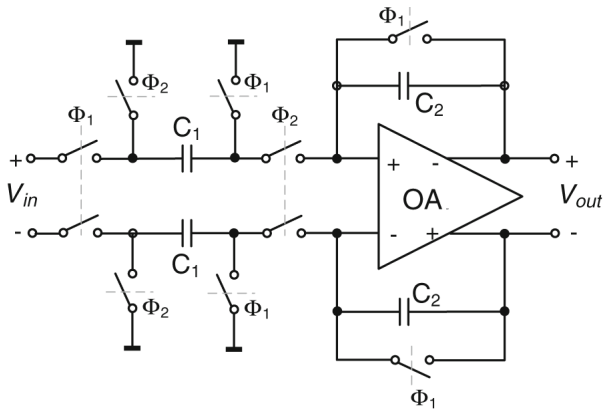


Figure 2.2: SC topology.

gain errors of 0.1% are achievable. The strength of this topology is, besides simplicity and area efficiency, a natural Rail-To-Rail input CM range, because capacitors block DC voltages, and high CMRR if accurate matching between capacitors is obtained (usually they have excellent matching parameters). On the other hand, capacitors used for sampling suffer from kT/C noise, that requires large areas to be kept low. Furthermore, while this technique naturally implements Correlating Double Sampling (see section 2.2.3) and so eliminates DC offset and flicker noise, it suffers from noise foldover since it is a discrete time topology and needs sampling. Finally, a high input impedance is often difficult to achieve: because of the switching behaviour, capacitors must be periodically charged by the source, resulting in an equivalent input resistance of

$$R_{in} = \frac{1}{2C_1 f_s} \tag{2.1}$$

This can be increased with small capacitors (that however cause a high kT/C noise) or with low sampling frequency, although this last solution might be impractical due to Nyquist band specifications.

2.1.3 Capacitively-Coupled Topology

Another topology that uses capacitors as feedback element is shown in figure 2.3 [4]. Unlike SC, this approach belongs to the Continuous-Time domain: signals undergo chopper modulation (their polarity is periodically inverted by switches: see section 2.2.5 for better explanation) in order to flow through capacitors and to eliminate low frequency errors (offset and flicker noise).

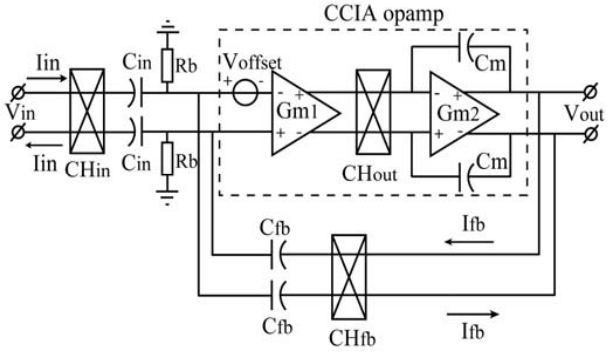


Figure 2.3: Capacitively-Coupled topology.

2.1.4 Current-Mode Topology

This continuous time topology uses two Op-Amps to obtain a high input resistance; through virtual short cirtuit, it sets a copy of the input voltage across a precision resistor, whose current is accurately mirrored and sent to a second resistor. The resulting voltage is then buffered and presented to the output, with a precise differential gain resulting equal to the ratio between the resistances. Figure 2.4 shows a block diagram of this topology.

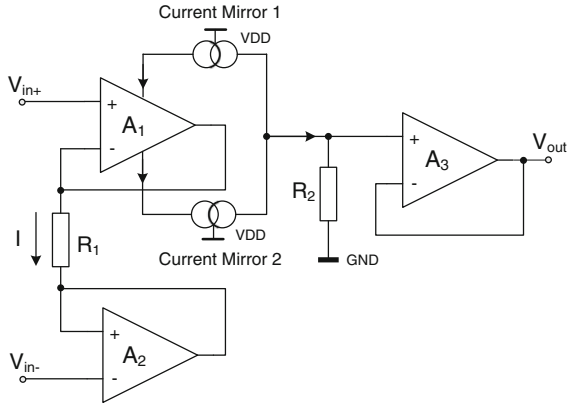


Figure 2.4: Current-Mode topology.

The CMRR of this topology depends on the matching of the current mirrors and the DC precision of the current mirrors is essential for the overall offset, gain accuracy, drift and linearity. Since the matching of an impedance-boosted current mirror can still be insufficient for the required DC precision, thin-film resistor-degenerated current mirrors can be used in order to achieve CMRR as large as 120 dB [5]. The use of Op-Amps as buffers, instead of amplifying elements, extends the input common mode

range, that though still doesn't include supply rails (unless the two input Op-Amps have true Rail-To-Rail output).

2.1.5 Current Feedback Topology

Current Feedback Instrumentation Amplifiers (CFIAs) are surely the most diffused in the field of sensor readout, thanks to their several advantages and excellent performances. Figure 2.5 shows a typical topology for a CFIA.

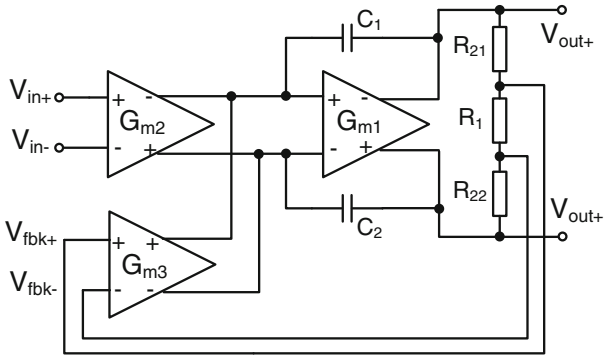


Figure 2.5: Current Feedback topology.

Two transonductors, G_{m2} and G_{m3}, transform respectively the input voltage and a fraction $\beta = \frac{R_1}{R_1+R_{21}+R_{22}}$ of the output voltage into differential currents, that are subtracted one from the other. G_{m1}, together with capacitors C_1 and C_2, form a Miller compensated integrator, whose high DC gain nulls the difference between the currents, thanks to negative feedback. Thus, if $I_{out2} = I_{out3}$, then it's simple to derive an expression for the differential gain:

$$A_{DD} = \frac{G_{m2}}{G_{m3}} \frac{1}{\beta} \tag{2.2}$$

This topology can achieve great CMRR, thanks to the input transconductors that convert input voltage into a current. Furthermore, transconductors are usually designed with differential pairs, whose input common mode range can easily include the positive rail (n pairs), the ground rail (p pairs) or both if both kinds of pairs are combined together. In addition, compared to the SC topology, it doesn't suffer from noise foldover, and its input impedance can be extremely high, since its input capacitances don't need to be periodically charged.

Actually, this topology, like all others when implemented in CMOS technology, is affected by huge offset and $1/f$ noise, if proper precautions are

not taken: almost in every precision application, thus, dynamic techniques for low frequency errors reduction will be implemented. If chopper modulation, as described in paragraph 2.2.5, does not cause noise foldover, in any case it lowers the input resistance. Next paragraph will describe in detail these issues.

Beyond that, the main problem of this architecture is gain accuracy: as seen in 2.2, differential gain is proportional to a ratio between the input and feedback transconductances (and, of course, a ratio between resistances that must be precise). It is then crucial to pay attention during the design and layout process, in order to make the two tranconductors match as closely as possible.

The first CFIA was introduced by Analog Devices [6] in 1971, and was implemented in bipolar technology, while the first CFIA implemented in CMOS technology was presented in 1987 by Steyaert for medical applications [1].

2.2 Dynamic techniques for low frequency errors reduction

As said, on one hand CMOS technology shows many advantages: first of all, it is much less expensive, as cost per unit of Silicon area, than bipolar or BCD technologies, mainly because often relatively old technological nodes are sufficient (analog transistors are usually large), so yields are high. Furthermore, since CMOS was initially thought for digital applications, it naturally allows integration between analog and digital blocks within the same chip: this is a reason for the increasing diffusion of Analog-Mixed Signal (AMS) chips or even Systems on a Chip (SoC) that may include MEMS structures.

On the other hand, analog CMOS transistors are particularly affected by flicker noise and mismatch, that in common topologies can cause high offset voltages. Flicker noise and offset are also called *low frequency errors*.

2.2.1 Offset voltage

Offset voltage is defined as the input differential voltage to be imposed in order to null output voltage. It can be considered the DC component of RTI noise. If a fully differential topology is designed with perfect symmetry of its components, there shoud be no reason for output to be different from zero if its input is zero. However, local variations (i.e. mismatch) of MOSFET parameters, like V_{th} (threshold voltage) and β, or passive components

parameters (mainly capacitances and resistances) can alter the symmetry and produce a finite and non null output when the input is zero. The following equations show an expression for the statistical variations of V_{th} and β:

$$\sigma_{V_{th}} = \frac{C_{V_{th}}}{\sqrt{wl}} \tag{2.3}$$

$$\sigma_{\frac{\Delta\beta}{\beta}} = \frac{C_{\frac{\Delta\beta}{\beta}}}{\sqrt{wl}} \tag{2.4}$$

where σ is the standard deviation of the two gaussian distributions, $C_{V_{th}}$ and $C_{\frac{\Delta\beta}{\beta}}$ are process parameters and w and l are MOSFET channel dimensions.

As seen from equations above, mismatch can be reduced with large areas, since it is mainly caused by a different concentration of dopant or imperfections: the larger the areas, the higher the probability that two trasistors have the same amount of them. Furthermore, if dopant concentration follows a cross-chip gradient, layout tricks can be used, like *common centroid*: two transistors are placed on the Silicon durface with central symmetry, so every possible gradient is compensated.

However, even with these precautions, CMOS circuits are usually affected by hughe offset voltages, that can reach tens of mV in worst cases. Usually sensitivities of a few µV are requested, so dynamic techniques are needed.

2.2.2 Flicker noise or $1/f$ noise

When carriers flow through a MOSFET channel in proximity of the substrate-gate oxide border, they can be trapped by impurities spread all over the oxide near the surface, and released after a random interval of time. This strongly depends on the purity of the gate oxide, so it can be very different from process to process. Anyhow, it can be demonstrated that the consequent drain current noise (or, similarly, the voltage noise referred to input v_{gs}) has a Power Spectral Density (PSD) proportional to $1/f$. A simple model for current Flicker noise PSD S_{In-F} can be described with the following equation [7]:

$$S_{In-F}(f) = \frac{N_f}{w_{eff} l_{eff}} g_m^2 \frac{1}{f} \tag{2.5}$$

where w_{eff} and l_{eff} are the effective width and length of the MOSFET

channel, f is frequency, g_m is the MOSFET tranconductance and N_f is a parameter (constant at first approximation) strongly dependant on the process.

Flicker noise is added to the thermal noise PSD that is constant with frequency (it is also called *wide band noise*): an expression for current thermal noise S_{In-Th} is shown for comparison.

$$S_{In-Th} = \frac{8}{3} kT g_m (1 + m) \tag{2.6}$$

where $k = 1.38 \times 10^{-23}$ J/K is the Kelvin constant, T is absolute temperature, g_m is the MOSFET tranconductance and $m = \frac{g_{mb}}{g_m} \approx 0.3$.

Evidently, flicker noise is dominant at low frequencies, while thermal noise is dominant at high frequencies, up to the circuit's band limit. These two regions are separated by a *corner frequency* f_k such that

$$S_{n-F}(f_k) = S_{n-Th} \tag{2.7}$$

Figure 2.6 shows a typical RTI noise PSD for an analog CMOS amplifier.

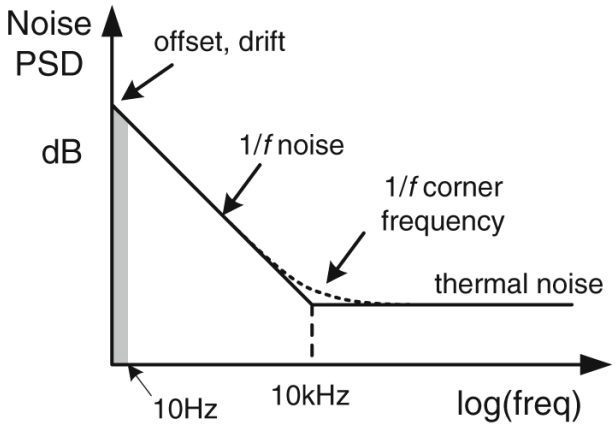


Figure 2.6: Typical RTI noise PSD.

Now, even for flicker noise a solution is to use large areas for transistors, but chip area is expensive and it could not be enough for precision applications. In fact, in most cases f_k can be as high as a few kHz, so fully inside the signal band: since it is usually intolerable, dynamic techniques must be taken into account in order to reduce it.

Next paragraphs will give an overall description of the three most diffused techniques, that are Correlated Double Sampling (CDS), Auto-Zeroing (AZ) and Chopper Stabilization (CHS) [8].

2.2.3 Correlated Double Sampling

CDS is a typical and natural technique for discrete time circuits, that is circuits that use *sampling* in signal elaboration. The most relevant examples are for sure Switched Capacitor (SC) topologies, already mentioned in paragraph 2.1.2. In these circuits, the output is valid only at the end of one of the two (or more) phases, but they can use the other phase to get information about offset and noise. Let us consider, for instance, the circuit in fig 2.7: it is a simple, single-ended, SC inverting amplifier based on Op-Amp with negative feedback.

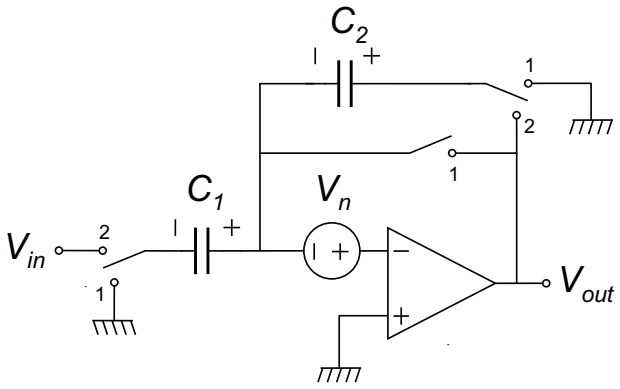


Figure 2.7: CDS in a SC amplifier.

At the end of phase 1, these voltages are present (voltages across capacitors are intended according to polarity):

$$V_{out}^{(1)} = -V_n^{(1)} \tag{2.8}$$

$$V_{C_1}^{(1)} = -V_n^{(1)} \tag{2.9}$$

$$V_{C_2}^{(1)} = V_n^{(1)} \tag{2.10}$$

At the clock edge, neglecting kT/C noise and charge injection, voltages change and a charge ΔQ flows from C_2 to C_1, according to polarity:

$$V_{C_1}^{(2)} = -V_n^{(2)} - V_{in} \tag{2.11}$$

$$\Delta Q = C_1(V_{C_1}^{(2)} - V_{C_1}^{(1)}) = C_1(-V_n^{(2)} - V_{in} + V_n^{(1)}) \tag{2.12}$$

$$V_{C_2}^{(2)} = V_{C_2}^{(1)} + \frac{\Delta Q}{C_2} = V_n^{(1)} + \frac{C_1}{C_2}(-V_n^{(2)} - V_{in} + V_n^{(1)}) \tag{2.13}$$

Finally, at the end of phase 2 (when data are valid), output voltage turns

out to be

$$V_{out}^{(2)} = -V_n^{(2)} + V_{C_2}^{(2)} = \underbrace{-\frac{C_1}{C_2}V_{in}}_{signal} \underbrace{-\left(1+\frac{C_1}{C_2}\right)\left[V_n^{(2)} - V_n^{(1)}\right]}_{CDS\ noise/offset} \tag{2.14}$$

It is evident that the offset is completely removed from the output, since it is constant during both phases. Similarly, those noise components that don't change too much from phase 1 to phase 2, that is the component with high *correlation* (that is low frequency components) are significantly reduced. On the other hand, components with strong negative correlation will be enhanced, but if the sampling frequency f_s is high enough, these component will involve thermal noise only.

Without further calculation, it can be proved that:

1. If sampling frequency is higher than the flicker corner frequency ($f_s > f_k$), flicker noise will be completely removed, leaving thermal noise only that can be decreased by increasing static current until needed.

2. Since noise is sampled, its frequencies higher than f_s (that are only from thermal noise S_{n-Th}) are shifted down to the Nyquist band according to the *foldover* phaenomenon, so thermal noise within the band is increased. In these conditions, it can be found that total noise PSD is

$$S_{n-tot} \approx 4\frac{B_n}{f_s}S_{n-Th} \tag{2.15}$$

   where B_n is the noise bandwidth. This is usually quite larger than f_s: it is indeed equal to the amplifier's band, that must be a few times (usually 5 or 6) larger than f_s in order to reach steady state between two successive clock transitions.

2.2.4 Auto-Zeroing

Auto-Zeroing is another technique that, like CDS, periodically samples noise, stores it and subtracts it from the output signal, but it's thought for continuous time circuits. A typical operation is divided again in two phases (figure 2.8):

1. Auto-Zeroing phase: during this phase input signal is disconnected, and noise is sampled and stored into a memory component (typically

capacitors). This phase should be as quick as possible, since during it the output is not valid: for a real continuous time application, a sample and hold (S/H) circuitry can be implemented at the output stage, but still the information could be degraded if this phase is too long.
There are two different approaches, depending whether the noise/offset is stored at the amplifier's input or output, called Input Offset Storage or Output Offset Storage. The latter, however, can lead to saturation of the output if the offset is not so low and the amplifier's gain is high: in these cases, the former approach is to be preferred.

2. Normal Operation phase: during this phase input signal is connected back to the amplifier, to be presented at the output amplified and with the sampled noise subctracted from it. The effective RTI noise, function of time, will then have an expression like

$$v_{n,eff}(t) = v_n(t) - v_n(kT_s) \tag{2.16}$$

where $v_n(kT_s)$ is the last noise sample stored, and it is constant through all the clock period T_s.

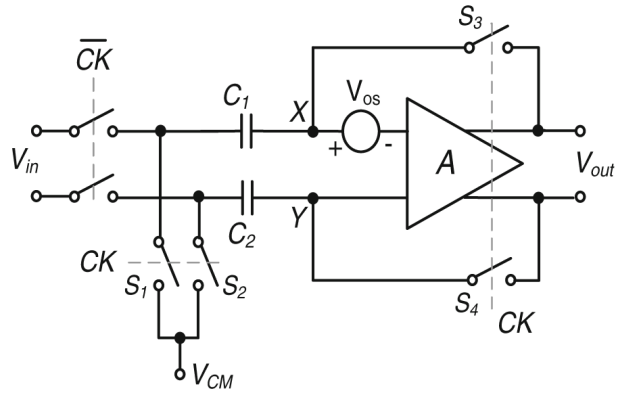


Figure 2.8: Auto-Zeroing with input offset storage.

It is evident, in this case as well, that offset voltage is completely removed, since it is not function of the time, while low frequency components of noise (i.e., flicker noise) are strongly reduced since they do not vary significantly within each clock period. Like CDS, it can be proved that if $f_s > f_k$, flicker noise is completely removed, but high frequency noise (thermal noise) will undergo foldover and will increase the total noise in the Nyquist band, since it is sampled. In this case, however, a mathematical explanation is

more complex, since it involves a continuous time signal and a sampled one: for a detailed analysis of noise in Auto-Zeroing, see [8].

Here, the most relevant results are reported. First of all, noise foldover affects only the Nyquist band (that is for $f < f_s$), while for $f > f_s$ noise will remain unalterated and equal to S_{n-Th}. A typical shape of AZ noise PSD is shown in figure 2.9

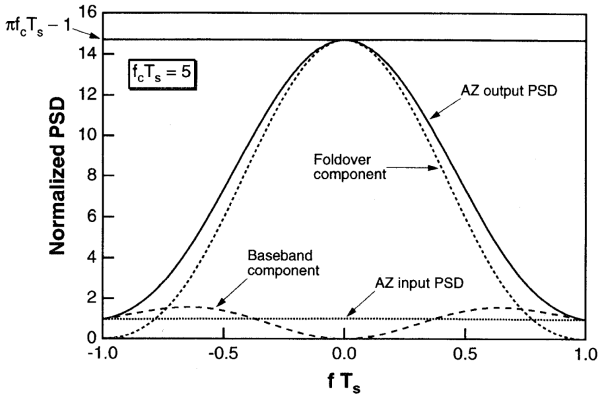


Figure 2.9: Noise in Auto-Zeroing.

For $f < f_s$ total noise will roughly be

$$S_{n-tot} \approx 2\frac{B_n}{f_s} S_{n-Th} \tag{2.17}$$

provided that flicker noise is completely removed. Compared to CDS, it shows a factor of 2 instead of 4; on the other hand, since the amplifier must sample noise in an interval of time as short as possible, it will need to have a band $B = B_n \gg f_s$, thus causing several copies of the noise to be folded into Nyquist band. A drastic solution to this issue is presented with the so called Ping-Pong configuration [9]. The whole circuit is doubled: when one of them is in Normal Operation, the other is auto-zeroing and vice versa. This time, since output is always valid, the two phases can be equally long, thus relaxing the speed specification of the amplifier (B_n comparable with the one in CDS). This way, noise results are better than in CDS and output is always valid (it is a real continuous time system), but doubling the whole system means doubling chip area and current consumption.

2.2.5 Chopper Stabilization

Unlike CDS and AZ, Chopper Stabilization is a continuous time modulation that does not cause noise foldover, resulting in a much more effective noise

cancellation approach: this is the main reason for its large diffusion in precision amplifiers. A block schematic of a chopper amplifier is shown in figure 2.10.

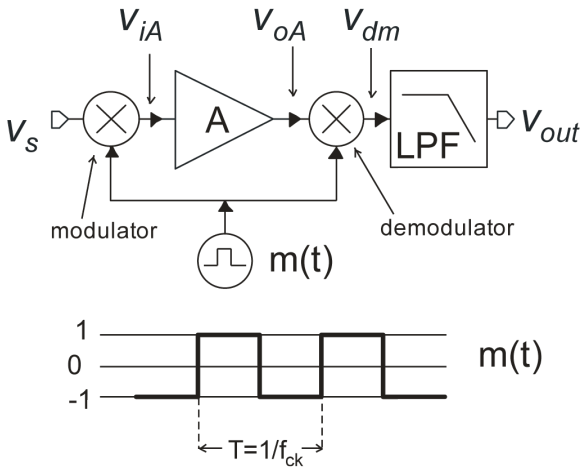


Figure 2.10: CHS block schematic.

Basically, a dimensionless square wave $m(t)$ with frequency f_{ch} and duty cicle $\delta = 0.5$ is used to modulate the input signal: its spectrum is thus shifted around f_{ch} and its multiples. The modulated signal is then amplified, together with offset and noise. The second modulator, finally, shifts the signal back to baseband, while offset and flicker noise are shifted to high frequency, where they can be filtered away by the Low Pass Filter (LPF). For a precise analysis of the effect of modulation, a frequency domain representation is helpful ([7], chapter 2.3).

Thanks to the Fourier tranform, we can express the square wave with a sum of harmonics:

$$m(t) = \sum_{k=-\infty}^{\infty} c_k e^{j2\pi f_{ch} t}, \quad with \quad c_k = \begin{cases} 0 & : k \ even \\ \frac{2}{\pi k} & : k \ odd \end{cases} \tag{2.18}$$

This signal then modulates the input signal spectrum $V_s(f)$ to frequencies kf_{ch}, with k odd, each multiplied by a factor c_k. Modulated signal V_{iA} can be expressed as

$$V_{iA}(f) = \sum_{k=-\infty}^{\infty} c_k V_s(f - kf_{ch}) \tag{2.19}$$

This signal is then amplified: supposing that the amplifier, as a general case, has a limited bandwidth B_a, it will amplify only the replicas at $kf_{ch} <$

B, while nulling the others. The demodulator, finally, will shift back non-null replicas, each multiplied again for $c_{-k} = c_k^*$, and the LPF will eliminate high frequency components. The output signal spectrum will then be

$$V_{out}(f) = V_s(f)A \underbrace{\sum_{k=-N}^{N} \|c_k\|^2}_{\alpha} = \alpha A V_s(f) \tag{2.20}$$

where A is the amplifier gain and $N = B/f_{ch}$ is roughly the numer of amplified replicas. If the amplifier had infinite bandwidth, N would be infinite and α, that is the sum, would actually be the square wave power, that is 1, and the signal would be perfectly recovered. Actually, since the power of a square wave is mostly contained in the first harmonics, even a limited bandwidth is often sufficient to let almost all the power pass, and the factor α will be not significantly lower than 1 (i.e. for $N = 10$, $\alpha \approx 0.95$).

The use of a sine wave, instead of a square wave, would completely remove this problem, but it is impractical since linear modulators like Gilbert cells usually introduce a large amount of noise/offset, and would be devastating with non-preamplified signals. On the other hand, a square wave is achievable by means of simple switches, that introduce minimal offset/noise, are power efficient and can easily make a modulator in a fully differential architechture, like shown in figure 2.11.

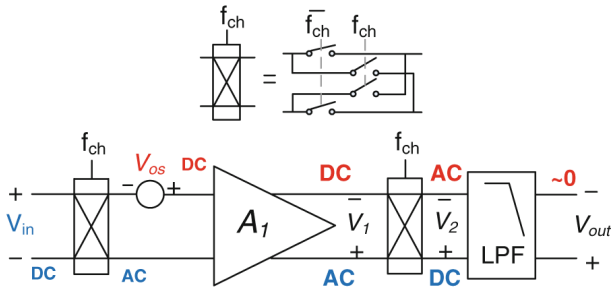


Figure 2.11: Chopper modulator implemented with switches.

Now an analysis of how chopper modulation affects offset and noise is presented. The whole noise PSD, after amplification by a factor A^2 is shifted by the second modulator to kf_{ch}, producing an infinite numer of replicas, each multiplied by a factor $\|c_k\|^2$ (thus they are null for k even). The LPF then keeps only the components that fall in the baseband, as shown in figure 2.12.

Supposed that $f_{ch} > f_k$ and that LPF cutoff frequency f_c is significantly lower than f_{ch}, the total output noise PSD will then be

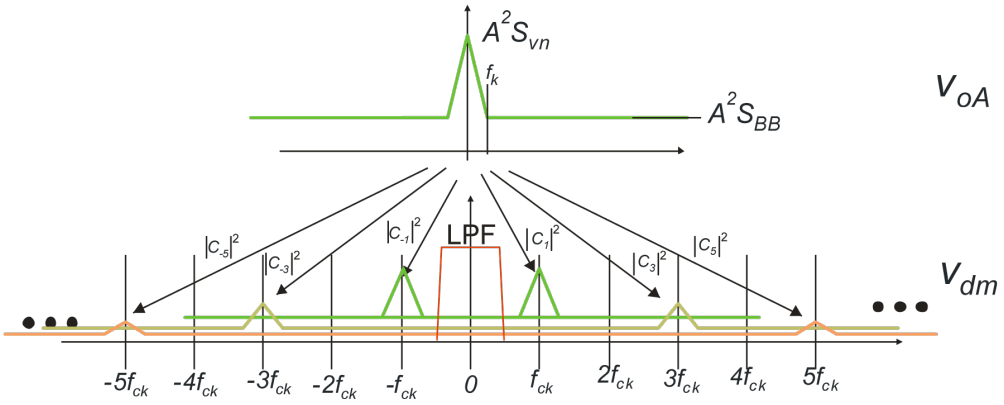


Figure 2.12: Noise in chopper modulation.

$$S_{n-out} = A^2 S_{n-Th} \sum_{k=-N}^{N} \|c_k\|^2 = \alpha A^2 S_{n-Th} \tag{2.21}$$

Thus, given that the effective signal amplification is αA, the equivalent total RTI noise PSD will be

$$S_{n-eff} = \frac{S_{n-out}}{(\alpha A)^2} = \frac{S_{n-Th}}{\alpha} \tag{2.22}$$

In a similar way, offset is amplified, modulated and filtered. It's important to notice that, like in Output Offset Storage AZ, the amplifier's gain cannot be too high: otherwise the output modulated offset, even if filtered, could saturate the amplifier output and prevent a proper amplification of the signal. This granted, at the LPF output offset will appear as a zero average (so with ideally no residual DC offset) square wave (for this it is also called *Chopped Offset* or *offset ripple*) at frequency f_{ch} and with an amplitude of

$$V_{out,CO} = AV_{io}H(f_{ch}) \tag{2.23}$$

where $H(f_{ch})$ is the attenuation of LPF at f_{ch} and V_{io} is the input offset voltage. Other techniques to reduce chopped offset will be presented in next paragraph.

Two main non-idealities must be taken into account, since they could cause a residual offset. First of all, a clock duty cycle different from 50%: such waveform, in fact, while not affecting the signal elaboration, produces an output chopped offset with non null average value, resulting in a residual offset. Suppose a modulating wave $m(t)$ with duty cycle $\delta = 0.5 + \epsilon$: the

residual output offset will then be

$$< AV_{OS}m(t) >= \frac{AV_{OS}(\delta+\epsilon)T_{ch} - AV_{OS}(\delta-\epsilon)T_{ch}}{T_{ch}} = 2AV_{OS}\epsilon \quad (2.24)$$

A solution to this problem can be producing a square wave with precise frequency, even with inaccurate duty cycle, at $2f_{ch}$, then using a digital divider like a T-Flip Flop. A second problem that causes residual offset is *charge injection*. When MOSFET switches are turned off, charges accumulated in the channel must flow out, and they usually gather in the source and drain capacitances, causing voltage spikes in certain nodes of the signal path. These spikes, amplified and demodulated, can result in a residual offset: again, a complete and accurate analysis can be found in [8].

2.3 CFIA implementations and error reduction

As said, Current Feedback Instrumentation Amplifier is the most suitable topology for readout of a great variety of sensors. Still, its main probems to be taken into account are low frequency errors and gain error.

Low frequency error, that is offset and flicker noise, is most commonly reduced with Chopper Stabilization technique, thanks to its continuous time nature that avoids noise foldover and thus makes it very power efficient. There are several examples in literature of chopped CFIAs: however, they can use different strategies in order to deal with chopped offset. Next paragraph will present an overwiev of relevant proposed solutions.

2.3.1 Chopped offset reduction

As said, at the output of the chopped amplifier a null-mean square wave with AV_{io} will be present and it must be eliminated. Two approaches can be distinguished: the first one is based simply on a linear low pass filtering, that will be called *static offset reduction*. The second one exploits a second modulation to shift chopped offset back to baseband and compensate it at the input of the amplifier: this can be called *dynamic offset reduction*.

Static offset reduction This approach is the oldest and ideally the simplest, since it entrust entirely the LPF behaviour of the amplifier to eliminate chopped offset. Since external LPF are not recommended for noise/off-

ste issues, usually filtering is embedded in the CFIA topology. One of the most simple solutions is based on Miller integrators[10]: in this implementation, shown in figure 2.13, the first two of the three stages are chopped and nested Miller compensation is used to filter the chopped offset, as well as to assure stability of the feedback loop.

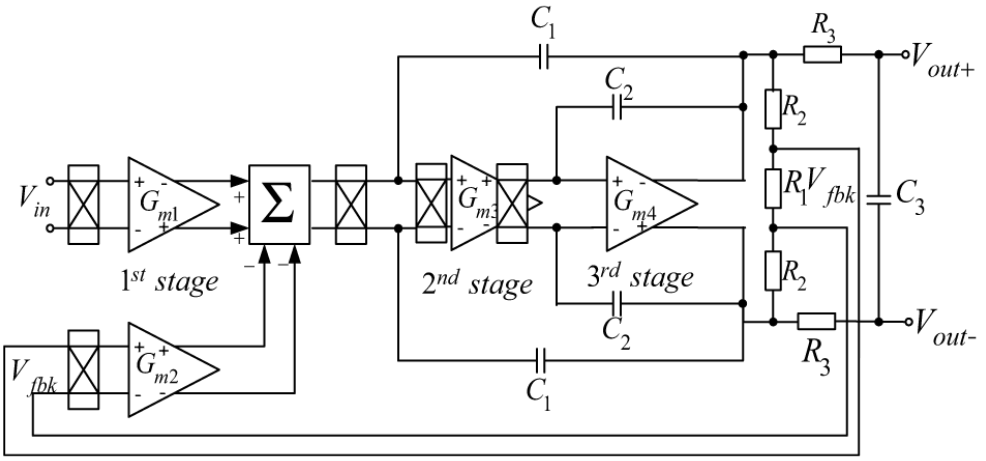


Figure 2.13: Miller integrators as LPF for a CFIA.

Another way to implement a LPF filter is the so called State Variable Filter (SVF), that uses a cascade of integrators whose input is the difference between the previous stage's output and a partition of the amplifier's output (see figure 2.14). A second order State Variable LPF is proposed in [11].

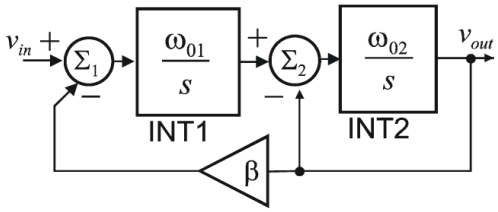


Figure 2.14: State Variable Filter block diagram.

In this implementation, the first integrator is a $G_m - C$ configuration, while the second one is a G_m-Op-Amp in order to have a low output resistance to drive the resistive feedback network. Figure 2.15 shows a more detailed block diagram for this topology.

Other filtering techniques have been proposed: a Switched Capacitor Notch Filter [12], a continuous time notch filter in a multipath architechture [13], or a notch filter integrated in an Auto Correction Loop [14]. These

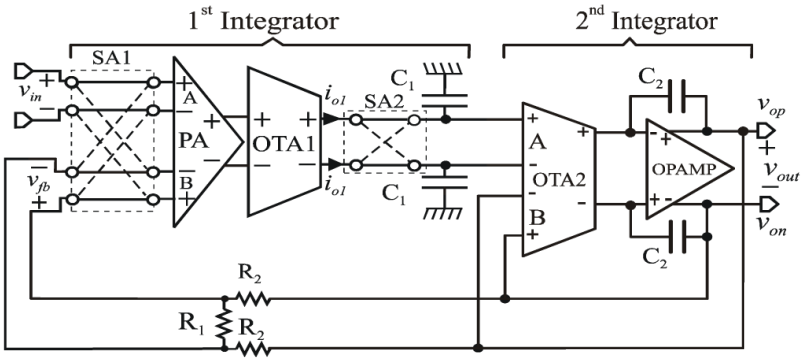


Figure 2.15: State Variable Filter as LPF for a CFIA.

architectures, however, are quite far from the purpose of this thesis and will not be further analyzed.

Dynamic offset reduction A different approach is presented to reduce output chopped offset, in case the attenuation of LPF alone is not sufficient or it would need too large capacitors. A three stage Miller-integrator-based architecture features a Ripple Reduction Loop (RRL - elsewhere called ORL or Offset Reduction Loop), whose operation is explained in figure 2.16 [15].

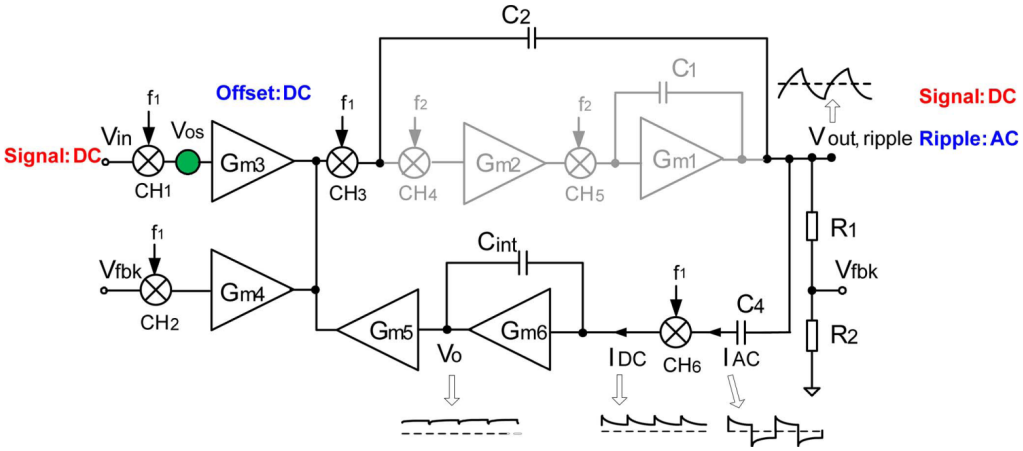


Figure 2.16: Ripple Reduction Loop in a CFIA.

Capacitor C_4 senses output ripple and transforms it into an AC current, that is demodulated and integrated into a voltage through C_{int}. This voltage is then converted into a current to be summed to the output current of the input transconductors, resulting in an offset compensation action.

2.3.2 Gain Error Reduction

As said in paragraph 2.1.5, another relevant issue in Current Feedback topologies is gain error. When open loop gain is high enough, that is if current virtual short circuit can be considered valid, closed loop gain can be expressed as

$$A_{DD} = \frac{G_{mi}}{G_{mf}} \frac{1}{\beta} \tag{2.25}$$

where G_{mi} and G_{mf} are the input and feedback OTA tranconductances, and β is a ratio between resistances. Thus, mostly three causes can affect the precision of the gain:

1. A mismatch between resistors can significantly alter the coefficient β, if proper precautions are not taken. In particular, layout techniques for matching can be considered, like increasing the area of resistors (while keeping constant proportions) or placing them in a common centroid configuration. If extremely high precision is needed, laser trimming can be the definitive solution, although it is a really expensive process step.

2. A mismatch between tranconductances can be a more difficult problem to be solved with traditional techniques: in fact, while large areas and common centroid configuration of the active MOSFETs can be of help, usually mismatch between active components is worse than between passive ones, and trimming is not a choice for them unless using resistor degenerated pairs [16]. Thus, dynamic techniques can be used, like Dynamic Element Matching (DEM), as described in the following paragraph.

3. Usually OTAs are based on differential pairs, biased by a current mirror with finite output resistance. Thus, a variation in the input common mode will be reflected into a variation in the mirror output voltage, that is a variation in the bias current that obviously will alter the tranconductor's G_m. As a consequence, if input and feedback voltages present different common modes, there will be a difference between the two G_m, resulting in gain error. While DEM is uneffective in this case, a Common Mode EQualization (CMEQ) technique can be used to solve this problem.

Dynamic Element Matching is a method to improve the matching between two elements (e.g. two resistors) by continuously swapping them in

order to balance their values, at least on an average basis. Here it can be used to dynamically reduce the effect of G_m mismatch on the gain precision, achieved by periodically swapping the input and the feedback transconductor in order to reduce the dependance of gain from G_m mismatch [17]. Since [11] the first stage formed by the two input OTAs can be regarded as a Difference Differential Amplifier (DDA, used as preamplifier) with two input ports, DEM is also called Port Swapping (PS).

Suppose that the mismatch between the two OTAs is expressed as:

$$G_{mf} = G_{mi}(1 + \Delta) \tag{2.26}$$

where G_{mi} and G_{mf} are the input and feedback OTA tranconductances, and Δ is the relative error. Without DEM, this would result in a gain equal to

$$A_{DD} = \frac{G_{mi}}{G_{mf}}\frac{1}{\beta} = \frac{1}{1+\Delta}\frac{1}{\beta} \approx (1-\Delta)\frac{1}{\beta} \tag{2.27}$$

so presenting a relative error exactly equal to Δ. Usual values for G_m mismatch can be as high as 2%, that is an unacceptable gain error for most applications.

When DEM is active, the amplifier works in one phase with $A_{DD}^{(1)} = \frac{1}{1+\Delta}\frac{1}{\beta}$, in the other with $A_{DD}^{(2)} = \frac{1+\Delta}{1}\frac{1}{\beta}$, giving rise to an output ripple as shown in figure 2.17.

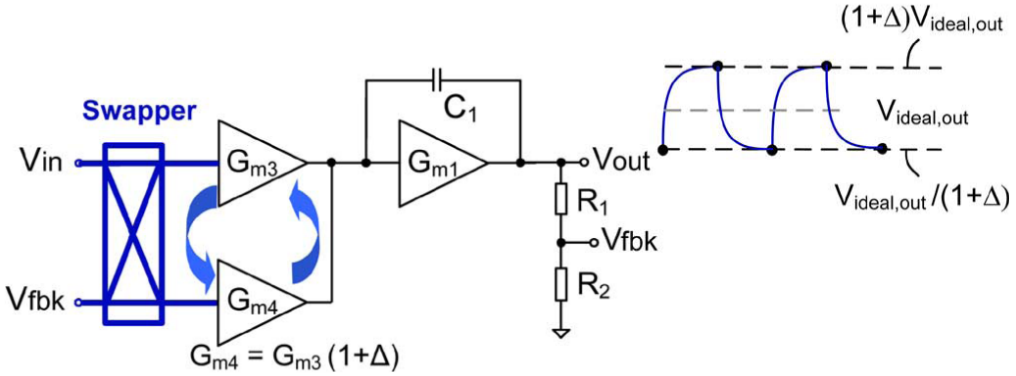


Figure 2.17: Dynamyc Element Matching in a CFIA.

Since this ripple is at high frequency (can be the same frequency of chopper clock), it is filtered away by the LPF: the effective gain is then the average between the two. The resultant gain error will then be

$$\frac{< A_{DD} >}{\frac{1}{\beta}} = \frac{\frac{1}{1+\Delta} + \frac{1+\Delta}{1}}{2} - 1 \approx \frac{(1-\Delta+\Delta^2)+(1+\Delta)}{2} - 1 = \frac{\Delta^2}{2} \tag{2.28}$$

In conclusion, a 2% mismatch between G_m would be just a 0.02% gain error, acceptable for most applications, granted that the DEM ripple is sufficiently neglectable after filtering.

As a side effect, it is relevant to notice that DEM has been proved to increase the input impedance [18], that usually high frequency chopping decreases.

Anyway, like for offset ripple, a static approach for reducing DEM ripple (i.e. simple low pass filtering) may not be sufficient or convenient: so, a dynamic ripple reduction technique can be implemented [17], called Gain Error Reduction Loop (GERL). Its operation is described in figure 2.18.

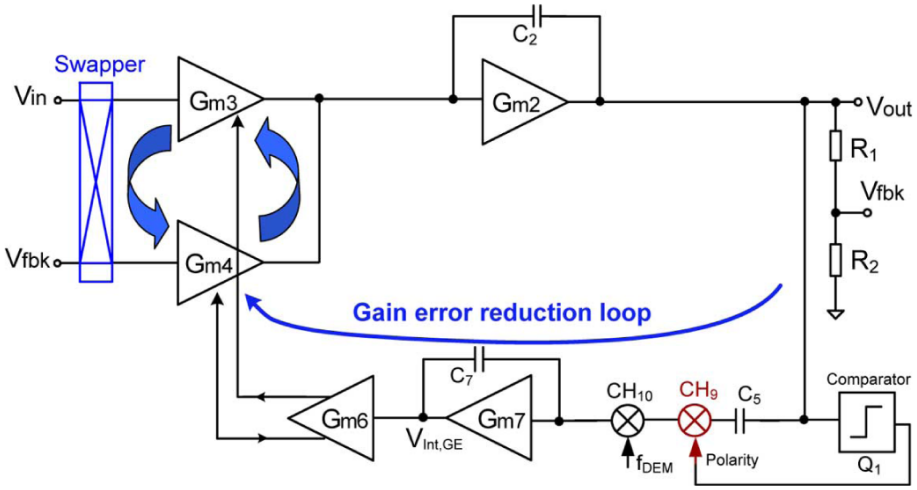


Figure 2.18: Gain Error Reduction Loop in a CFIA.

It works exactly like ORL, with two main differences. The first one, the GERL output current proportional to DEM ripple amplitude is not added to the output current of input OTAs, but it is used to correct the bias current of input OTAs, thus compensating G_m mismatch. In other words, while ORL provides an addictive correction, GERL provides a multiplicative one.

The second one is represented by modulator CH_9: in fact, the DEM ripple polarity (or phase, if preferred) depends on the input signal's polarity. So, before current demodulation, polarity must be inverted according to the signal: a comparator is then used to detect signal polarity and, if needed, invert the ripple current, thus preventing the feedback to turn from negative to positive.

Common Mode Equalization As said, if input and feedback common mode are different, the two OTAs can have different G_m and so affect the gain precision. This is actually a quite common situation: for instance,

since the output common mode voltage is usually stabilized at a certain fixed value, any variation on the input common mode will produce a G_m difference. Therefore, a large common mode drift or disturbance will cause a time dependant gain error, while if the amplifier is designed to work with a wide range of input common modes, gain error will depend on the source. This is particularly relevant for Rail-To-Rail input common mode amplifiers: such topologies are in fact caracterized by a double input differential pair (an n pair and a p pair), so that they can work together or one at a time depending on the input common mode, that can vary from one rail to the other. In such architectures G_m variations can reach 50%, if input common mode is near one rail while the feedback common mode is in the middle. And even if constant-G_m solutions are implemented [19], variations can still be as high as 6%, so countermeasures must be taken.

It must be noticed that DEM is not a solution for this problem (see again [17]). In fact, suppose that the feedback CM is higher than input CM: G_{mf} will then have a relative error Δ_{CM} that is added to the mismatch error Δ. We will have then

$$A_{DD}^{(1)} = \frac{1}{\beta}\frac{1}{1+\Delta+\Delta_{CM}} \tag{2.29}$$

During the other phase, the two OTAs switch roles, so now CM difference affects the other OTA (the one that did not have mismatch error Δ), so gain will be

$$A_{DD}^{(1)} = \frac{1}{\beta}\frac{1+\Delta}{1+\Delta_{CM}} \tag{2.30}$$

If relative gain error ϵ is evaluated in this case, the following expression can be found:

$$\epsilon = \left|1 - \frac{\frac{1}{1+\Delta+\Delta_{CM}} + \frac{1+\Delta}{1+\Delta_{CM}}}{2}\right| \approx \left|\frac{\Delta^2}{2} + \frac{\Delta\Delta_{CM}}{2} - \Delta_{CM}\right| \tag{2.31}$$

In conclusion, DEM does not reduce gain error due to a CM difference. When this difference can be high (like in R2R topologies) and the resulting gain error can not be tolerated, a Common Mode Equalization (CMEQ) technique can be used [20]. In this approach, the problem is solved radically by setting feedback CM to input CM. In this example, a Common Mode Difference Amplifier (CMDA), together with strong negative feedback, is used to achieve this result (figure 2.19).

CMDA is a high gain, high input impedance, low output impedance

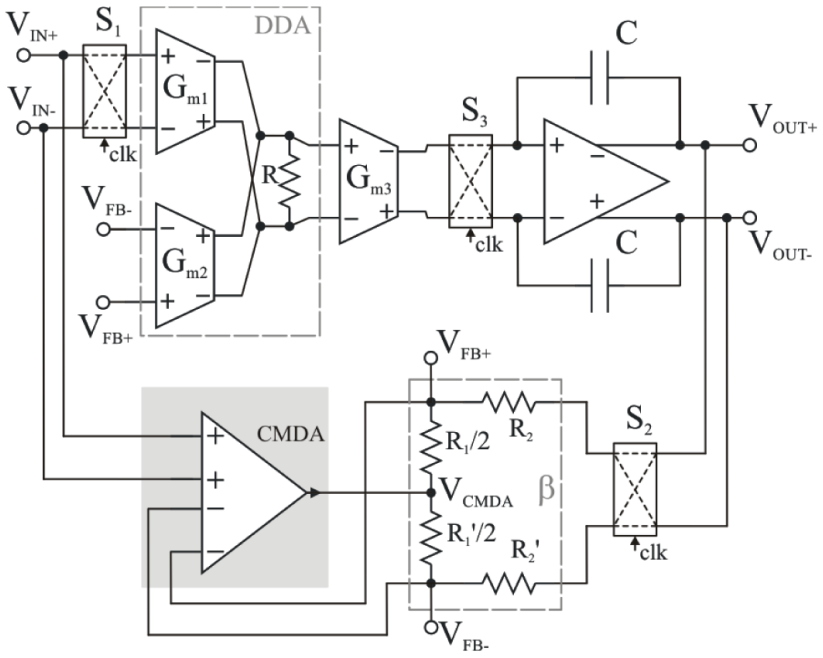


Figure 2.19: Common Mode Equalization in a CFIA.

amplifier characterized by such an expression:

$$V_{CMDA} = A\,(V_{1+} + V_{1-} - V_{2+} - V_{2-}) = A\,(V_{1CM} - V_{2CM}) \tag{2.32}$$

where 1 and 2 identify the two ports of the CMDA. It can be proved (see [20]) that, if connected in negative loop (shown in figure 2.19), the resulting feedback common mode is

$$V_{fbCM} = \frac{V_{fb+} + V_{fb-}}{2} = \beta V_{outCM} + (1-\beta)V_{inCM} \approx V_{inCM} \tag{2.33}$$

if $\beta \ll 1$ (in [20] $\beta = 200^{-1}$), and supposing perfect matching between resistors. It is relevant to notice that the feedback chopper modulator is placed *before* the resistive voltage divider: this is because otherwise a mismatch between resistors would turn a CM difference into a differential error in the voltage V_{fbD}, causing residual offset. This way, this offset is not modulated at the input of Gm2 and is then rejected through modulator S3, such as the offset from OTA1 and OTA2.

Chapter 3

High level desgin

This chapter will provide a detailed description of the proposed solution, highlighting the goals and the consequent high level topology and dimensioning choices. Only a block diagram architecture will be discussed, since transistor level design will be the object of chapter 4.

3.1 Reference design

The reference design for the solution proposed in this thesis is available in [18]. It is a Current Feedback Instrumentation Amplifier based on chopper stabilization and a second order State Variable LPF for chopped offset rejection: a high level schematic is shown in figure 3.1.

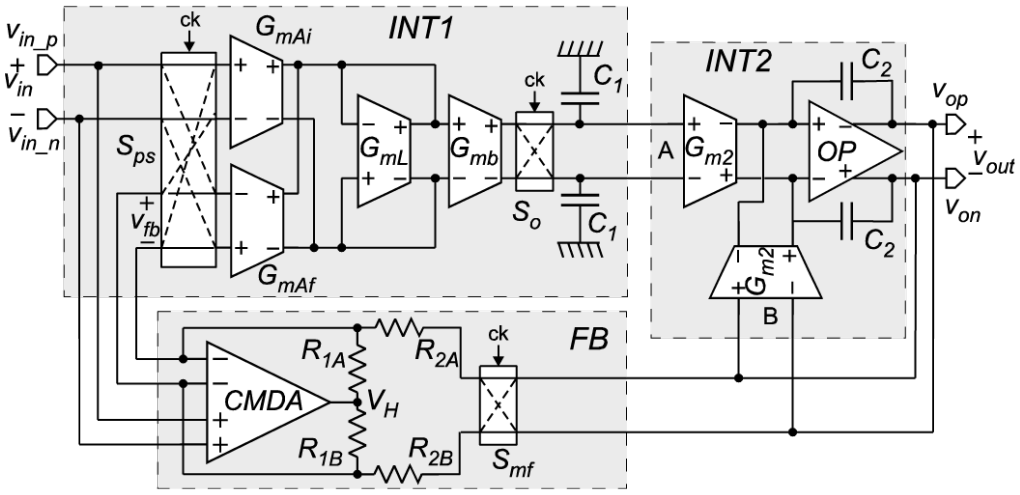


Figure 3.1: High level schematic of the reference design.

This device is thought to be part of a Data Acquisition System (DAS) for integrated thermocouples readout: it is then designed to have an input differential range of a few mV, while its input common mode range is not rail-to-rail (the input stage is based only on n differential pairs, whose input CM range does not extend to ground). The DC gain A_0 is set to 200 (nominal) through a resistive voltage divider:

$$\beta = \frac{R_{1A} + R_{1B}}{R_{2A} + R_{1A} + R_{1B} + R_{2B}} = A_0^{-1} = 200^{-1} \tag{3.1}$$

Futhermore, bandwidth specifications for thermal sensors applications required a cutoff frequency of a few hundreds Hz: a 200 Hz bandwidth is achieved in this design.

The key goals of this project are low power consumption and low offset/noise, with attention to area occupation as well: of course, gain accuracy is a prerogative too, as in every CFIA. Here follows a synthetic description of the main features of this implementation, with an eye on the problems caused by them, since solving these problems will be the main point of the original topology proposed in this thesis.

First of all, as a countermeasure against offset and flicker noise, **chopper modulation** is implemented: thanks to that, flicker noise is completely rejected and its corner frequency is as low as 0.2 Hz. As said, a **second order State Variable LPF** is in charge to reject the chopped offset. Dynamic techniques like ORL are not implemented, in order to avoid excessive schematic complexity and to contain chip area and current consuption. On the other hand, since only an LPF is responsible for ripple reduction, an attenuation of around 80 dB is requested. Now, chopper frequency f_{ch} was chosen high enough to completely reject flicker noise, but not too high as well, in order to reduce parasitic effects such as charge injection and to keep input impedance high: a value of 20 kHz was considered a good trade off. Furthermore, since a second order LPF has a 40 dB/dec slope, an attenuation of 80 dB was achieved with a 200 Hz cutoff frequency, that was coherent with bandwidth specifications.

A brief digression about the implementation of the filter is needed. Figure 3.2 shows again the block diagram of a second order State Variable LPF; the two integrators are represented by their ideal transfer function

$$H_{int}(s) = \frac{\omega_{0i}}{s} \tag{3.2}$$

where ω_{0i} is, for the i-th integrator, the unity-gain pulsation, that is the pulsations at which gain is 0 dB.

Now, different transfer functions can be identified. On one hand, it can

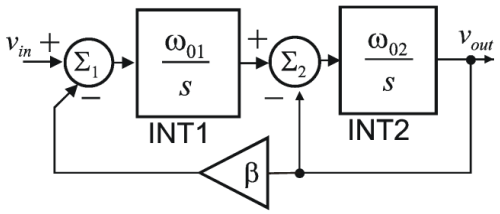


Figure 3.2: Second order State Variable LPF block diagram.

be found that the Signal Transfer Function (STF) has an expression as follows:

$$STF(s) = \frac{v_{out}}{v_{in}} = \frac{A_0}{1 + A_0\frac{s}{\omega_{01}} + A_0\frac{s^2}{\omega_{01}\omega_{02}}} \tag{3.3}$$

where $A_0 = 1/\beta = 200$ is the DC gain. For a second order filter, the characteristic frequency (cutoff frequency) ω_c and the quality factor Q can be defined: in this case, they turn out to be

$$\omega_c = \sqrt{\beta\omega_{01}\omega_{01}} \tag{3.4}$$

$$Q = \sqrt{\beta\frac{\omega_{01}}{\omega_{02}}} \tag{3.5}$$

The quality factor Q sets the position of the poles on the complex plane, and consequently the shape of the transfer function, that is the kind of filter. In this work, a Butterworth filter was chosen, due to its property to be maximally flat: for this filter, $Q = 1/\sqrt{2}$.

On the other hand, Noise Transfer Functions (NTF) can be defined, to describe how the i-th integrator's RTI noise v_{ni} is proposed at the output. Now, while obviously v_{n1} follows the same path of the signal, and so $NTF_1 = STF$, the second has such an expression:

$$NTF_2(s) = \frac{A_0\frac{s}{\omega_{01}}}{1 + A_0\frac{s}{\omega_{01}} + A_0\frac{s^2}{\omega_{01}\omega_{02}}} \tag{3.6}$$

This means that for v_{n2} this architecture acts as a Band Pass Filter (BPF), with peak of 0 dB at ω_c: noise generated from the second integrator INT2 will then be much less relevant at the output than the one from INT1, because not only it has a 0 dB peak, but above all because it filters away low frequencies (that is, flicker noise of INT2).

While this is a good advantage, noise of INT1 will be critical. First of all, integrators are usually (and in this case as well) implemented in a Gm-C

configuration, that is an OTA followed by a capacitor. This way, it is easy to get the ideal transfer function:

$$V_{out} = \frac{1}{Cs} I_{out} = \frac{G_m}{Cs} V_{in} = \frac{\omega_0}{s} V_{in} \tag{3.7}$$

if we assume that $\omega_0 = \frac{G_m}{C}$ (G_m is the the OTA's transconductance, that is in turn proportional to the active input MOSFETs' transconductance g_m). Then, considering the OTA's RTI thermal noise (flicker noise is ignored, supposed completely rejected by chopper modulation), it can be estimated as

$$S_{Vn-Th} = n \cdot 4 \frac{kT}{g_m} \tag{3.8}$$

where n is a coefficient depending on the topology. Now, in [18] the cutoff frequency $f_c = 200\,\text{Hz}$ was relatively low, so for eq. 3.4 the unity-gain frequencies ω_{0i} will be low as well (their ratio is fixed by Q). In addiction, capacitors are meant to be on-chip, meaning that their capacitance will not be higher than few tens of pF, if not using extremely large areas. So, low cutoff frequency and low capacitances intevitably lead to low $G_m = \omega_c C$, that means high RTI thermal noise. This problem, if large areas are not desirable, can be solved by means of a **preamplifier** preceding INT1 (as said, noise from INT2 is less critical).

Suppose that the preamplifier has a DC gain A_{pre}. Then, in order to mantain the same overall G_{m1}, the OTA will need to have a transconductance

$$G'_{m1} = \frac{G_m}{A_{pre}} \tag{3.9}$$

The resulting RTI thermal noise PSD will then be:

$$S_{Vn-Th} = \frac{1}{A_{pre}^2} n \cdot 4 \frac{kT}{G'_m} = \frac{1}{A_{pre}} n \cdot 4 \frac{kT}{G_m} \tag{3.10}$$

with a net reduction of thermal noise by a factor A_{pre}. Of course, the preamplifier's RTI noise has now to be taken into account, but it is only a matter of current consumption and it is independent from frequency specifications. In [18], A_{pre} was set to be 600, and the total RTI noise density was $18\,\text{nV}/\sqrt{\text{Hz}}$.

It is relevant to notice that, while a high gain preamplifier can strongly reduce RTI noise, it may cause dynamic-related problems. Indeed, while at low frequencies virtual short circuit nulls the preamplifier's input voltage

difference ($v_{inD} - v_{fbD}$), this is no more true when signals approach the cutoff frequency. Suppose an input signal within the range, that is with an amplitude less than $V_{out,max}/A_0$. At low frequencies, virtual short circuit is valid, input voltage difference ($v_{inD} - v_{fbD}$) is negligible and the preamplifier output is still near zero. But if this signal has a frequency near f_c, the output signal will have a phase shift of around 90°, since it is a second order filter. The preamplifier's input voltage difference then can be even higher than the amplitude of the input signal itself: while such amplitude may not saturate the CFIA's output, it most probably will saturate the preamplifier's output, since $A_{pre} > A_0$. For this reason, that work showed extremely high Total Harmonic Distortion (THD) when the frequencies approach the upper band limit (already at 90 Hz, see figure 3.3) and relatively high amplitudes, that, on the other hand, produce negligible distortion at much lower frequencies.

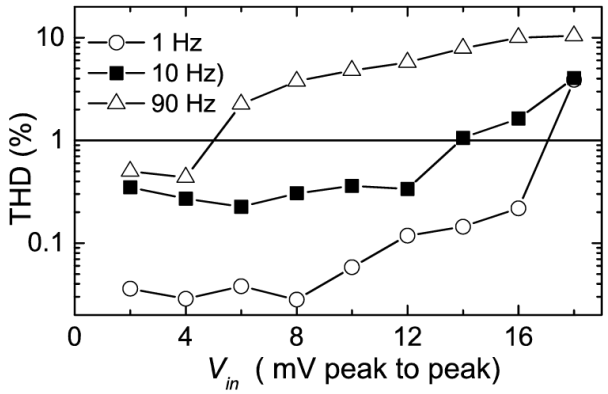


Figure 3.3: THD vs input amplitude for different frequencies.

As a matter of fact, INT2 as well can suffer from this problem, that this time has a different nature: while for the preamplifier the output range was critical, here INT2 input range can be a limitation. In fact, INT2 is designed in a Gm-OpAmp configuration, in order to have a low output resistance and drive properly the resistive feedback network, but however its first stage is a pair of OTAs, each based on a p differential pair. These should deal with differential voltages as large as the output voltage, that is impossible for a differential pair since their input range usually does not exceed a few hundreds of mV.

This problem can be partially solved with a routing trick: instead of connecting each signal (input v_{in2} and feedback $v_{fb2} = -v_{out}$) to a different pair, it is possible to give ($v_{in2+} - v_{fb2+}$) to one pair and ($v_{in2-} - v_{fb2-}$) to the other. This way, when virtual short circuit is valid and supposed that the v_{in2} and v_{out} have the same common mode, these voltage difference would be zero and INT2 will work in linearity region. But again, if signals

approach the cutoff frequency, they can be shifted in phase and the non-zero voltage difference can saturate the differential pairs.

A much more critical problem, finally, affects INT2 when it deals with large signals, and this time it is not frequency-dependent (see again figure 3.3). As said, its two differential pairs have $(v_{in2+} - v_{out+})$ and $(v_{in2-} - v_{out-})$ as input differential voltages, that at low frequencies are negligible ($(v_{in2+} \approx v_{out+})$ and $(v_{in2-} \approx v_{out-})$). On the other hand, their input common mode is strongly dependent on the output signal: thus, for large signal, one pair will have an input CM near the positive rail and, the other, near ground. This can create a mismatch between the two pairs' transconductances, and more severely it can turn off one of the pairs if the input CM is high enough.

This problem was mitigated by setting an output CM closer to ground, but of course it was a limitation for the output range.

Finally, among the other techniques used in this work, **Port Swapping** for G_m equalization (elsewhere called DEM, see paragraph 2.3.2) must be mentioned. As demonstrated in [18], in this case the combination of chopping and port swapping has a positive side effect in increasing input impedance. In fact, for the particular alternation of signals on the input ports (figure 3.4), slowly varying signals do not cause the input capacitances to be charged and discharge periodically, strongly reducing input current.

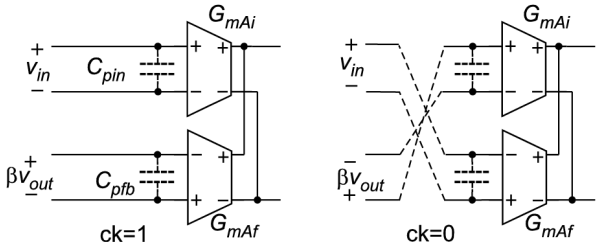


Figure 3.4: Input connections in the two clock phases.

Since the input capacitances switch between v_{in} and $-v_{fb} = \beta v_{out}$, the resulting current over a complete clock cycle is $i_{in} = 2(v_{in} - \beta v_{out}) f_{ch} C_{inD}$, at least for $f \ll f_{ch}$. So the following expression can be derived [18]:

$$z_{in}^{(ch-ps)}(\omega) = \frac{1}{2[1 - \beta STF(j\omega)] f_{ch} C_{inD}} \tag{3.11}$$

This would predict an infinite input impedance for f that tands to zero, while in a real implementation it would be proportional to the finite open loop DC gain of INT1 A_{OL1}. At first order approximation (and considering INT2 ideal), it is

$$STF(j\omega = 0) = \frac{A_{OL1}}{1 + \beta A_{OL1}} = \frac{1}{\beta} \frac{1}{1 + \frac{1}{\beta A_{OL1}}} \approx \frac{1}{\beta} \left(1 - \frac{1}{\beta A_{OL1}}\right) \tag{3.12}$$

that, substituted in 3.11, gives

$$z_{in}^{(ch-ps)}(0) = \frac{1}{2[1 - \beta STF(j0)] f_{ch} C_{inD}} = \frac{\beta A_{OL1}}{2 f_{ch} C_{inD}} \tag{3.13}$$

that is an increment of a factor βA_{OL1} with respect to the classical chopper input impedance

$$z_{in}^{(ch)}(0) = \frac{1}{4 f_{ch} C_{inD}} \tag{3.14}$$

The last technique implemented by this reference design that is worth mentioning is, as shown in figure 3.1, **Common Mode Equalization**. Thus, a Common Mode Differential Amplifier (CMDA) is implemented, whose operation has been described in paragraph 2.3.2, even if this work has not a Rail-To-Rail input CM range.

3.2 Proposed solution

3.2.1 Goals

The proposed solution aims to solve the problems highlighted for the reference design, with a change in topology and in the value of parameters. Furthermore, the basic concept to achieve with this work is **flexibility**, that is the ability of the amplifier to interface the widest variety of sensors, above all in terms of bandwidth and input CM range. Thus, the major new features implemented are listed below.

- A **third order state variable LPF** is implemented. Chopping is a compulsory feature for such a low noise amplifier: so this technique is mantained in this project, and the chopping frequency $F_{ch} = 20\,\text{kHz}$ is the same as well, for the reasons mentioned in paragraph 3.1. However, a $-60\,\text{dB/dec}$ slope of a third order filter allows, being equal the attenuation at f_{ch}, a quite higher cutoff frequency. This, in turn, will increase the integrators' unity-gain frequencies ω_{0i} and, capacitance values being equal, will increase the tranconductances G_{mi}, decreasing the RTI thermal noise. In addiction, a wider band is coherent with

the idea of flexibility introduced above: the bandwidth will result to be $f_c = 1\,\text{kHz}$, quite higher than the previous 200 Hz.

- A **preamplifier** is still needed to reduce the critical INT1 RTI noise, but since this is already lower (thanks to the higher G_{m1}, it will allow the preamplifier to have a lower gain, RTI noise being equal. With a lower gain, then, dynamic problems for large signals at band limits will be less critical and THD will be improved (even because the frequencies that in the reference design were at band limits now are fully inside the band). The preamplifier gain will be set at $A_{pre} = 50$, much less than the previous 600.

- A **voltage divider feedback network** β_2 was added for the INT2 (and now INT3 as well) loop. This, while slightly increasing their output noises, that in any case are not critical, will solve or at least reduce the already mentioned dynamic problems. In fact, now INT2 and INT3 will deal with lower voltages, that are more likely to be within the input differential or CM ranges of a differential pair. A good trade off between noise and dynamic was chosen to be a value of 2, that is a coefficient $\beta_2 = 0.5$

- A **Rail-To-Rail input CM range** was achieved. This feature, while it tends to increase RTI total noise, current consumption and chip area, was implemented in order to create a really flexible device, that could interface a wide varieties of sensors: while it is still thought to interface thermocouples and it is optimized to work with a $v_{inCM} = V_{DD}/2$, it works with other input CM voltages as well, from V_{DD} to ground.

- A **Rail-To-Rail output swing** is still present, in order to maximize the dynamic. A closed loop gain $A_0 = 200$ is mantained, because it was a good trade off for thermocouples sensing; on the other hand, however, supply voltage has been decreased from 3.3 V to $V_{DD} = 1.8\,\text{V}$, so that this device could be suitable for most of portable/mobile application and for lower power consumption as well. This decrease in supply rail requires a R2R output swing, in order to mantain a good dynamic that, however, is a bit narrower than in the previous work. As a bright side, the increased linearity over a wider range of voltages let the dynamic to be exploited completely with low distortion.

3.2.2 Third order LPF design

The first step of the high level design was the sizing of the state variable filter. During this phase, Python programming language was widely use, thanks to its libraries: SciPy and mostly the Signal package were used to design the filter and evaluate the filter's parameters, MatPlotLib was then used to plot graphics and compare different results.

Transfer function analysis First of all, it is necessary to analyze a third order LPF transfer function, that can be expressed in a general form as follows:

$$H_{LPF}(s) = \frac{b_0}{a_0 + a_1 s + a_2 s^2 + a_3 s^3} \tag{3.15}$$

The SciPy function `signal.butter()`, given the filter order and the cutoff frequency (or, more accurately, the characteristic frequency ω_N), outputs the arrays `b[]` and `a[]`, that are the coefficients of Butterworth LPF's transfer function polynomials. For a Chebyshev filter, the function `signal.cheby1()` does the same, but it needs as third parameter the maximum ripple allowed inside the band.

In case of a state variable implementation, a block diagram of the filter would be as shown in figure 3.5.

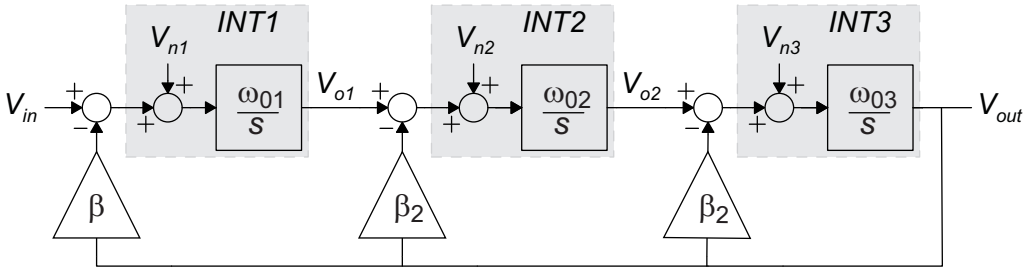


Figure 3.5: Block diagram of a 3rd order State Variable LPF.

From the schematic, it is easy to obtain an expression of the Signal Transfer Function $STF = v_{out}/v_{in}$ as a function of the schematic parameters ω_{0i}, β and β_2. In fact:

$$v_{out} = \frac{\omega_{03}}{s}\left\{-\beta_2 v_{out} + \frac{\omega_{02}}{s}\left[-\beta_2 v_{out} + \frac{\omega_{01}}{s}\left(-\beta v_{out} + v_{in}\right)\right]\right\} \tag{3.16}$$

from which, after simple algebric steps, we obtain:

$$STF(s) = \frac{v_{out}}{v_{in}} = \frac{1}{\beta + \beta_2 \dfrac{s}{\omega_{01}} + \beta_2 \dfrac{s^2}{\omega_{01}\omega_{02}} + \dfrac{s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.17}$$

This equation is coherent with the specification about DC gain:

$$\lim_{s \to 0} STF(s) = \frac{1}{\beta} = A_0 \tag{3.18}$$

In filter design, when comparing different kinds of filters, it is often useful to work on *normalized* transer functions, that is functions that have a 0 dB DC gain and a characteristic pulsation $\omega_N = 1\,\text{rad/s}$. The normalized STF for this filter will then be

$$STF_N = \frac{STF(s\omega_N)}{STF(0)} = \frac{1}{1 + \beta_2 A_0 \dfrac{\omega_N}{\omega_{01}} s + \beta_2 A_0 \dfrac{\omega_N^2}{\omega_{01}\omega_{02}} s^2 + A_0 \dfrac{\omega_N^3}{\omega_{01}\omega_{02}\omega_{03}} s^3} \tag{3.19}$$

Thanks to this relation, it is possible to link the numerical coefficients produced by the Python functions to the physical parameters of the schematic. As said, the goal is to design a filter that, given an attenuation of 80 dB at f_{ch}, would **maximize the value of** ω_{01}, in order to have an high G_{m1} for INT1's OTA an so a low RTI noise (it will be proven that, even in this case, noises from INT2 and INT3 are not critical).

Butterworth vs Chebyshev The choice of the filter type was limited between Butterworth and Chebyshev, the most diffused and studied in literature. The first one has the property of being *maximally flat* within the band, that is a good feature that an InAmp should have. On the other hand, a Chebyshev filter allows a ripple within the band, but its slope for frequencies immediately higher than cutoff is quite higher: this, being equal the attenuation at f_{ch}, could allow a higher characteristic frequency and, potentially, higher ω_{0i}.

In any case, it is necessary to compare the two frequency responses *being equal the attenuation at* f_{ch}: this could lead to different ω_N that must be evaluated. Thus, normalized filter polynomials are generated with the following Python code (Chebyshev filter is evaluated with 2 dB maximum ripple within the band):

```
bn_butt, an_butt = signal.butter(3, 1, analog=True)
bn_cheb, an_cheb = signal.cheby1(3, 2, 1, analog=True)
```

```
w_n = np.logspace(-2,2,501)
w_n, magn_butt, phn_butt = signal.bode((bn_butt, an_butt),w_n)
w_n, magn_cheb, phn_cheb = signal.bode((bn_cheb, an_cheb),w_n)
if(plot):   #plot normalized frequency responses
    plt.semilogx(w_n, magn_butt, 'b', w_n, magn_cheb, 'r')
    plt.show()

wp_n_butt = w_n[find_nearest(magn_butt, -80)]
wp_n_cheb = w_n[find_nearest(magn_cheb, -80)]
```

The function `find_nearest()` then scans the magnitude vs. frequency arrays generated with `signal.bode()`, and finds the pulsation at which attenuation is around 80 dB. For the two filters, this pulsation is called $\omega_{p-n-butt}$ (`wp_n_butt`) and $\omega_{p-n-cheb}$ (`wp_n_cheb`). Since these filters are normalized, this value will be the ratio between $\omega_{ch} = 2\pi f_{ch}$ and ω_N in the real filters.

The function gives the following values:

$$\omega_{p-n-butt} = 21.677 \rightarrow \quad \omega_{N-butt} = \frac{2\pi f_{ch}}{\omega_{p-n-butt}} = 5794\,\text{rad/s} \approx 2\pi\ 920\,\text{Hz} \tag{3.20}$$

$$\omega_{p-n-cheb} = 14.723 \rightarrow \quad \omega_{N-cheb} = \frac{2\pi f_{ch}}{\omega_{p-n-cheb}} = 8530\,\text{rad/s} \approx 2\pi\ 1330\,\text{Hz} \tag{3.21}$$

Now that the characteristic frequencies have been estimated, it is possible to equate the coefficient of denominator polynomial of STF_N (equation 3.19) with the values given by the Python functions. First of all, it is important to notice that the arrays `a[]` and `b[]` are normalized in order to make the maximum order term of `a[]` equal to 1, while for a better comparison we should have the zero-order term equal to 1 (and thus the nominator as well). Such operation is achieved by the following lines:

```
an_n_butt = an_butt/bn_butt
an_n_cheb = an_cheb/bn_cheb
```

Table 3.1 shows the values of the normalized denominator's coefficients for both filters, compared with the corrispondent coefficients in STF_N. It must be noticed that, for implementation reasons in Python, the first element of an array (`an_n_butt[0]`, for instance) corresponds to the coefficient of the highest-order term of the polynomial (the third-order term, in this case), and so on.

It is now possible to find the unity-gain frequencies in both cases, by solving the following linear system:

	an_n[3]	an_n[2]	an_n[1]	an_n[0]
Butterworth	1	2	2	1
Chebyshev	1	3.127	2.257	3.059
STF_N	1	$\beta_2 A_0 \frac{\omega_N}{\omega_{01}}$	$\beta_2 A_0 \frac{\omega_N^2}{\omega_{01}\omega_{02}}$	$A_0 \frac{\omega_N^3}{\omega_{01}\omega_{02}\omega_{03}}$

Table 3.1: Normalized transfer function coefficients comparison.

$$\begin{cases} \beta_2 A_0 \dfrac{\omega_N}{\omega_{01}} & = \texttt{an\_n[3]} \\ \beta_2 A_0 \dfrac{\omega_N^2}{\omega_{01}\omega_{02}} & = \texttt{an\_n[2]} \\ A_0 \dfrac{\omega_N^3}{\omega_{01}\omega_{02}\omega_{03}} & = \texttt{an\_n[1]} \end{cases} \tag{3.22}$$

Finally, table 3.2 shows the evaluated values of ω_{0i} for both filters, using the values of ω_N found in eq. 3.20 and with $\beta_2 = 0.5$.

	Butterworth	Chebyshev
ω_{01}	289 krad/s	267 krad/s
ω_{02}	5.78 krad/s	11.58 krad/s
ω_{03}	23.12 krad/s	12.33 krad/s

Table 3.2: Unity-gain pulsations comparison.

Since the most critical value is, as said, the first one, the two filters seem more or less equivalent, with Butterworth slightly better (even if ω_N for Chebyshev was quite higher).

Time domain analysis: step response For a better decision, a time domain test has been made, and the step response was evaluated. The test consists in stimulating the filter with a step waveform and analyze the output waveform. Because of the poles of the system, the output waveform will not follow immediately the input, but it will rise slowlier, it may have overshoots until it will settle at its steady-state value. Many quantitative parameters can be considered in evaluating a step response: in this work, *settling time* has been taken as parameter of merit.

The settling time to α is the interval of time that passes from the beginning of the input stimulus to the moment in which the output waveform reaches for the first time the error band $V_{SS}(1 \pm \alpha)$, without ever leaving it. In mathematical terms, if $t_0 = 0$ is the moment in which the input step

rises and V_{SS} is the output steady-state value, settling time to α is defined as

$$t_{s-\alpha} : \left| \frac{V_{out}(t)}{V_{SS}} - 1 \right| < \alpha \quad \forall t > t_{s-\alpha} \tag{3.23}$$

Figure 3.6 shows more clearly the definition of settiling time.

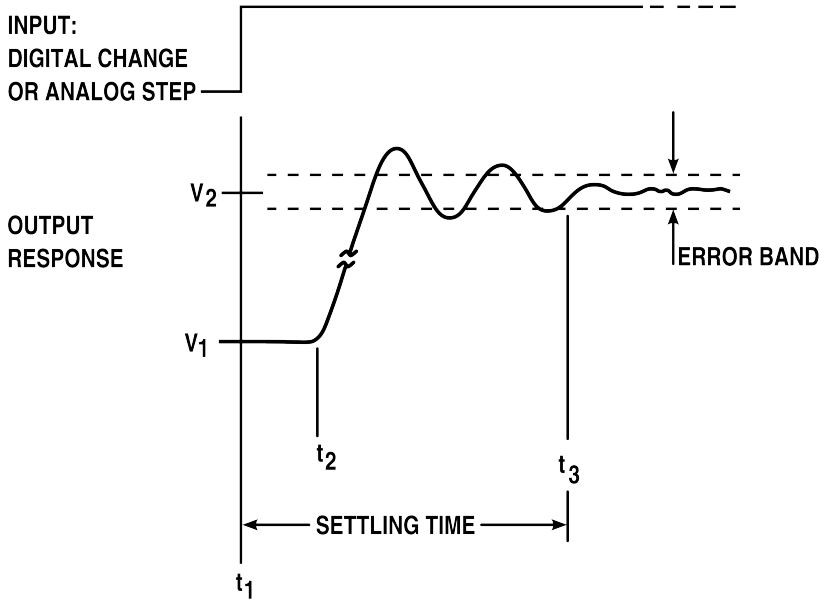


Figure 3.6: Graphic representation of settling time.

Since the two filters have a different ω_N, it is necessary to generate their transer functions normalized only in amplitude and not in frequency. Thus, the following code was used to generate such transfer functions, while figure 3.7 shows the magnitude of their Bode diagram, for comparison.

```
wn_butt = 2*np.pi*920
wn_cheb = 2*np.pi*1330
b_butt, a_butt = signal.butter(3, wn_butt, analog=True)
b_cheb, a_cheb = signal.cheby1(3, 2, wn_cheb, analog=True)
```

Then, the SciPy function `signal.step()` was used to evaluate the step response for both filters. The user-defined function `t_settle()` finally is in charge to evaluate the settling time for a given step response `y`, function of time, and parameter α, here called `threshold`, as shown by the following lines.

```
def t_settle(t, y, threshold):
    i_last = 0
    for i in range(len(t)):
        if (abs(1-y[i])>threshold):
```

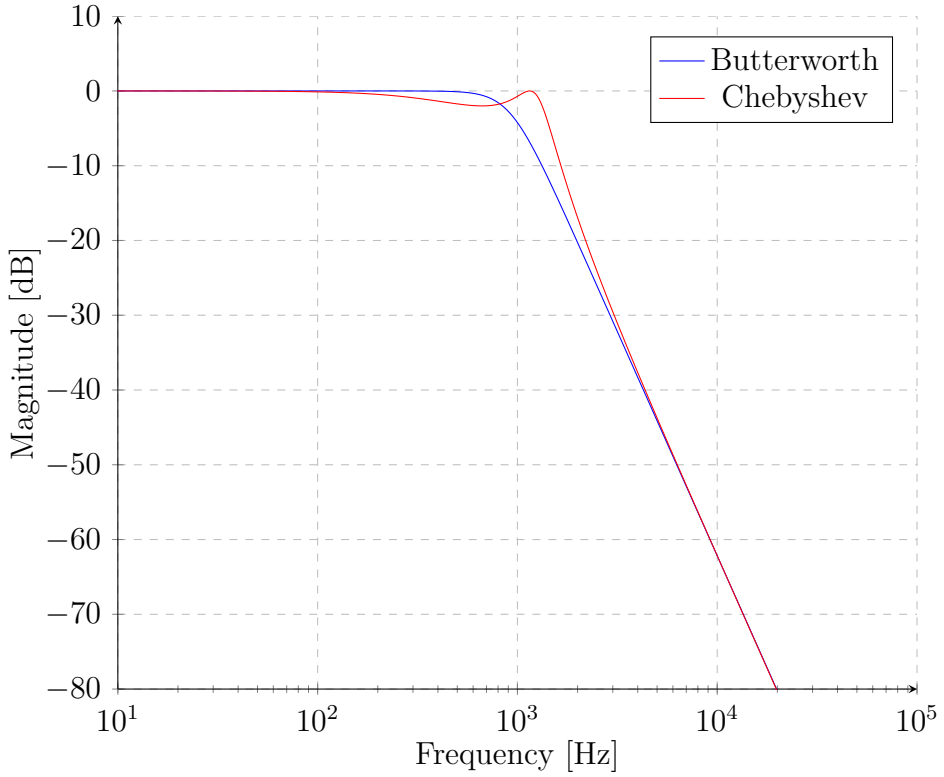


Figure 3.7: Magnitude frequency response comparison.

```
            i_last = i
    return t[i_last + 1]
```

Here is the portion of code in which the two step responses and settling times are evaluated. As values of α, 1% and 0.1% were chosen, since the former is a standard parameter for settling time, while the latter is more coherent with the expected gain error of the amplifier.

```
t = np.linspace(0,0.01,1001)
t, y_butt = signal.step((b_butt, a_butt),T=t)
t, y_cheb = signal.step((b_cheb, a_cheb),T=t)
t_settle1_butt = t_settle(t,y_butt, 0.01)
t_settle1_cheb = t_settle(t,y_cheb, 0.01)
t_settle01_butt = t_settle(t,y_butt, 0.001)
t_settle01_cheb = t_settle(t,y_cheb, 0.001)
```

Figure 3.8 shows the plot of both step responses vs time. As we can see, Butterworth filter's step response has a larger overshoot than Chebyshev, but the oscillations are damped faster, so that the settling time turns out to be shorter for Butterworth. Table 3.3 reports the exact values of settling times evaluated by the funtion `t_settle()`.

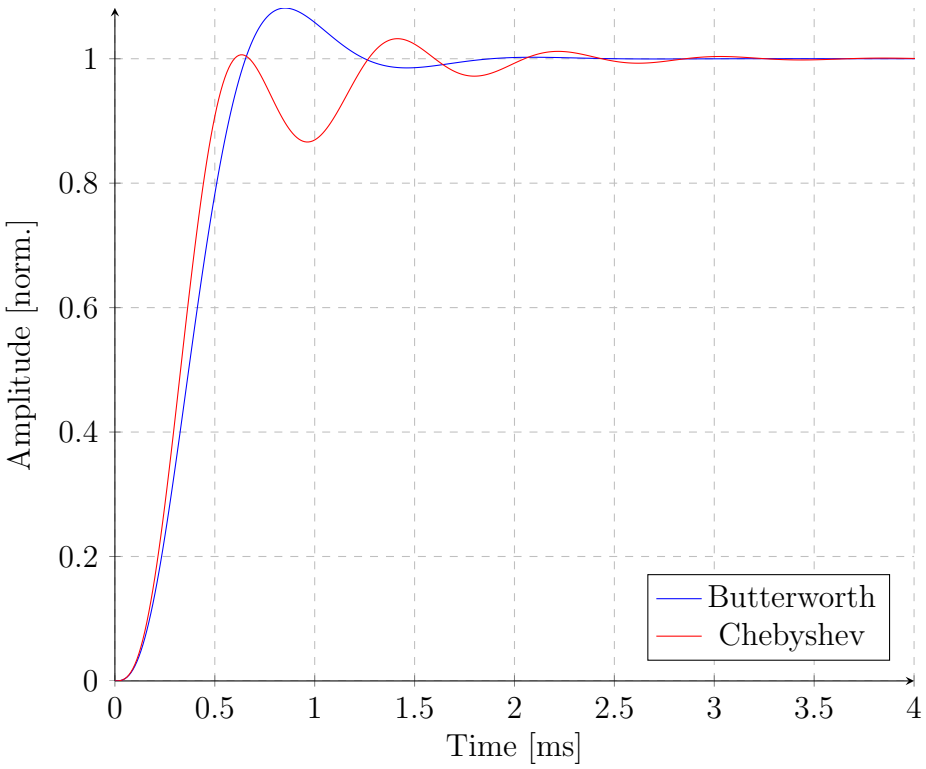


Figure 3.8: Step response comparison.

	$t_{s-1\%}$	$t_{s-0.1\%}$
Butterworth	1.63 ms	2.35 ms
Chebyshev	2.29 ms	3.88 ms

Table 3.3: Settling time comparison.

So, it is evident that settling times of our Butterworth filter are, in this case, much better than the ones of our Chebyshev filter. Therefore, in the end, the final choice for the filter implementation was Butterworth, that - must be reminded - has as well a maximally flat frequency response within the band, a desirable property for an intrumentation amplifier.

3.2.3 Final high level design

Here we summarize the final values of all parameters of the high level design.

First of all, the characheristic frequency $f_N = \omega_N/2\pi$ (that is, for a Butterworth filter, also the cutoff frequency f_c at $-3\,\mathrm{dB}$) was rounded up

from 920 Hz to 1 kHz.

Signal and Noise Transfer Functions The following code was used to generate the final transfer function (neither normalized in frequency nor in amplitude) and the noise transfer functions: these can be found to have the following expressions (with, obviously, $NTF_1 = STF$):

$$STF = \frac{v_{out}}{v_{in}} = \frac{A_0}{1 + \beta_2 A_0 \frac{s}{\omega_{01}} + \beta_2 A_0 \frac{s^2}{\omega_{01}\omega_{02}} + A_0 \frac{s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.24}$$

$$NTF_2 = \frac{v_{out}}{v_{n2}} = \frac{A_0 \frac{s}{\omega_{01}}}{1 + \beta_2 A_0 \frac{s}{\omega_{01}} + \beta_2 A_0 \frac{s^2}{\omega_{01}\omega_{02}} + A_0 \frac{s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.25}$$

$$NTF_3 = \frac{v_{out}}{v_{n3}} = \frac{A_0 \frac{s^2}{\omega_{01}\omega_{02}}}{1 + \beta_2 A_0 \frac{s}{\omega_{01}} + \beta_2 A_0 \frac{s^2}{\omega_{01}\omega_{02}} + A_0 \frac{s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.26}$$

$$\tag{3.27}$$

```
wn_butt_d = 2*np.pi*1000
b_butt_d, a_butt_d = signal.butter(3, wn_butt_d, analog=True)
A0 = 200
B = 0.5
a = a_butt_d/b_butt_d
b = np.array([0, 0, 0, A0])
b2 = np.array([0, 0, a[2]/B, 0])
b3 = np.array([0, a[1]/B, 0, 0])
w, mag, ph = signal.bode((b, a),w)
w, mag2, ph2 = signal.bode((b2, a),w)
w, mag3, ph3 = signal.bode((b3, a),w)
```

Figure 3.9 shows a plot of the magnitude of these three transfer functions. We found that both NTF have a peak value of 9 dB around f_c. NTF_2 rises from DC to the peak with a slope of 20 dB/dec, while it falls to infinity with -40 dB/dec; NTF_3, instead, rises from DC to the peak with a slope of 40 dB/dec, while it falls to infinity with -20 dB/dec. This means that, as said, both integrators' flicker noise will be completely rejected, and in general noise from INT2 and INT3 will be less crtitical than the one from INT1. However, it must be noticed that, although this is a third order filter, overall high frequency noise will be dominated by v_{n3}, that is filtered like in a first order LPF: so when integrated it could be a problem if it is too large, and could be a significant portion of the total output noise power.

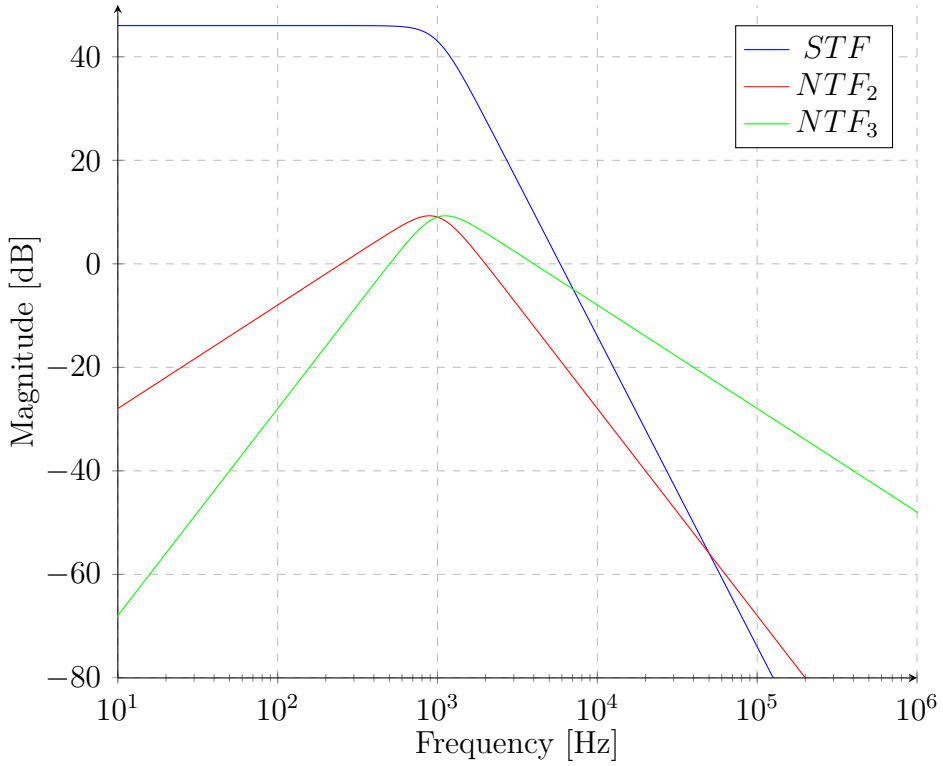


Figure 3.9: Signal and Noise Transfer Functions.

Noise At this point, it is interesting to estimate some noise parameters. Since actual transistor-level architechture has not been designed yet, it could be difficult to set the values for some parameters: for instance, we know that capacitances should fall in the range of tens of pF, but we cannot know at this point what will be a good value for them. Then, we will use the values that were set after a transistor-level design, as if we already knew them, for a more precise estimation.

Thus, first of all it is possible to evaluate the unity-gain frequencies, by solving again the linear system 3.22 with the new value of $\omega_N = 2\pi\ 1\,\text{kHz}$. We obtain:

$$\omega_{01} = 314\,\text{krad/s} = 2\pi\ 50\,\text{kHz} \tag{3.28}$$
$$\omega_{02} = 6.28\,\text{krad/s} = 2\pi\ 1\,\text{kHz} \tag{3.29}$$
$$\omega_{03} = 25.1\,\text{krad/s} = 2\pi\ 4\,\text{kHz} \tag{3.30}$$

Then, we can evaluate the INTi transconductances G_{mi} by setting values

for the capacitances, and reminding that INT1 is formed by a preamplifier with $A_{pre} = 50$ followed by the actual integrator, that will have an effective $G_{m1} = G_{m1-eq}/A_{pre}$. The chosen values for the capacitances are

$$C_1 = 65\,\text{pF} \qquad C_2 = 100\,\text{pF} \qquad C_3 = 50\,\text{pF} \tag{3.31}$$

from which we find the transconductances (eq. 3.7):

$$G_{m1-eq} = C_1\omega_{01} = 20.42\,\mu\text{A/V} \tag{3.32}$$

$$G_{m1} = \frac{G_{m1-eq}}{A_{pre}} = 408.4\,\text{nA/V} \tag{3.33}$$

$$G_{m2} = C_2\omega_{02} = 628.3\,\text{nA/V} \tag{3.34}$$

$$G_{m3} = C_3\omega_{03} = 1.257\,\mu\text{A/V} \tag{3.35}$$

From these values, we obtain a rough estimation of the three RTI noises. Of course, it would be impossible to do it without any idea of the transistor-level implementation of the block, so some anticipation is needed.

First of all, INT1 will have an OTA (called OTA1) with a single input differential pair, since the subtraction between v_{in} and v_{fb} has already been done by the preamplifier. Furthermore, OTA1 will have only one differential pair, p or n, since its input CM will be set by the preamplifier, with no need of R2R input CM range. Neglecting the flicker noise, supposed completely rejected by chopping modulation, its RTI noise will be only thermal and will have such an expression:

$$S_{Vn1-th} = \frac{1}{A_{pre}^2} \cdot 4\frac{kT}{G_{m1}} \cdot \alpha_t \cdot 2 = 1.62 \times 10^{-16}\,\text{V}^2/\text{Hz} \tag{3.36}$$

As said, OTA1 noise PSD is reduced by a factor A_{pre}^2 thanks to preamplification. The coefficient α_t is a topological factor: the expression $4\frac{kT}{G_{m1}}$ alone, in fact, expresses the RTI noise of only one input MOSFET of the pair, while in OTA1 two input MOSFETs will be present and other transistors will contribute to the total noise, hence a factor $\alpha_t = 5$ was used in this case (coherent with following simulations). The coefficient 2 is used to include the preamplifier's RTI noise, supposed that the total RTI noise is equally due to OTA1 and preamplifier (preamplifier noise is a DoF - Degree of Freedom - and depends only on current consumption). Anyway, from this formula, INT1 RTI noise density turns out to be

$$v_{n1} = \sqrt{S_{Vn1-th}} \approx 12.7\,\mathrm{nV}/\sqrt{\mathrm{Hz}} \tag{3.37}$$

Of course, this is just a rough estimation, since many factors can alter this value, first of all flicker noise, that will not be completely rejected, and process-related variations, that can be controlled only with a simulator. In any case, a maximum total RTI noise density value of $20\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ could be a good target for such architecture.

For INT2 and INT3, in a similar way, RTI noise PSD can be evaluated, considering that this time there are 2 input differential pairs per OTA. With proper coefficient adjustments, a rough estimation for their RTI noise densities could be:

$$v_{n2} \approx 360\,\mathrm{nV}/\sqrt{\mathrm{Hz}} \qquad\qquad v_{n3} \approx 250\,\mathrm{nV}/\sqrt{\mathrm{Hz}} \tag{3.38}$$

In this case as well only a simulation could give some relevant information. Anyway, again supposing to neglect flicker noise, is it possible to integrate this contributes at the output, that is after being processed by their NTF_i, to estimate the total output noise power. Dividing it by A_0^2 then we get the input noise power. A square root of input noise power will give us the input RMS (Root Main Square) noise, that is strongly linked to the Dynamic Range parameter. In order to evaluate the output power, Python user-defined functions were used, so that the transfer functions could be integrated with respect to frequency and not pulsation. Since the input quantities (thermal noise PSDs) were supposed to be constant with frequency, first of all only the frequency responses were integrated, giving the *equivalent bands*, to which noise PSDs were multiplied. The following code shows the user defined function and the integration, achieved by means of the SciPy Integrate pakage function `scp_int.quad()`.

```
def h_mod2(f):
    jw = np.array([(2*np.pi*f*1j)**3,(2*np.pi*f*1j)**2,(2*np.pi*f*1j),1])
    return abs(np.dot(b,jw)/np.dot(a,jw))**2

def h2_mod2(f):
    jw = np.array([(2*np.pi*f*1j)**3,(2*np.pi*f*1j)**2,(2*np.pi*f*1j),1])
    return abs(np.dot(b2,jw)/np.dot(a,jw))**2

def h3_mod2(f):
    jw = np.array([(2*np.pi*f*1j)**3,(2*np.pi*f*1j)**2,(2*np.pi*f*1j),1])
    return abs(np.dot(b3,jw)/np.dot(a,jw))**2


e, err = scp_int.quad(h_mod2,0,np.inf)
e2, err2 = scp_int.quad(h2_mod2,0,np.inf)
e3, err3 = scp_int.quad(h3_mod2,0,np.inf)
```

```
Pno1 = Svth1*e
Pno2 = Svth2*e2
Pno3 = Svth3*e3
Pnout = Pno1 + Pno2 + Pno3
Pnrti = Pnout/(A0**2)
vnrti = Pnrti**0.5
```

These are the results for said calculations:

$$P_{no1} = \int_0^\infty S_{Vn1-th} \cdot NTF_1(j2\pi f)df = 6.79\,\text{nW} \tag{3.39}$$

$$P_{no2} = \cdots = 1.10\,\text{nW} \tag{3.40}$$

$$P_{no3} = \cdots = 1.10\,\text{nW} \tag{3.41}$$

As expected, noise power is mostly dominated by the INT1 contribute, and this justifies the use of a preamplifier only for it. In any case, the following values are obtained for total output and input (RTI) noise:

$$P_{no} = 9.0\,\text{nW} \qquad P_{ni} = 0.225\,\text{pW} \tag{3.42}$$

$$v_{noRMS} = 94.9\,\mu\text{V} \qquad v_{niRMS} = 0.47\,\mu\text{V} \tag{3.43}$$

Finally, it is interesting to evaluate the Dynamic Range, whose expression was mentioned in eq. 1.15. In this case, we have (output quantities are considered)

$$DR = \frac{\Delta V_{FS}}{v_{n-pp}} = \frac{2\Delta V_{out-swing}}{4v_{noRMS}} \approx \frac{2 \cdot 1.6\,\text{V}}{4 \cdot 94.9\,\mu\text{V}} = 8430 \tag{3.44}$$

that, reminding eq. 1.12, gives us an Equivalent Number Of Bits of

$$ENOB = \lfloor \log_2(DR) \rfloor = \lfloor 13.04 \rfloor = 13 \tag{3.45}$$

Partial Signal Transfer Function As additional check, we should make sure that intermediate signals are within the integrators' input range for all the frequencies: in fact, while we assured that output response is flat (and, scaled by β_2) within the integrators' input ranges), partial STFs could have resonances that could saturate intermediate stages. Thus, we want to find a closed loop transfer function between the input signal v_{in} and INT1 and INT2 outputs, respectively v_{o1} and v_{o2}. From the block diagram in figure 3.5 at page 54, it is easy to come to the following relations.

$$v_{o1} = \frac{\omega_{01}}{s}\left[v_{in} - \beta v_{out}\right] = \frac{\omega_{01}}{s}\left[1 - \beta STF(s)\right] v_{in} \tag{3.46}$$

from which we obtain

$$STF_1 = \frac{v_{o1}}{v_{in}} = \beta_2 A_0 \frac{1 + \dfrac{s}{\omega_{02}} + \dfrac{s^2}{\beta_2 \omega_{02} \omega_{03}}}{1 + \beta_2 A_0 \dfrac{s}{\omega_{01}} + \beta_2 A_0 \dfrac{s^2}{\omega_{01}\omega_{02}} + A_0 \dfrac{s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.47}$$

In the same way we proceed for v_{o2}:

$$v_{o2} = \frac{\omega_{02}}{s} \left[v_{o1} - \beta v_{out} \right] = \frac{\omega_{01}}{s} \left[STF_1 - \beta STF(s) \right] v_{in} \tag{3.48}$$

from which we obtain

$$STF_2 = \frac{v_{o2}}{v_{in}} = \beta_2 A_0 \frac{1 + \dfrac{s}{\beta_2 \omega_{03}}}{1 + \beta_2 A_0 \dfrac{s}{\omega_{01}} + \beta_2 A_0 \dfrac{s^2}{\omega_{01}\omega_{02}} + A_0 \dfrac{s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.49}$$

Figure 3.10 shows the magnitude of these frequency responses. STF is shown as well for comparison, scaled by a factor β_2 so that all the functions have the same DC gain of 40 dB. As we can see, neither STF_1 nor STF_2 have particularly high resonance peaks (STF_2 reaches only 40.2 dB around f_c), so this problem is definitely not relevant.

STF with different input common modes A final consideration is due about what to expect when input CM changes. As mentioned before, a R2R input CM range needs two differential pairs for each input port (a detailed description is object of chapter 4.3): these work together when input CM is in a middle region of the input range; if it is near ground, only the p pairs will be active, while if it is near the positive rail only the n pairs will be active. This way, if G_m is one pair's transconductance, when they work together the equivalent tranconductance will be $2G_m$. In other words, with an input CM near the rails, the first stage's overall G_m will be half of the one that we would have with an input CM in the middle of the range.

Since the target of this work is to deal with $v_{inCM} \approx V_{DD}/2$, the component sizing will aim to achieve a G_{m1-eq} (eq. 3.32) for the desired input CM. When input CM is near the rails, then, we will have half the trasconductance, that is a $G^{(h)}_{m1-eq} = G_{m1-eq}/2$. Thus, capacitances being the same, ω_{01} will be half the original value, while ω_{02} and $\omega 03$ do not change: this

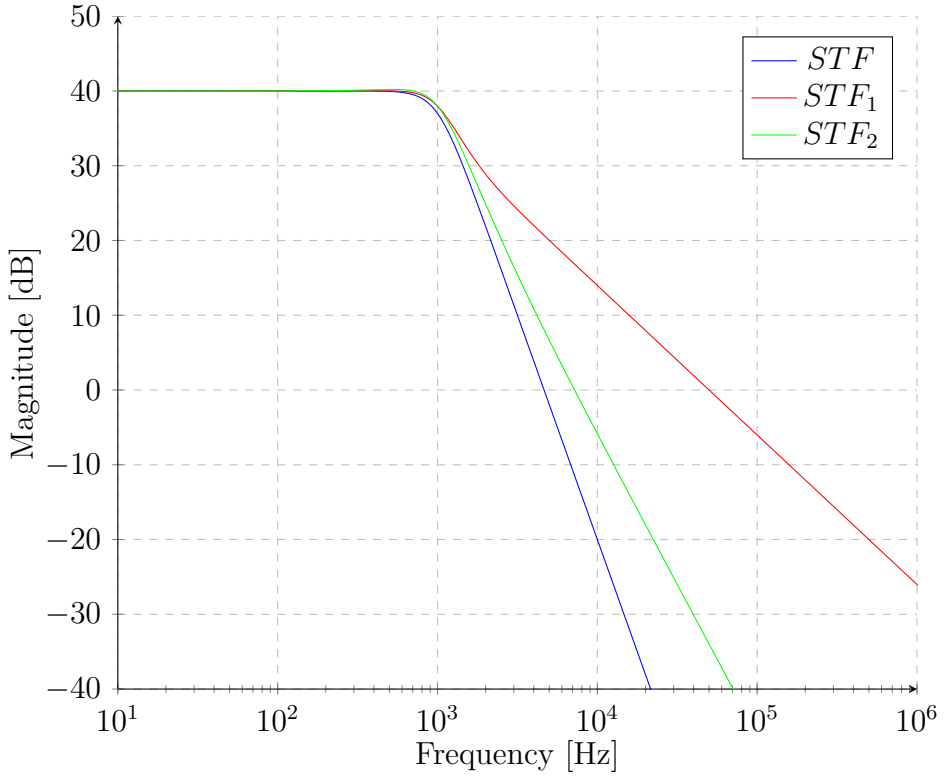


Figure 3.10: Partial Signal Transfer Functions.

way, the closed loop frequency response may change, since, for instance, this would not be a Butterworth filter anymore. So, we can define a STF_h with all ω_{01} halved, as follows:

$$STF_h = \frac{A_0}{1 + \beta_2 A_0 \frac{2s}{\omega_{01}} + \beta_2 A_0 \frac{2s^2}{\omega_{01}\omega_{02}} + A_0 \frac{2s^3}{\omega_{01}\omega_{02}\omega_{03}}} \tag{3.50}$$

We can then generate it with the usual Python functions and compare the graphics (figure 3.11). The last two code lines are used to find the cutoff frequency, that is the frequency at which the frequency response is $-3\,\text{dB}$ compared to the DC gain (in this case $STF(0) - 3 = 43\,\text{dB}$).

```
a_h = 2 * a
a_h[3] = 1
w, mag_h, ph_h = signal.bode((b, a_h),w)
fc = w[find_nearest(mag, (mag[0]-3))]/(2*np.pi)
fc_h = w[find_nearest(mag_h, (mag[0]-3))]/(2*np.pi)
```

We have that, obviously, the frequency response is different: the new

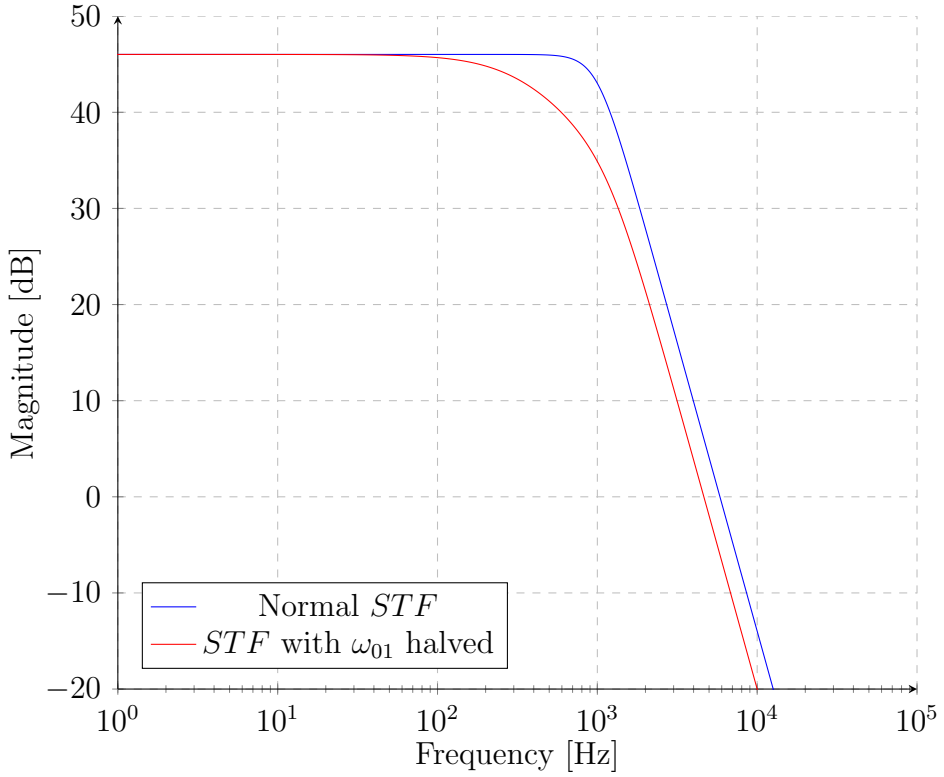


Figure 3.11: STF comparison with different ω_{01}.

one has a softer transition from inside to outside the band, so it begins to decrease it magnitude for lower frequencies. Trying to quantify this loss, we compare the two cutoff frequencies, found with the last two Python code lines. We have:

$$f_c = 1\,\text{kHz} \qquad\qquad f_{c-h} = 355\,\text{Hz} \tag{3.51}$$

While this is quite a huge loss in terms of bandwidth, we must notice that bandwidth specifications are still met, especially if compared with the reference design that had only a 200 Hz cutoff frequency. As we expected, DC gain has not changed, since it depends only on the feedback network β.

This peculiarity of the proposed solution could, however, give rise to a different problem. Indeed, suppose that a large, low frequency common mode disturbance is present at the input (for example, the 50 Hz mains voltage): if a signal at relatively high frequency (a few hundreds Hz) is

presented at the same time at the differential input, this signal will be amplified with a gain that depends on the instantaneous value of the input CM voltage. Thus, a sort of intermodulation between common mode and differential mode is expected, and has to be tested.

Chapter 4

Transistor level design

4.1 Overall schematic

After a detailed description of the high level behaviour of the proposed architecture, now a discussion on the electrical implementation of the single blocks is given. A detailed description of each block's topology, up to the dimensions of single transistors, will be given in next parapragh. Before that, an overview of the complete system is useful in order to understand the role of each block within the whole architecture: figure 4.1 shows the block diagram, where all electrical signals are represented and all the blocks correspond to the actual electrical subcircuits. As we can see, all blocks except CMDA and CMFB are fully differential.

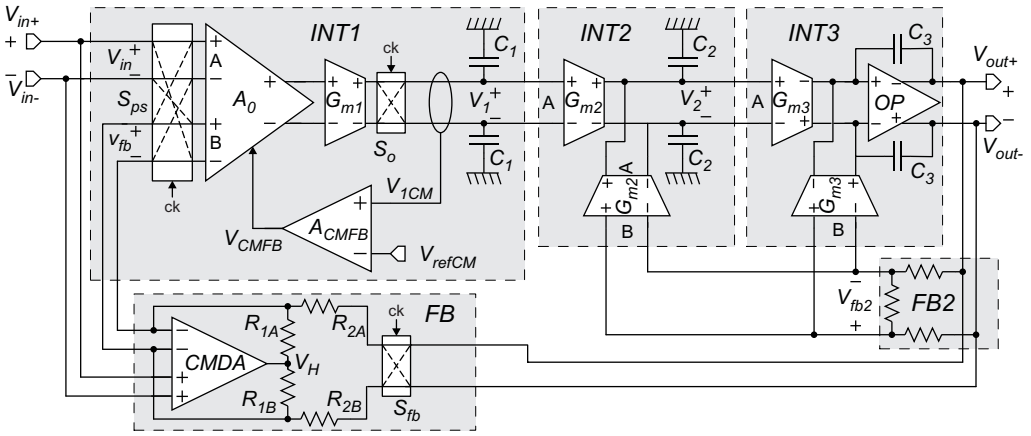


Figure 4.1: Block schematic of the proposed InAmp.

Some key points of this implementation should be considered:

- The main error reduction techniques used are substantially the same as in the reference design. In particular, **chopper modulation** and **port swapping** is implemented, as evident from the switch modulators S_o, S_{fb} and S_{ps}. Since chopping and port swapping are driven by the same clock frequency $f_{ck} = 20\,\text{kHz}$, the input modulator S_{ps} can swap both polarities and ports together with a simplified mechanims, as we can see from the lines inside its symbol: some in-out paths are not necessary, so a considerable number of switches can be saved.

  Finally, block FB shows how Common Mode Equalization (CMEQ) is achieved: in this case as well, there are no significant differences with the reference design, so no additional explanation is necessary.

- The internal structure of the integrators is now evident. As said, INT1 is formed by a preamplifier and the actual integrator: the former is actually a Difference Differential Amplifier (DDA), while the latter is a Gm-C integrator, that is an OTA (called OTA1) followed by capacitors. INT2 as well is a Gm-C integrator: since it does not need a preamplifier, its OTA (called OTA2) must have two input ports in order to deal with input and feedback voltages. Finally, INT3 is in Gm-OpAmp configuration, since a low output resistance was needed to drive the resistive feedback network. OTA3 as well presents two input ports.

- As said, the preamplifier is actually a DDA. A simple yet effective way to implement such device is with a two input port OTA, followed by a resistive load, as shown in figure 4.2.

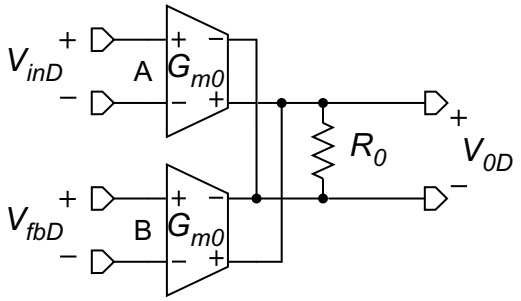


Figure 4.2: Block schematic of the preamplifier.

Evidently, since $v_{0D} = \frac{R_0}{2} I_{out-preD} = \frac{R_0}{2} G_{m0}(v_{inD}+v_{fbD})$, the resulting differential voltage gain will be

$$A_{pre} = G_{m0}\frac{R_0}{2} \tag{4.1}$$

It must be noticed that, with this topology, A_{pre} is not function of homogeneic quantities, so it will be affected by global process errors instead of matching errors. As known, these can be very high and gain will most probably will be affected by a significant error. Anyway, this error would be reflected only on OTA1 RTI noise and on G_{m1-eq}, that in turn influences only INT1 unity-gain frequency ω_{01}: while the overall frequency response may change around f_c, this will not affect the accuracy of the in-band gain A_0.

- The signal path has been widely discussed in the previous chapter, so the expected differential voltages around the circuit are known. However, common mode voltages must be controlled as well, to guarantee the proper operation of the circuit. In this architecture, the idea is to control OTA1 output CM voltage and let it be copied to the following stages up to the output. So, first of all, a **Common Mode Feedback** (CMFB) loop is created within INT1, in order to stabilize the preamplifier's and OTA1's output common mode voltages to proper values. In particular, OTA1 CM output v_{1CM} was set to $V_{DD}/2$, in order to maximize the output swing. Then, the following stages are designed to feature a high common-mode to common-mode gain A_{cc}: this way, since the feedback network β_2 does not change the common mode ($\beta_{2CM} = 1$), a strong negative CM feedback sets INT2 and INT3 closed loop CM gain to

$$A_{cc-loop2,3} = \frac{A_{cc}}{1 + \beta_{2CM} A_{cc}} \approx \frac{1}{\beta_{2CM}} = 1 \tag{4.2}$$

  This way, we achieve the desired goal of setting $v_{1CM} = v_{2CM} = v_{outCM} = V_{DD}/2$.

4.2 Process and device models

4.2.1 Process components

At this point, it is convenient to discuss the characteristics of the devices used in this work, for a better comprehension of the design choices that will be described in next paragraphs.

In this work, we used the technological process **BCD8S** from STMicroelectronics. For the electrical design of this work, we used the schematic editor of the Cadence Virtuoso suite, that allowed the integration of the design kit. The following components were employed:

M - Transistors are low voltage MOSFETs designed for digital applications. They are designed to work with a 1.8 V supply voltage, and were used since they have better parameters, especially noise-related, than the 5 V analog ones. For this reason, supply voltage V_{DD} was decreased from 3.3 V, used in the refernce design, to 1.8 V. For these devices, the process allows the following minimal channel dimensions:

$$w_{min} = 280\,\text{nm} \qquad\qquad l_{min} = 180\,\text{nm}$$

These values, actually, are fixed in the design steps and are used in the schematic editor: in reality, the pocess used is the shrinked version of BCD8, so before the mask fabrication all the geometries are scaled by a factor 0.92. The simulator takes this effect into account through the parameter `shrfact`.

R - High resistance polysilicon resistors were used. These devices are basically polysilicon lines, insulated from the substrate by the field oxide (that forms a parasitic capacitance) and protected from Silicide and from further doping by a dedicated mask. This way, a high square resistance can be achieved: in our process, we have around

$$R_{\square} \approx 6\,\text{k}\Omega/\square$$

C - Capacitor are basically a MOS strucrture, that is N+ polysilicon, over low voltage gate oxide ($V_{max} = 1.8\,\text{V}$), over a CPC implantation. These devices were chosen since they have the highest capacitance-to-area ratio: as reported by the process manual, we have

$$\frac{C}{A} \approx 8\,\text{nF/mm}^2 \qquad \rightarrow \qquad \frac{A}{C} \approx 100\,\mu\text{m}^2/\text{pF}$$

Furthermore, their capacitance is quite dependant on the voltage drop, and in a non-linear way (file: electrical char, pag 263). Thus, whenever a capacitor was expected to undergo both positive and negative large voltages, an antisymmetric design (figure 4.3) was implemented, in order to eliminate even-order non-linearities and reduce distortions.

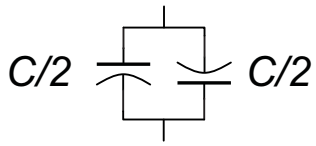


Figure 4.3: Antisymmetric design of capacitors.

4.2.2 Device models

In this section, we aim to summarize the analytical models of the devices used, trying to highlight the most useful equations for the preliminary design steps. A great amount of literature is available on this topic, since more and more accurate models have been proposed. Among these, the most accurate ones (BSIM3 is an example) are used by simulators to evauate the behavior of electronic circuits with high precision: without them, high performance analog design would not be possible.

Anyway, these models rely on complex equations, that on one hand are extremely accurate, but on the other are impossible to deal with in a preliminary stage of the design, when manual component sizing is performed by the designer. So, we will consider a much simpler model, that however allows a first rough estimate of the components' dimensions: these, in most cases, were then adjusted, or sometimes even strongly changed, after the simulations.

Since the most relevant components used in this work are MOSFETs, we will concentrate on the most frequently used equations of this device.

Drain current First of all, it is necessary to distinguish different working regions, classified according to overdrive voltage $V_{OD} = V_{GS} - V_{th}$, that determines the level of *channel inversion*, and to drain-source voltage V_{DS}, that determines the level of *channel saturation*. These regions are then classified as follows:

- Inversion ($V_{GS} - V_{th} = V_{OD}$):
  - Strong inversion: $V_{GS} - V_{th} > 4V_T \approx 100\,\mathrm{mV}$
  - Weak inversion: $V_{GS} - V_{th} \ll 4V_T$ (for $V_{GS} - V_{th} < 0\,\mathrm{V}$ it is also called sub-threshold region)
- Saturation (V_{DS}):
  - Saturation region: $V_{DS} > V_{DSat}$
  - Triode (or linear) region: $V_{DS} < V_{DSat}$

where $V_T = kT/q$ is the thermal voltage and is equal to 25.8 mV at 27 °C, while V_{DSat} is the saturation voltage, and its value depends on the inversion:

$$V_{DSat} = \begin{cases} V_{GS} - V_{th} & \text{in strong inversion} \\ 4V_T \approx 100\,\text{mV} & \text{in weak inversion} \end{cases} \tag{4.3}$$

The equation for a MOSFET in weak inversion region are often too complex to be used, so we will ignore them in this work. As a matter of fact, *all MOSFETs in this work will be considered in strong inversion region*, unless otherwise specified. For the strong inversion region, the following large signal equations are often very useful.

- Drain current in saturation region (parabolic equation):

$$I_D = \frac{\beta}{2}(V_{GS} - V_{th})^2(1 + \lambda V_{DS}) \tag{4.4}$$

  where the parameter β has the following expression:

$$\beta_n = \mu_n C_{ox}\frac{W}{L} \qquad\qquad \beta_p = \mu_p C_{ox}\frac{W}{L} \tag{4.5}$$

  for a n-channel and a p-channel MOSFET, respectively. W and L are the channel dimensions (width and length), C_{ox} is the MOS structure capacitance per area unit, and μ is the carriers mobility in the channel. These parameters can vary from process to process, but usually it is $\mu_n = 3 \sim 5\ \mu_p$.

- Drain current in triode region:

$$I_D = \beta\left[(V_{GS} - V_{th})V_{DS} - \frac{V_{DS}^2}{2}\right] \tag{4.6}$$

- Threshold voltage dependance from body-source voltage (body effect):

$$V_{th} = V_{th0} + \gamma\left(\sqrt{\phi - V_{BS}} - \sqrt{\phi}\right) \tag{4.7}$$

For small signal parameters, in addition, the following relations can be useful.

- Small signal transcharacteristic:

$$i_d = g_m v_{gs} + g_d v_{ds} + g_b v_{bs} \tag{4.8}$$

where, usually, $g_m/g_d \approx 100$ and $g_b/g_m \approx 0.3$.

- Transconductance:

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \beta(V_{GS} - V_{th}) = \sqrt{2\beta I_D} = \frac{2I_D}{V_{GS} - V_{th}} \tag{4.9}$$

  These relationships are useful to express g_m as function of any two of the three parameters, that are β (process/dimensions), I_D (drain current) and V_{OD} (overdrive voltage). If we define an *equivalent thermal voltage* $V_{TE} = (V_{GS} - V_{th})/2$, we can write a general expression for g_m that is always valid for both MOSFETs (even in weak inversion) and BJTs, simply by changing the definition of V_{TE} in each case:

$$g_m = \frac{I_D}{V_{TE}} \quad \text{with} \quad V_{TE} = \begin{cases} \dfrac{V_{gs} - V_{th}}{2} & \text{strong inversion} \\ \xi V_T \approx 35\,\text{mV} & \text{weak inversion} \\ V_T \approx 25\,\text{mV} & \text{bipolar} \end{cases} \tag{4.10}$$

- Output resistance:

$$r_d = \frac{1}{g_d} = \frac{1}{\dfrac{\partial I_D}{\partial V_{DS}}} = \frac{1}{\lambda I_D} = \frac{\lambda^{-1}}{I_D} \tag{4.11}$$

  The process parameter λ^{-1} is proportional to the effective channel length L_{eff} and, as we can see by this equation, is the MOSFET equivalent of the Early voltage V_A for bipolar transistors.

As we have seen, some process-dependant parameters compare in these relations, so an estimation of them is needed. In particular, V_{th} and μC_{ox} were considered for both nMOS and pMOS. For the former, the process manual gives some information, while no trace of the latter is present: in both cases, testbenches were set in order to evaluate them in correspondence of different bias points. All of them seemed quite constant with the bias point and with the channel dimensions (the values given below are averages), except for nMOS threshold voltage V_{th-n}, that showed a strong dependance from channel length. These average values were found (just a rough estimation):

$$V_{th-p} \approx 500\,\text{mV} \qquad V_{th-n} \approx \begin{cases} 500\,\text{mV} & L \approx 1\,\mu\text{m} \\ 400\,\text{mV} & L \approx 10\,\mu\text{m} \end{cases} \tag{4.12}$$

$$\mu_p C_{ox} \approx 50 \times 10^{-6}\,\text{A/V}^2 \qquad \mu_n C_{ox} \approx 250 \times 10^{-6}\,\text{A/V}^2 \tag{4.13}$$

Noise As introduced in the previous chapters, a MOSFET is mostly affected by thermal noise and flicker noise. At low frequencies, we can neglect the gate current, so the only noise source would be drain current. As said, thermal noise has a flat PSD, since it is a gaussian white noise, while flicker noise has its peculiar $1/f$ slope: we remind their simplified expression for drain current noise PSD.

$$S_{In-Th} = 4kTg_m \qquad S_{In-F}(f) = \frac{N_f}{W_{eff}L_{eff}}\frac{1}{f}g_m^2 \tag{4.14}$$

Actually, a more frequently used expression has a coefficient $(8/3)(1+n)$, with $n \approx 0.3$, intstead of 4. However, the numerical value of these coefficients is almost equal (in fact 4 is slightly higher than the other, providing an overestimation of the noise); in addition, with a coefficient 4, the expression becomes equal to the one of the thermal noise of a resistor with $R = 1/g_m$.

If preferred, a RTI voltage noise PSD expression can be obtained dividing eq. 4.14 by a factor g_m^2:

$$S_{Vn-Th} = \frac{4kT}{g_m} \qquad S_{Vn-F}(f) = \frac{N_f}{W_{eff}L_{eff}}\frac{1}{f} \tag{4.15}$$

We remind, finally, that corner frequency f_k is defined as that frequency at which thermal and flicker noise PSDs are equal:

$$S_{Vn-F}(f_k) = S_{Vn-Th} \tag{4.16}$$

All terms in these relations have already been mentioned: this leaves us to find an estimation of N_f in this process.

The manual reports a value for nMOSFETs only, that is

$$N_f = 5.62 \times 10^{-10}\,\text{V}^2\mu\text{m}^2$$

Anyway, with a dedicated testbench N_f has been evaluated for both n and p MOSFETs, and an unusual dependance of N_f from the overdrive

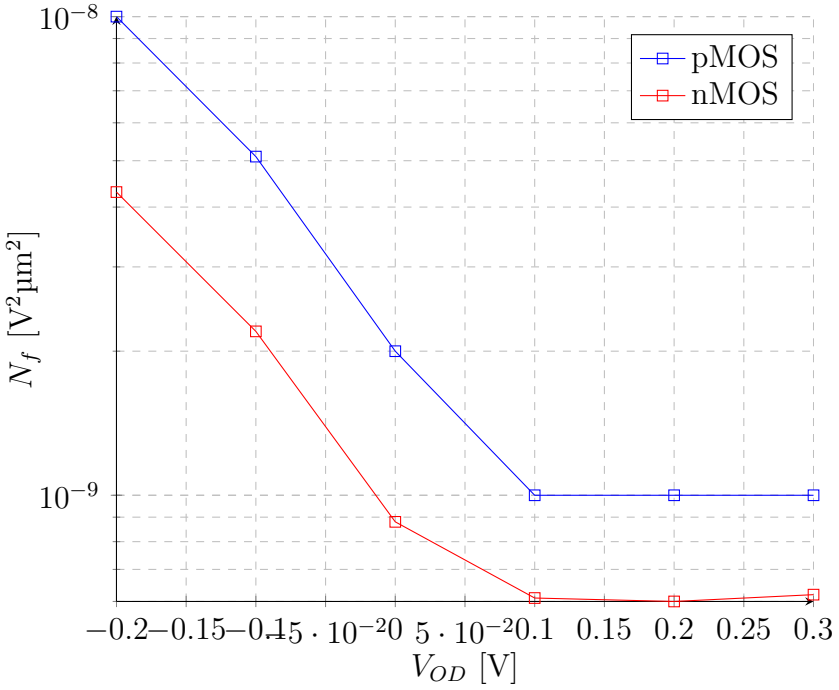


Figure 4.4: N_f values vs. V_{OD} for p and n MOSFETs.

voltage was found. Basically, N_f was near the given value for nMOSFETs in strong inversion, while weak inversion leads to a progressive increase of N_f. The same effect affected pMOSFETs, with slightly higher values at all overdrive voltages. Figure 4.4 shows a plot of the values found from the testbench ($W = L = 100\,\mu\text{m}$, $V_{DS} = 0.95\,\text{V}$).

It must be said, actually, that this increment of flicker noise observed in sub-threshold MOSFETs might be simply due to an inaccuracy in the model of the devices. In any case, since we do not have any other characterization, we must rely on the models used by the simulator.

4.3 Preamplifier

4.3.1 Topology and input range

As said, this preamplifier is actually a DDA, which can be easily designed as a two-input-port, fully differential OTA with a resistive load, as shown in figure 4.2, whose differential gain will be

$$A_{pre} = G_{m0} R_0 \tag{4.17}$$

It must be noticed that, for a certain value of $A_{pre} = 50$, we have now one additional degree of freedom (DoF). We will prove that G_{m0} is strongly linked to the RTI thermal noise: thus, in this case, noise considerations will lead us to a value for the transconductance, then we will find the consequent value of R_0.

For the choice of the topology, some architectural considerations must be made. First of all, as basic element for an OTA we have a differential pair that operates the differential voltage-to-current conversion. Since we aim to a Rail-To-Rail input CM range, both n and p pairs must be employed together and their output currents must be added one to the other. However, it is known that the output nodes (drains) of a n pair are usually kept at a voltage that is near the positive rail, while those of a p pair are kept near ground. This fact, which is dictated by input CM range considerations, prevents a direct sum of the currents by simply connecting the output nodes together. This problem is solved by the **folded cascode** topology: common gate transistors create low impedance nodes, one near V_{DD} and one near ground, that are perfect to direct the pairs' output currents towards the same output nodes. Each of these two low impedance nodes are suitable to be connected to more than one pair of the same kind (n or p), since all the output currents will flow towards the common gate's source terminals, thanks to their low impedance. This is how we can sum together the output currents of the input port's pair and the feedback port's pair. For this reason, the subsircuit that contains the common gate transistors is often called *summing stage*.

A relevant problem for R2R folded cascode architectures is the **biasing** of pairs that, depending on input CM, could turn off. To better understand the problem, consider the simplified, folded cascode n pair shown in figure 4.5.

As we can see, the current flowing in each branch of the pair is I_0, while for each common gate is I_1 (usually the two pairs of generators are merged into a single pair of generators that supplies $I_0 + I_1$). When input CM decreases towards ground, the pair's tail (the MOSFET under the pair that normally sinks $2I_0$) exit saturation region and progressively nulls its current. The two I_0 currents sourced from above then can do nothing but flowing into the common gates and hence to the output nodes, altering the output CM voltage, most probably beyond the capability of the output CM control loop.

Thus, **self-biased pairs** are used [20]. See figures 4.6a and 4.6b: each

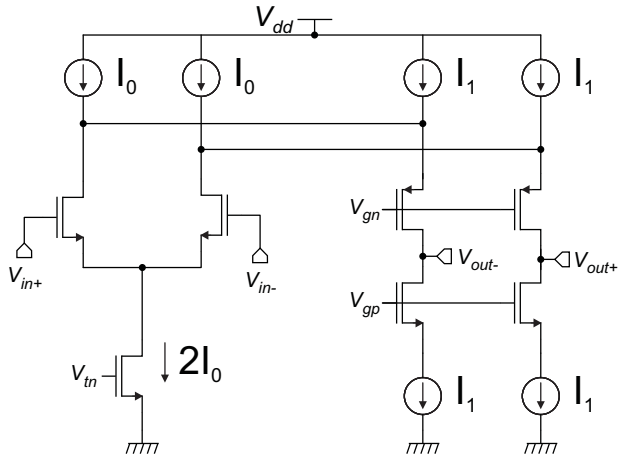
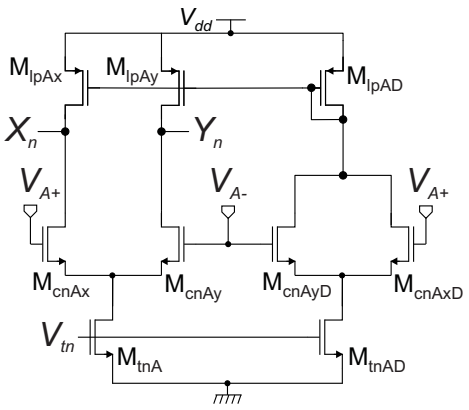


Figure 4.5: Classical folded cascode topology.

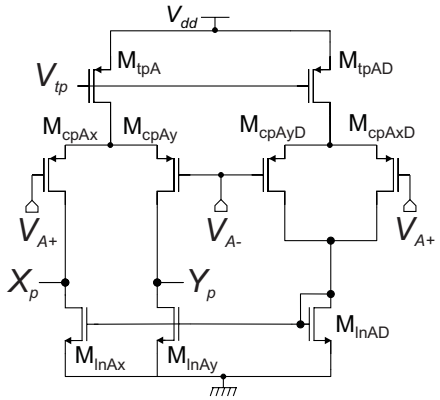
pair is accompanied by a dummy pair, which reproduces on its tail the current variations of the main tail, and mirrors this current to the main pair branches from above. This way, if the input CM changes the bias current of a pair, no current in excess flows into the common gates, thus producing only minor effects (due mainly to mismatch) on the output CM voltage. Figures 4.6a, 4.6b and 4.6c show, respectively, a n pair, a p pair (self-biasing as well) and the summing stage.

A brief digression on the meaning of the labels. For components and bias current or voltages, upper case letters are used (M for MOSFET). Than, the first subscript letter means the role of a component in a circuit: c for pair ("coppia" in italian), l for load, m for mirror, t for tail, g for common gate, s for common source (not present in this block). The second letter is the kind of MOSFET, n or p. The third is the relative port, A or B (they are equivalent). The fourth, x or y, distinguish left side from right side of a circuit, while D stands for Dummy in self-biasing pairs and H means a helpful device for a certain goal (e.g., loop stabilization). The pairs shown are relative to input port A: obviously, an exact copy of these pairs is actually implemented, with the same connections, for port B.

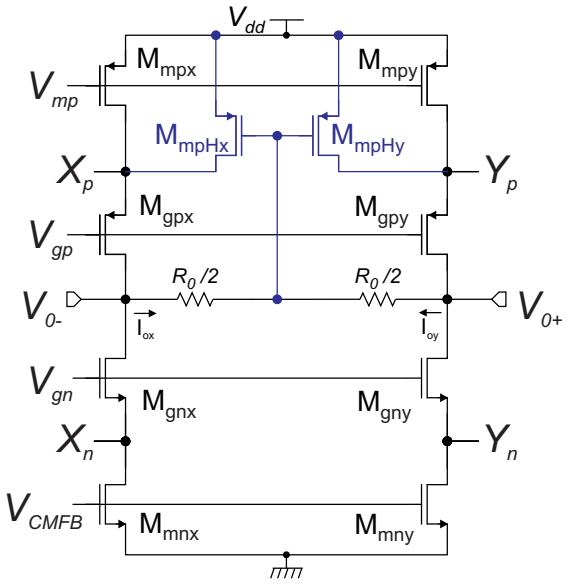
First of all, a simple consideration about bias currents can be made, in order to find an optimum value for I_1, at least compared to I_0 (figure 4.5). For this purpose, let's suppose that $v_{inD} = v_{Dmax}$, that is M_{cny} is off and the whole $2I_0$ sinked by the tail flows through M_{cnx}. This current must be taken from I_1, that will have to be greater that I_0. The excess current $I_1 - I_0$ then flows into M_{gpix}. Similarly, we can find that $I_1 + I_0$ flows through M_{gnx}, thus the output current turns out to be $I_{o-} = -2I_0$, independent from I_1. This means that the current excess $I_1 - I_0$ is useless and only adds current noise



(a) Self-biasing n pair.



(b) Self-biasing p pair.



(c) Cascode summing stage.

Figure 4.6: Schematic of the preamplifier

at the output. Thus, the optimum value is $I_1 = I_0$. Actually, since both input and feedback pairs can work independently, we must set $I_1 = 2I_0$, or, with the new notation,

$$I_m = 2I_l = I_t \tag{4.18}$$

We have actually $I_m = I_{mn} = I_{mp} + I_{mpH}$: transistors M_{mpH} contribute to the bias current of the summing stage, but they also help the stabilization of the CMFB, as described in paragraph 4.5.4. In addition, we can define $I_R = I_{oy} - I_{ox}$, such that $V_{0D} = R_0 I_R$. If a single pair worked, its

current contributions to I_R would be, as usual, $g_{mc}v_{inD}$, where g_{mc} is the transconductance of each MOSFET in any of the pairs. Since for each input port two pairs are active (if input CM is adequate), the resulting current will be double: hence we find the important relation

$$G_{m0} = 2g_{mc} \tag{4.19}$$

4.3.2 Thermal noise

For a more accurate noise estimation, it is necessary to study how the noise currents are presented at the output nodes, and we will consider only the ones that affect the differential output. It can be proved that:

- currents from each of M_c, M_l, M_m and M_g flow into only one of the output nodes, thus affecting the differential output;
- among these, M_g are common gates, so their noise current will flow into the output nodes reduced by a factor $(g_m r_d) \approx 100$, so it can be neclected;
- current from M_t, with no input signal, is equally divided between the two branches, so is a common mode disturb. Actually, when a large signal is applied, current is no more equally divided and this causes a differential disturb. Though, this noise will be proportional to input signal, and it will be relevant only with large input signals: we will neglect this contribution.
- currents from dummy pairs are, of course, common mode disturbs.

So, eventually we find:

$$i_{n-out} = 8i_{n-c} + 8i_{n-l} + 4i_{n-m} \tag{4.20}$$

Since these noise generators are independent, so incorrelated, the total output current PSD (that is the variation of the gaussian distributions) is simply the sum of the single noise current PSDs:

$$S_{In-Th} = 8S_{Ic} + 8S_{Il} + 4S_{Im} = 4kT\left[8g_{mc} + 8g_{ml} + 4g_{mm}\right] \tag{4.21}$$

Then we divide by G^2_{m-pre} to find RTI voltage noise PSD, and using eq. 4.19 we get

$$S_{Vn-Th} = \frac{S_{In-Th}}{(2g_{mc})^2} = \frac{4kT}{g_{mc}}\left[2 + 2\frac{g_{ml}}{g_{mc}} + \frac{g_{mm}}{g_{mc}}\right] = \frac{4kT}{g_{mc}}\left[2 + 2F_l + F_m\right] \tag{4.22}$$

Where the following parameters have been defined:

$$F_l = \frac{g_{ml}}{g_{mc}} = \frac{I_{Dl}}{I_{Dc}}\frac{V_{TEc}}{V_{TEl}} = \frac{V_{TEc}}{V_{TEl}} \tag{4.23}$$

$$F_m = \frac{g_{mm}}{g_{mc}} = \frac{I_{Dm}}{I_{Dc}}\frac{V_{TEc}}{V_{TEm}} = 2\frac{V_{TEc}}{V_{TEm}} \tag{4.24}$$

where equations 4.10 and 4.18 were used, and the evident fact that $I_{Dl} = I_{Dc}$ when no input signal is applied. These equations will be the starting point for the final sizing of the components.

4.3.3 Flicker noise

A similar expression can be found for flicker noise. From equations (4.14) and (4.21) we obtain (notation W_{eff} is omitted)

$$S_{Vn-F} = \frac{N_f}{(WL)_c}\frac{1}{f}\left[2 + 2t_l F_l^2 + t_m F_m^2\right] \tag{4.25}$$

where coefficients t are defined as the inversed ratio between the areas:

$$t_l = \frac{(WL)_c}{(WL)_l} \qquad\qquad t_m = \frac{(WL)_c}{(WL)_m} \tag{4.26}$$

4.3.4 Input and output ranges

Other relationships that are useful for a ballpark design are related to both differential an common mode input and output ranges. We leave out the output CM, since it will be stabilized by a dedicated control loop.

Output differential range First of all, it must be noticed that the voltage of nodes X and Y is almost constant, since output voltage variations are masked by common gates. So, considering V_{Xn} and V_{Yn} constant (low side is considered, the same is valid for the high side), we have that the output stage stops working correctly when common gates exit their saturation region. So the output stage works as long as

$$V_{DSg} = V_{0x} - V_{DSm} > V_{ODg} \quad \rightarrow \quad V_{0x-min} = V_{ODg} + V_{DSm} \tag{4.27}$$

Now, we simply have $V_{DSm} = V_{Kg} - V_{GSg}$, so, given a drain current, V_{GSg} is set and V_{Kg} sets V_{DSm}. In addition, we have to guarantee the saturation of M_m: since their V_{DS} is almost constant, we can make them work always at the limit of saturation region, that is

$$V_{DSm} = V_{ODm} \tag{4.28}$$

Finally, we get an expression for minimum output voltages (for branch y is the same, for symmetry):

$$V_{0x-min} = V_{ODm} + V_{ODg} \tag{4.29}$$

Now, as said, at low frequencies, thanks to virtual short circuit differential output voltage will be almost zero, so there will be no range-related problems. However, consider the example of a high frequency signal, whose phase is shifted by 90° (it happens at $f = 700\,\text{Hz}$: when the input is at a maximum, the feedback signal will be null. As maximum input value, we expect $v_{inDmax} = V_{outDmax}/A_0 \approx 8\,\text{mV}$: then we will have

$$V_{0x-min} = v_{0CM} - A_{pre}\frac{v_{inDmax}}{2} \approx 0.7\,\text{V} \tag{4.30}$$

supposing a $v_{0CM} = V_{DD}/2$.

Input CM range Consider a n pair: we would like input CM to reach the positive rail. From the schematic, we can easily obtain the following expression:

$$v_{inCM} - (V_{ODc} + V_{thc}) = V_{DD} - V_{DSl} - V_{DSc} \tag{4.31}$$

Now, like for mirrors, we can set $V_{DSl} = (V_{DD} - V_{Kgp}) = V_{ODl}$ and make them work at limit of their saturation region. This voltage is almost constant, so the pair works as long as M_{cn} is in saturation region, that is $V_{DSc} > V_{OD}c$. We get then, after a few algebraic steps, the condition

$$v_{inCM} < V_{DD} - (V_{ODl} - V_{thc}) \tag{4.32}$$

If the term between brakets is negative, input CM range includes the positive rail.

On the other hand, when input CM decreases, the tail's voltage decreases as well, until it will eventually exit the saturation reason. A condition for an n pair's good operation is then

$$v_{inCM} > V_{ODc} + V_{thc} + V_{ODt} \tag{4.33}$$

These value must be less than $V_{DD}/2$, otherwise for $v_{inCM} = V_{DD}/2$ none of the two pairs would work (supposing a symmetric sizing of p pairs). The lower this value, the wider the input CM range in which both pairs work.

4.3.5 Final design and dimensions

All the equations obtained so far give us costraints for sizing I_{DC} and all overdrive voltages. In particular:

- I_{Dc} will be a trade-off between power consumption and noise. It is a critical parameter, since the overall current consumption of the preamplifier will be $12I_{Dc}$ (2 MOSFETs per pair, 4 pairs and 2 branches in the summing stage, each with double current).
- V_{ODc} should be as low as possible. However, we have seen that flicker noise in deep sub-threshold strongly increases, so extremely large areas would be necessary to contain it: it may not be convenient.
- Other overdrive voltages should be as high as possible, but they limit input and output ranges.

A good trade-off for overdrive voltages was found by setting the following values:

- $V_{ODm} = V_{ODl} = 400\,\text{mV}$
- V_{Kg} such that $V_{DSm} = V_{DSl} = 450\,\text{mV}$, to assure saturation of loads and mirrors. This way, since $V_{thc} > 500\,\text{mV}$ thanks to body effect, rails are included in input CM range.
- $V_{ODg} = 100\,\text{mV}$, since their noise is negligible, and to guarantee an output swing of $0.9 \pm 0.35\,\text{V}$. This guarantees linearity for all possible signals, even very large, at all frequencies. That is, a maximum differential output $v_{0Dmax} = 700\,\text{mV}$ is achieved with a $v_{0CM} = 0.9\,\text{V}$.
- $V_{ODc} = 100\,\text{mV}$ to assure strong inversion, to be descreased a bit if necessary.

In other terms, coefficients F were set to be $F_l = 1/4$ and $F_m = 1/2$, so that thermal noise is 4/6 from pairs, 1/6 from loads and 1/6 from mirrors. Then, if we set a target noise of $0.81 \times 10^{-16}\,\mathrm{V^2/Hz}$, we get

$$S_{Vn-Th} = 3\frac{4kT}{g_{mc}} = 0.81 \times 10^{-16}\,\mathrm{V^2/Hz} \tag{4.34}$$

From this we obtain a value of $g_{mc} = 613\,\mu\mathrm{A/V}$ that, with a $V_{TEc} = 50\,\mathrm{mV}$, gives a drain current of $30\,\mu\mathrm{A}$. However, this solution brings a few problems. First of all, current consumption was really high; secondly, in order to obtain such g_{mc} with such current, MOSFETs had to work in deep sub-threshold (evidently our simple equations were incorrect in this operation region), so the flicker coefficients were much higher than its normal values. The target for flicker noise was to set to be $S_{Vn-F}(f_{ch}) = 0.5\ S_{Vn-Th}$ (that is, a corner frequency $f_k = 0.5\ f_{ch}$): but with these coefficients, extremely large areas would have been necessary for this purpose. A much more efficient design has been made after a few simulations, with the following values:

- $I_{Dc} = 20\,\mu\mathrm{A}$, for an overall current consumption of around $260\,\mu\mathrm{A}$
- $g_{mc} = 350\,\mu\mathrm{A/V}$, obtained with an overdrive of $42\,\mathrm{mV}$ (nMOS) and $7\,\mathrm{mV}$ (pMOS).
- resulting RTI thermal noise: $S_{Vn-Th} = 1.26 \times 10^{-16}\,\mathrm{V^2/Hz}$
- desired flicker noise $S_{Vn-F}(f_{ch}) \approx 0.5\ S_{Vn-Th} = 0.58 \times 10^{-16}\,\mathrm{V^2/Hz}$, achieved with reasonably large areas (flicker noise contributes were set to be around 4/6 from pairs, 1/6 from loads and 1/6 from mirrors, as for the thermal noise).
- total RTI noise density (noise and flicker) $v_{n-pre}(f_{ch}) = 13.58\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$.
- Load resistance has a consequent value of $R_0 = 2\dfrac{A_{pre}}{G_{m0}} \approx 146\,\mathrm{k\Omega}$

Table 4.1 shows the devices' dimensions (bias stage and dummies excluded). An integer factor is often present to multiply width: this is the value of parameter `m` (multiplier) of the simulator, that means that a device is actually implemented as a parallel of `m` devices. This is equivalent to a single MOSFET with a channel width `m` times larger.

The total area of the preamplifier (only active areas, dummy pairs and bias circuitry included) is $49\,300\,\mu\mathrm{m^2} = 0.0493\,\mathrm{mm^2} \approx (222\,\mu\mathrm{m})^2$.

Device	W/L [µm/µm]	Device	W/L [µm/µm]
M_{cn}	$10 \cdot 26/9.8$	M_{mn}	$10 \cdot 3.4/16$
M_{lp}	$5 \cdot 10/10$	M_{gn}	$100/5$
M_{tn}	$10 \cdot 4/5$	M_{gp}	$270/3$
M_{cp}	$10 \cdot 114/5.6$	M_{mp}	$10 \cdot 7.6/9$
M_{ln}	$5 \cdot 3.5/17$	M_{mpH}	$10 \cdot 1/9$
M_{tp}	$10 \cdot 10/2.5$		

Table 4.1: Preamplifier devices dimensions.

4.4 OTA1

4.4.1 Topology

We now proceed to the design of the transconductor of the first integrator. As mentioned before, we aim to obtain a $\omega_{01} = 2\pi\ 50\,\text{kHz}$, so with capacitors $C_1 = 65\,\text{pF}$ we need a tranconductance of

$$G_{m1} = 408.4\,\text{nA/V} \tag{4.35}$$

This value is too low to be achieved with classical differential pairs; furthermore, their input differential range is usually too narrow. In fact, we have said that for low frequency signals the preamplifier's input (and so its output as well) will be almost zero, but for large signals at high frequencies a large differential voltage can be presented at the preamplifier's output: for this purpose, we have designed it with an output voltage swing from 0.55 to 1.25 V, that is $v_{0Dmax} = 700\,\text{mV}$ and we would like to have a comparable input CM range.

Since the preamplifier's CM output can (and will) be stabilized by a control loop, we can implement this OTA with a **pseudo-differential pair**. In addition, since we need a high output impedance in order to have a good current output, a **folded cascode** topology is employed. Figure 4.7 shows the transistor level schematic of OTA1. Since a low tranconductance is needed, a p pair is implemented, so that, for equal β, we will have a relatively shorter channel length than in the case of an equivalent n pair. S_{op} and S_{on} are chopper demodulators. Instead of using only one demodulator at the output nodes, this solution is preferreable since they are placed in correspondence of low impedance nodes, created by common gates: this way, the impact of the switched capacitor parasitic resistance associated to the modulators (due to swiching back and forth the drain/source parasitic capacitances) is reduced.

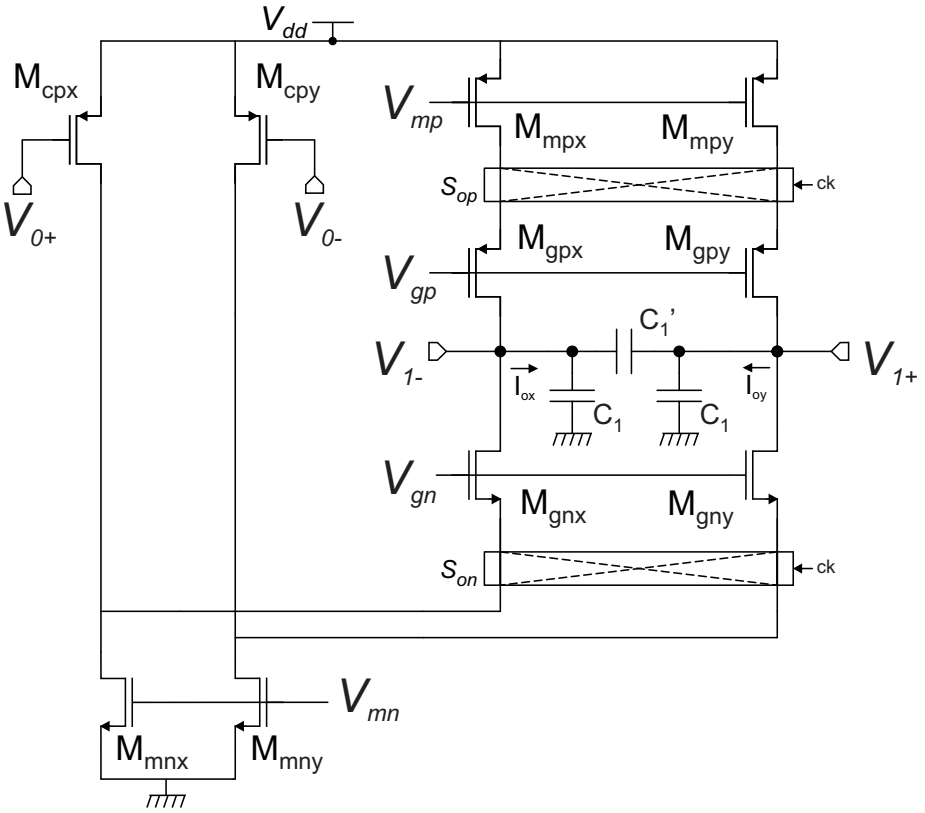


Figure 4.7: Schematic of OTA1.

Again, it must be noticed that the optimum ratio between the quiescent currents is $I_{mp} = I_c$, for the same reasons discussed for the preamplifier.

In addition, we can see that, instead of two common mode 65 pF capacitors, two $C_1 = 25\,\text{pF}$ common mode capacitors and one $C_1' = 20\,\text{pF}$ differential capacitor were used. This way, being equal the effective differential capacitance seen in both cases ($C_{diff} = C/2 + C' = (65/2 + 0)\text{pF} = (25/2 + 20)\text{pF}$), the capacitors' area is strongly reduced. Common mode capacitors were made relatively small but could not be eliminated, since they are needed by the CMFB, as discussed in paragraph 4.5.4.

Furthermore, it is worth finding from the schematic a relationship that links G_{m1} to the device parameters. We can easily obtain the following expression for differential output current:

$$I_{outD} = I_{oy} - I_{ox} = \frac{\beta_p}{2}\left[V_{ODcy}^2 - V_{ODcx}^2\right] = \frac{\beta_p}{2}\left[V_{ODcy} + V_{ODcx}\right]\cdot(V_{ODcy} - V_{ODcx}) \tag{4.36}$$

Now, we have (voltages are negative for pMOSFETs - consider $V_{th} = |V_{th-p}|$):

$$V_{ODcx} = |V_{GScx}| - V_{th} = V_{DD} - v_{0+} - V_{th} \tag{4.37}$$
$$V_{ODcy} = |V_{GScy}| - V_{th} = V_{DD} - v_{0-} - V_{th} \tag{4.38}$$
$$\tag{4.39}$$

Then, combining the last two relationships, we get an expression for G_{m1}

$$G_{m1} = \beta_p \left[V_{DD} - v_{0CM} - V_{th}\right] = \beta_p V_{ODc} = g_{mc} \tag{4.40}$$

This expression, if made equal to the target value of tranconductance, gives us a value for parameter β_p:

$$\beta_p = \frac{G_{m1}}{V_{DD} - v_{inCM} - V_{th}} = 1.02 \times 10^{-6}\,\mathrm{A/V^2} \tag{4.41}$$

where $V_{th} \approx 0.5\,\mathrm{V}$ for a pMOS and a value of $V_{DD}/2 = 0.9\,\mathrm{V}$ has been used for v_{inCM}, in order to maximize the preamplifier's output range (this value will be fixed by the CMFB control loop). From last relationship, then, we can obtain a value for the quiescent current in the pair's transistors, reminding that $V_{ODc} = V_{DD} - v_{inCM} - V_{th} = 0.4\,\mathrm{V}$ and using eq. (4.40):

$$I_{Dc} = \frac{\beta_p}{2} V_{ODc}^2 = \frac{1}{2} G_{m1} V_{ODc} = 81.6\,\mathrm{nA} \tag{4.42}$$

As said, this will be the quiescent current in common gates as well, so it will be the maximum output current for each branch. We will see that this design is not the best, and it will be changed.

4.4.2 Input range

Indeed, a rough estimation of the input differential range is the voltage that would produce the maximum differential output current:

$$v_{inD-max} \approx \frac{I_{outD-max}}{G_{m1}} = \frac{I_{ox-max} - I_{oy-min}}{G_{m1}} = \frac{2I_{Dc}}{G_{m1}} = V_{ODc} \approx 400\,\mathrm{mV} \tag{4.43}$$

Since the preamplifier was designed to have a $v_{outDmax} = 700\,\mathrm{mV}$, this would be the bottleneck for voltage swing: thus, a different design was preferred.

In particular, to increase the input differential range we should increase the overdrive, that is decrease input CM: this was then set to 0.8 V, giving us an input differential range of

$$v_{inDmax} = V_{ODc} = V_{DD} - v_{inCM} - V_{th} = 500\,\text{mV} \tag{4.44}$$

While this is an improvement, obviously the preamplifier's output differential range will be now lower, since

$$v_{outx-pre-min} = v_{outCMpre} - \frac{v_{outDpre-max}}{2} = 550\,\text{mV} = \text{constant} \tag{4.45}$$

In fact, this gives us a value of output range of $v_{outDpre-max} = 500\,\text{mV}$, that perfectly matches the value found for OTA1 input. It is interesting - and easy - to prove that, for our values of $v_{outx-pre-min} = 0.55\,\text{V}$, V_{th}, V_{DD} and G_{m1}, this is the maximum achievable range. In fact, the following linear sistem can be found:

$$\begin{cases} v_{inDmax} = V_{ODc} = V_{DD} - V_{th} - v_{CM} & (4.43) \\ v_{outDpre-max} = 2(v_{CM} - v_{outxpre-min}) & (4.45) \\ v_{Dmax} = \min\{v_{inDmax}, v_{outDpre-max}\} \end{cases} \tag{4.46}$$

The following plot proves graphically that $v_{Dmax} = 0.5\,\text{V}$ is the widest range achievable, and it is achieved with $v_{CM} = 0.8\,\text{V}$.

It must be reminded that this relationships provide a rough estimation of input differential range, that only a simulator can evaluate precisely.

We now proceed to evaluate the current consumption. Eq. (4.42) gives

$$I_{Dc} = \frac{1}{2} G_{m1} V_{ODc} = 102\,\text{nA} \tag{4.47}$$

for a total current consumption of $I_{DD} = 4I_{Dc} = 404\,\text{nA}$.As expected, this value is negligible, compared to the preamplifier's current consumption. Actually, it must be said that, after simulations, we found a value $I_{Dc} = 110\,\text{nA}$ for a resulting $G_{m1} = g_{mc} = 410\,\mu\text{A/V}$.

4.4.3 Output range

Since the output stage is a folded cascode like in the preamplifier, similar considerations can be made. Now, we must consider that the output voltage of OTA1 will follow STF_1, that had a DC gain of $\beta_2 A_0$: that is, an output differential range of $V_{out-max}/2 \approx 0.8\,\text{V}$ must be achieved. This, supposing that CMFB sets OTA1 CM output at a value of $v_{1CM} = 0.9\,\text{V}$ to maximize

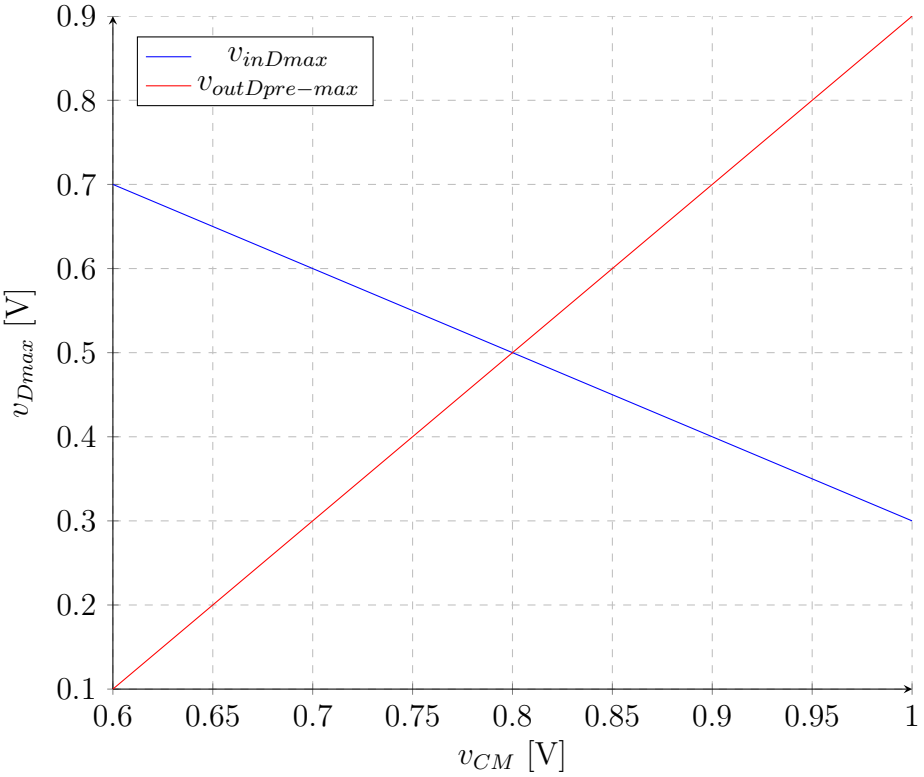


Figure 4.8: Ranges intersection.

the range, means that each branch of output stage must have a swing of

$$v_{1x} = (0.9 \pm 0.4)\text{V} \tag{4.48}$$

This range is determined by common gates' saturation voltages and by loads and mirrors' drain-source voltages, as in the preamplifier, that in turn cannot be made too low for noise considerations. Thus, they have been set as the maximum voltages that met range specifications: in particular, we have

- $V_{ODm} = 370\,\text{mV}$
- $V_{DSm} = 390\,\text{mV}$ to guarantee mirrors' saturation
- $V_{ODg} = 100\,\text{mV}$, since common gates' noise is negligible compared to the other contributes

With these voltages, an output differential voltage $v_{1Dmax} = 0.8\,\text{V}$ is achieved.

4.4.4 Noise

For both thermal and flicker noise, the same approach used for the preamplifier can used. In this case, for the former we get

$$S_{Vn-Th} = \frac{4kT}{g_{mc}}\left[2 + 2\frac{g_{mmn}}{g_{mc}} + 2\frac{g_{mmp}}{g_{mc}}\right] = 2\frac{4kT}{g_{mc}}\left[1 + 2\frac{V_{TEc}}{V_{TEmn}} + \frac{V_{TEc}}{V_{TEmp}}\right] \tag{4.49}$$

This time, all variables in this equation are fixed. The RTI thermal noise PSD, according to this equation, is $S_{Vn-Th} = 4.1 \times 10^{-13}\,\mathrm{V^2/Hz} = A_{pre}^2 \cdot 1.64 \times 10^{-16}\,\mathrm{V^2/Hz}$, that is not far from half the preamplifiers' RTI thermal noise. With the final design, simulated RTI thermal noise turns out to be

$$S_{Vn-Th} = 3.53 \times 10^{-13}\,\mathrm{V^2/Hz} = A_{pre}^2 \cdot 1.41 \times 10^{-16}\,\mathrm{V^2/Hz} \tag{4.50}$$

About flicker noise as well a similar relationship can be found, that is not reported here. In OTA1 design, flicker noise was not a concern: in fact, since thermal noise is very high (we used a preamplifier just for this purpose), flicker corner frequency can be kept low without extremely large areas. In fact, even with commonly large areas (see dimensions in table 4.2), simulated flicker noise PSD at f_{ch} resulted in

$$S_{Vn-F}(f_{ch}) = 0.36 \times 10^{-13}\,\mathrm{V^2/Hz} = A_{pre}^2 \cdot 0.144 \times 10^{-16}\,\mathrm{V^2/Hz} \tag{4.51}$$

That is a factor 10 between the two noises' PSD and, consequently, a corner frequency $f_k \approx f_{ch}/10$.

In conclusion, putting together preamplifier's and OTA1's contributes at f_{ch} and neglecting the noise from the following stages, with chopper working we can expect a total wideband RTI noise of

$$S_{Vn-tot} = 3.4 \times 10^{-16}\,\mathrm{V^2/Hz} \qquad \rightarrow \qquad v_n = 18.4\,\mathrm{nV}/\sqrt{\mathrm{Hz}} \tag{4.52}$$

that would be coherent with the target of maximum $20\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ total RTI noise.

One consideration must be made about common gates. Since these are the only source of unchopped flicker noise in INT1 (we said that flicker noise from others integrators is negligible), they will most probably be responsible

for the effective corner frequency of the chopped system. The final Periodic Steady State (PSS) analysis, however, showed good results (see paragraph 5.4) even with relatively small areas.

4.4.5 Final dimensions

The main electrical parameters of the relevant devices have been already discussed and evaluated in the previous paragraphs. In table 4.2, all the devices' dimensions (bias stage excluded) are reported.

Device	W/L [µm/µm]
M_{cp}	0.6/32.5
M_{mn}	0.6/38
M_{gn}	6/92
M_{gp}	9/33
M_{mp}	0.6/18

Table 4.2: OTA1 devices dimensions.

4.5 INT1 CMFB

4.5.1 Operating principle

Control of output common mode in fully differential systems is a compulsory feature, without which CM output offset would devastating. In fact, consider a simplified folded cascode output stage, like the one we have in the preamplifier and in OTA1. Currents I_0 sourced from pMOS mirrors are designed to be equal to the ones sinked by nMOS. If a common mode current error I_ϵ affects these currents, like shown in figure 4.9, this current error will flow into the output CM resistance R_{oCM}, that in a cascode configuration can be as high as a few GΩ. Thus, a current I_ϵ of a few nA, common case in such topologies, would cause a CM output offset of some V, making the device impossible to use.

A traditional approach is to use a control loop, called Common Mode Feedback (CMFB), whose operation principle is shown in figure 4.10. Its role is to read output CM, compare it to a V_{refCM} provided externally as reference and then produce a voltage V_{CMFB} that corrects the bias currents I_0; eventually, if the control loop has an high negative gain, virtual short circuit will null the difference $(V_{outCM} - V_{refCM})$.

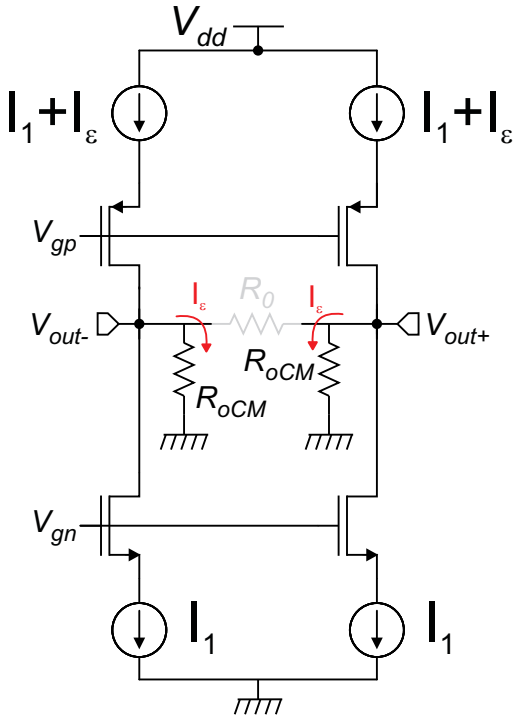


Figure 4.9: Output CM offset in a cascode structure.

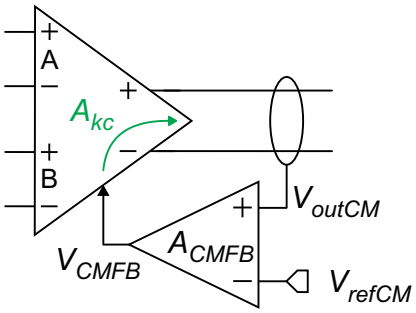


Figure 4.10: Classical CMFB block diagram.

Suppose that V_{CMFB} is connected to the gate of n mirrors M_{mn}, and that a voltage $V_{CMFB} = V_{k0}$ produces a nominal current I_0. Output CM then will be $V_{outCM} = V_0 + V_\epsilon$, where V_0 is the nominal output voltage, that is $V_{DD}/2$ for a symmetric configuration, and $V_\epsilon = R_{out} I\epsilon$. Suppose now that CMFB circuitry produces

$$V_{CMFB} = V_{k0} + V_c \quad \text{where} \quad V_c = A_{CMFB}(V_{outCM} - V_{refCM}) \tag{4.53}$$

Then, output CM will be

$$V_{outCM} = V_0 + V_\epsilon + A_{kc} V_c \tag{4.54}$$

where $A_{kc} = -g_{mm} R_{out}$ is the CMFB-to-outCM gain of our amplifier and R_{out} is the output common mode resistance of a cascode structure, that can be as high as a few GΩ.

With closed loop, finally, we will have

$$V_{outCM} = \frac{V_0 + V_\epsilon}{1 - A_{kc} A_{CMFB}} + \frac{-A_{kc} A_{CMFB}}{1 - A_{kc} A_{CMFB}} V_{refCM} \approx V_{refCM} \tag{4.55}$$

if $|A_{kc} A_{CMFB}| \gg 1$, and of course if feedback is negative, that is if $A_{kc} A_{CMFB} < 0$.

In reference design, preamplifier and OTA1 had one dedicated CMFB loop each. In this design, we try to make the CM stabilization more efficient by using a single loop, that exploits the high common mode gain A_{cc} of the pseudo-differential pair. If we look back at the schematic (figure 4.7, it is easy to understand that it is

$$A_{cc} = -G_{m1} R_{out} \tag{4.56}$$

Our goal, then, is to create a loop with the following operating principle: CMFB compares the OTA1 CM output V_{1CM} with V_{refCM}, producing a voltage V_{CMFB} that changes the preamplifier CM output, that is OTA1 CM input V_{0CM}: this voltage then influences back V_{1CM}, that thanks to virtual short circuit should be locked on V_{refCM}. Figure 4.11 shows this operating principle.

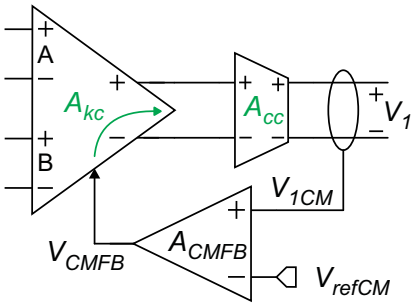


Figure 4.11: Proposed implementation of CMFB.

This time, we have for CMFB

$$V_{CMFB} = V_{k0} + V_c \qquad \text{where} \qquad V_c = A_{CMFB}(V_{1CM} - V_{refCM}) \tag{4.57}$$

Then, the preamplifier responds as it did in the previously example. Seen from the OTA1 point of view, this output voltage can be considered as sum of three parts: a V_{k1}, that is the voltage that would produce a nominal $V_0 = V_{DD}/2$ OTA1 CM output (that is, with perfectly matched currents: in our case this would be $V_{k1} = 0.8\,\mathrm{V}$), an error voltage V'_ϵ and of course the same increment V_c. So we have

$$V_{outCMpre} = V_0 + V_\epsilon + A_{kc}V_c = V_{k1} + V'_\epsilon + A_{kc}V_c \tag{4.58}$$

in which $V_0 + V_\epsilon = V_{k1} + V'_\epsilon$. Now, at the output of OTA1, we have that V_{k1} produces the nominal value of V_0, while the rest is amplified by A_{cc}. Then we have

$$V_{1CM} = V_0 + V_{\epsilon 2} + A_{cc}(V'_\epsilon + A_{kc}V_c) \tag{4.59}$$

where $V_{\epsilon 2}$ is the offset given by OTA1 output current mismatch. Finally, when the loop is closed, we have:

$$V_{1CM} = \frac{V_0 + V_{\epsilon 2} + A_{cc}V'_\epsilon}{1 - A_{cc}A_{kc}A_{CMFB}} + \frac{-A_{cc}A_{kc}A_{CMFB}}{1 - A_{cc}A_{kc}A_{CMFB}}V_{refCM} \approx V_{refCM} \tag{4.60}$$

if, again, $|A_{cc}A_{kc}A_{CMFB}| \gg 1$, and of course if feedback is negative, that is if $A_{cc}A_{kc}A_{CMFB} < 0$.

If the control is working correctly, then we should check what voltage is being set by the control loop at the preamplifier's output. From eq. (4.58) and (4.57) we get

$$V_{outCMpre} = V_{k1} + V'_\epsilon + +A_{kc}A_{CMFB}(V_{1CM} - V_{refCM}) \tag{4.61}$$

Finally, using eq. (4.60), after some algebraic steps we obtain

$$V_{0CM} = V_{k1} + \frac{V'_\epsilon + (V_0 + V_{\epsilon 2})A_{kc}A_{CMFB} - V_{refCM}}{1 - A_{cc}A_{kc}A_{CMFB}} \approx V_{k1} \tag{4.62}$$

that is exactly what we expected, that is our $0.8\,\mathrm{V}$.

4.5.2 Topology

A few different solutions for a CMFB have been proposed in literature, both static and dynamic. The latter are mainly Switched Capacitor architectures, that are efficient and compact but affected by problems of charge injection. Among the static solutions, the most diffused is a double differential pair, whose output current is $I_{CMFB} = kI_0 + g_m^*(V_{ox} - V_{refCM}) + g_m^*(V_{oy} - V_{refCM}) = kI_0 + g_m^*(V_{oCM} - V_{refCM})$: this architecture, however, has limited input differential and CM range, not suitable for the quite wide voltage swing that we have at OTA1 output. Even resistor degenerated double differential pair were tested, but they proved unsuitable after simulations.

The solution proposed is based on a resistive voltage divider placed across a buffered output voltage, so that the network will not cause loading effect on OTA1 output, in order to extract the common mode V_{1CM} from the output voltages; then, a simple differential pair compares it with V_{refCM} and produces V_{CMFB}. Since at steady-state this difference will be small, this pair can have small input differential range and will work with an almost constant input CM. A block diagram of the solution is shown in figure 4.12.

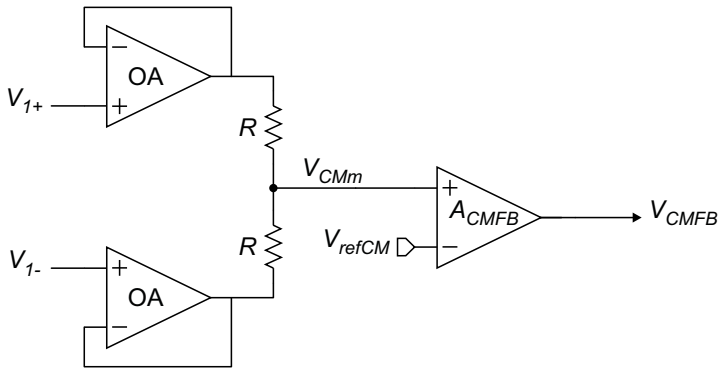


Figure 4.12: Block diagram of CMFB.

Elementary buffers Buffers are basically OpAmps with negative feedback: for this purpose, very simple OpAmps could be sufficient, in order to avoid excessive current consumption and circuit complexity. So, for closed loop buffers we used the following topology, that we will indicate as Simple Buffer (SB, see figure 4.13).

This is basically a simple differential amplifier with an open loop gain $A_0 \approx g_{mc}r_d \gg 1$ (a few hundreds), enough to set a good replica of input voltage on the output. With closed loop, its output resistance is $R_{out} =$

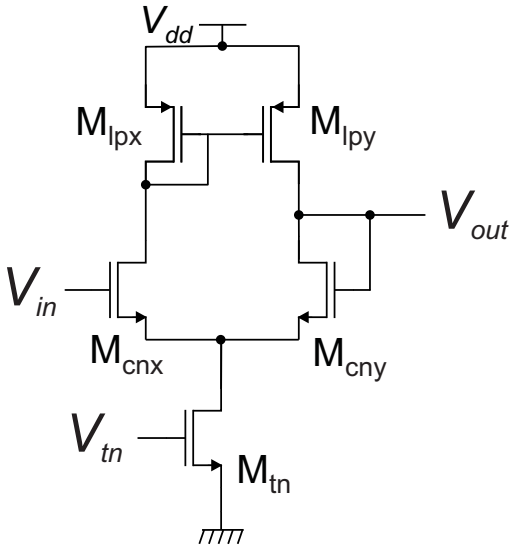


Figure 4.13: Schematic of the Simple Buffer.

$1/g_{mc}$, since the output MOSFET is mounted as a diode. If we load the two buffers' Thevenin equivalent with the resistive divider, we obtain such a circuit (figure 4.14):

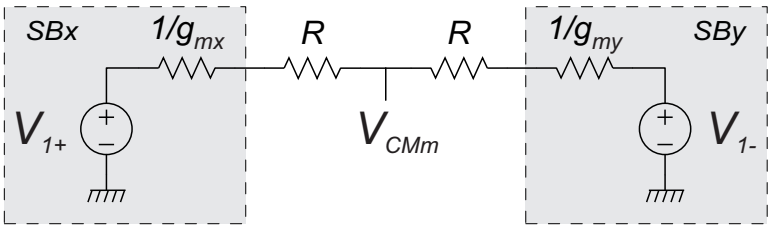


Figure 4.14: Thevenin equivalent of the circuit.

From this schematic, we can find that the output voltage has the following expression:

$$v_{out} = V_{1CM} + \frac{\dfrac{1}{g_{mx}} - \dfrac{1}{g_{my}}}{2\left[2R + \dfrac{1}{g_{mx}} + \dfrac{1}{g_{my}}\right]} V_{1D} \tag{4.63}$$

If V_{1D} is large, output resistances $1/g_m$ get different, since the two buffers work with different input CM. However, if g_m is not too small and R is high, the second term in the sum can be neglected. On the other hand, too high a g_m is unnecessary and it will cause large current consumptions. Obviously,

to optimize traconductance efficiency g_m/I_D, MOSFETs in the pair will in sub-threshold region; then, after simulations, an $I_0 = 2\,\mu A$ was considered a good trade-off between high g_m (the resulting value is $43\,\mu A/V$) and current consumption.

For the value of the resistances R, it must be noticed that the maximum output current of buffers is I_0, while the maximum voltage drop across the resistors is $V_{1Dmax} = 0.8\,V$. Thus, we must assure that

$$R \geq \frac{1}{2}\frac{V_{1Dmax}}{I_0} = 200\,k\Omega \tag{4.64}$$

So, a value of $R = 200\,k\Omega$ was set (higher values would increase area).

Differential amplifier Then, a second stage must sense the difference $(V_{1CM} - V_{refCM})$ and produce a proper current to be mirrored into the preamplifier's output stage. The proposed schematic is shown in figure 4.15.

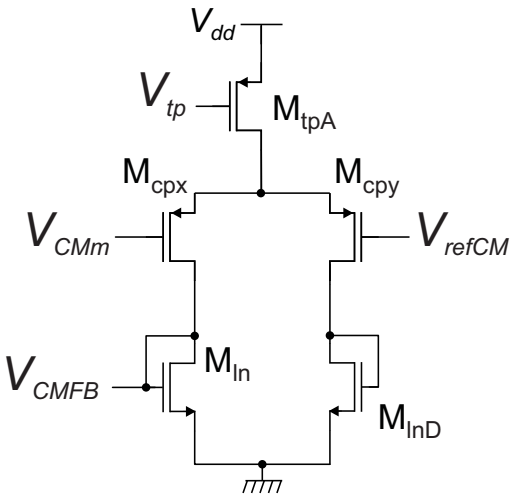


Figure 4.15: Schematic of CMFB.

If V_{CMFB}, as said, is connected to the gates of M_{mn} of the preamplifier (figure 4.6c), quiescent current I_0 is mirrored into the preamplifier's output stage. There, we had a quiescent current of $I_{mn} = 40\,\mu A$; here, like in the preamplifier's bias stage, we can use a mirroring factor of 10 to avoid waste of current, thus setting a current $I_1 = 4\,\mu A$. This current, present when differential input is zero (that is when $V_{1CM} = V_{refCM}$) will produce a $V_{CMFB} = V_{k0}$, as said in the previous paragraph. Now, when $V_{1CM} \neq V_{refCM}$, the current flowing in M_m will be

$$I_{CMFB} = I_0 - \frac{g_{mc}}{2}(V_{1CM} - V_{refCM}) \tag{4.65}$$

and the consequent CMFB voltage will be

$$V_{CMFB} = V_{k0} - \frac{g_{mc}}{2g_{mm}}(V_{1CM} - V_{refCM}) \tag{4.66}$$

where we have set $g_{mc} = 70\,\mu A/V$ (M_c again works in sub-threshold region to optimize tranconductance efficiency) and $g_{mm} = 19\,\mu A/V$. The latter was not a Degree of Freedom, since for a good mirroring it must have the same dimensions of the preamplifier's M_{mn}, apart from the multiplier factor `m` that sets the mirroring factor.

We have then found an expression for CMFB gain, that is (exact value obtained from simulations)

$$A_{CMFB} = -\frac{g_{mc}}{2g_{mm}} = -1.78 \tag{4.67}$$

4.5.3 Final dimensions

Range considerations have already been discussed, while we omit a noise analysis since all noise coming from CMFB circuitry will only affect the output common mode.

As usual, table 4.3 shows all the devices' dimensions.

Device	W/L [µm/µm]	Device	W/L [µm/µm]
M_{cn}	10/1	M_{cp}	20/0.5
M_{mp}	1.5/2	M_{ln}	3.4/16
M_{tn}	2.5/1	M_{tp}	10/2

Table 4.3: CMFB devices dimensions.

4.5.4 Loop stabilization

As known, the correct operation of a negative feedback loop is guaranteed if the feedback remains negative for all frequencies, that is the loop is stable. At high frequencies, in fact, the loop transfer function $\beta A(j\omega)$ can have poles and zeroes that invert the phase of the signal and turn the feedback from negative to positive, causing instability. In particular, an oscillation will rise at frequency $\omega_x = 2\pi f_x$ if the well known Barkhausen criterion is satisfied, that is if

$$\begin{cases} |\beta A(j\omega_x)| \geq 1 \\ \angle \beta A(j\omega_x) = 0 \end{cases} \tag{4.68}$$

Now, consider a system like our CMFB. We have found that, in this case, loop transfer function has a DC gain of

$$\beta A = A_{CMFB} A_{kc} A_{cc} \qquad \text{where} \begin{cases} A_{CMFB} = -\left.\dfrac{g_{mc}}{2g_{mm}}\right|_{CMFB} \\ A_{kc} = -\left. g_{mm} R_{out}\right|_{pre} \\ A_{cc} = -\left. g_{mc} R_{out}\right|_{OTA1} \end{cases} \tag{4.69}$$

Now, the transfer function found above has a DC phase of 180° and a high DC gain. At high frequencies, however, parasitic capacitances will introduce poles and zeroes, that could cause a phase shift of −180° or more that would give rise to oscillations at certain frequencies, since the gain may still be high.

Usually, compensation techniques are used: among these, the most diffused is the dominant pole compensation, that is used in this work as well. Basically, we introduce a low frequency pole f_p with a large capacitor (at least much larger than parasitic capacitances): then, the phase for $f > f_p$ will settle to 90°, while gain decreases at a slope of −20 dB/dec. In this way, if the system is correctly designed, when the phase shift caused high frequency poles has effect, the gain will be already less than 0 dB and the system will be stable.

In our design, the dominant pole is evidently given by the 25 pF common mode capacitors at OTA1 output, that introduces a pole at frequency at

$$f_p = 75\,\text{mHz} \tag{4.70}$$

as found from simulations. However, this pole was not sufficient to stabilize the system: in fact, simulations showed oscillations in the output common mode.

This was because on one hand the dominant pole was correctly set at low frequency, but on the other hand the DC gain was extremely high, and the dominant pole alone was not sufficient to decrease it enough to make it less than 0 dB when parasitics come into play (that is, at a few tens of kHz, as found from simulations). We could have increased a lot the value of C_1 to shift f_p to a much lower frequency, but at the cost of extremely large areas; instead, a different approach was used. The basic idea is that to decrease A_{kc}: since the preamplifier's g_{mc} is fixed by noise considerations, this leaves

us to decrease the output common mode resistance of the summing stage. A local feedback loop employing transistors M_{mpH} was implemented in order to achieve this target. The solution is presented in figure 4.16: the red subcircuit represents the local feedback of M_{mpH}.

Since we are analyzing a common mode in a symmetric circuit, it has been cut along its symmetry axis and only one half has been represented. We can distinguish the three blocks of CMFB, preamplifier and OTA1 (buffers are omitted), while in blue are highlighted the parasitic capacitances that cause poles and positive zeroes.

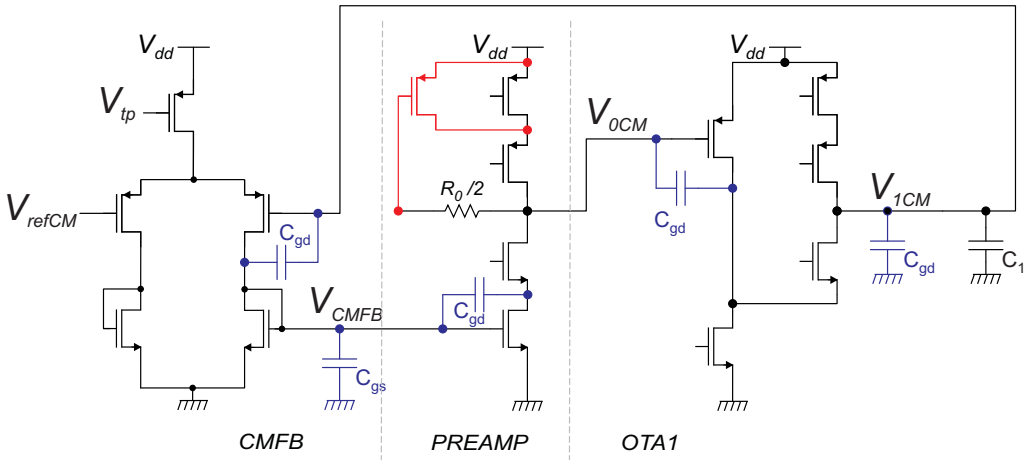


Figure 4.16: Schematic of the CMFB loop, with the main devices involved and the most relevant parasitic capacitances.

Now, it is easy to understand that the output common mode resistance of the preamplifier is around

$$R'_{out} = 1/g_{mpH} \tag{4.71}$$

Indeed, M_{mpH} is in a diode configuration, if we notice that no current flows through $R_0/2$ and that the common gate M_{gp} is uneffective. So, a value of g_{mpH} was chosen in order to decrease A_{kc} enough to stabilize the loop: simulations showed that a value of $A_{kc} = 18\,\text{dB}$ (much lower than a classical cascode structure) was sufficient, and it could be achieved by setting $g_{mpH} = 23.7\,\mu\text{A/V}$.

In other terms, the local feedback loop lowers the output resistance, since a portion of the bias current sourced from above depends on the output CM voltage itself. Suppose that the $40\,\mu\text{A}$ nominal bias current is partially given by the mirror M_{mp} ($34\,\mu\text{A}$ in this case), while transistors M_{mpH} are designed to give the remaining part when the output CM voltage is at the

desired nominal value (800 mV in this case). If, for any reason, the total bias current tends to increase, it flows into the parasitic resistances (the classical R_{out} of a cascode structure), increasing the output CM: this way, feedback transistors M_{mpH} will have a lower overdrive voltage and will decrease their current, thus mitigating the effect. As a consequence, the equivalent output resistance is much lower than in a classical cascode structure.

Besides the effect on the gain, the first non dominant pole, associated to the preamplifier common mode output resitance, is shifted to much higher frequencies, thus improving the phase margin. Note that no effect is caused to the differential mode resistance.

After this considerations, the feedback loop turned out to have a DC gain 113 dB, so still high enough to guarantee a good accuracy on the output CM voltage, and a phase margin of 60° to guarantee stability.

4.6 INT2

4.6.1 Topology

For INT2, implemented as well as a Gm-C integrator, we need an OTA with such a transcharacteristic:

$$I_{out2D} = G_{m2}(v_{1D} + v_{fb2D}) \tag{4.72}$$

where $G_{m2} = \omega_{02}C_2 = 2\pi\ 1\,\text{kHz}{\cdot}100\,\text{pF} = 628\,\text{nA/V}$, v_{1D} is INT1 differential output and $v_{fb2D} = -\beta_2 v_{out}$ ($\beta_2 = 0.5$). At first sight, a topology similar to the one used for the preamplifier could be appropriate, but two main problems make it not suitable.

First of all, for low frequencies, we are expecting $v_{1Dmax} \approx v_{fb2Dmax} \approx 0.5v_{out-max} = 0.8\,\text{V}$, that are very difficult to achieve even with resistor degenerated differential pairs.

Secondly, and most of all, we need a high common mode gain A_{cc}, as said in paragraph 4.5.1, in order to close a CM feedback loop and have an effective CM gain of $1/\beta_{2CM} = 1$. This way, the common mode set by CMFB will be transmitted to output and no more CM control loops will be needed.

So, following the example of the reference design, a similar topology was implemented, that is shown in figure 4.17.

First of all, we can notice that the equivalent CM capacitors of 100 pF were replaced by a differential capacitor $C_2' = 40\,\text{pF}$ and two smaller CM capacitors $C_2 = 20\,\text{pF}$, that however could not be eliminated for stability

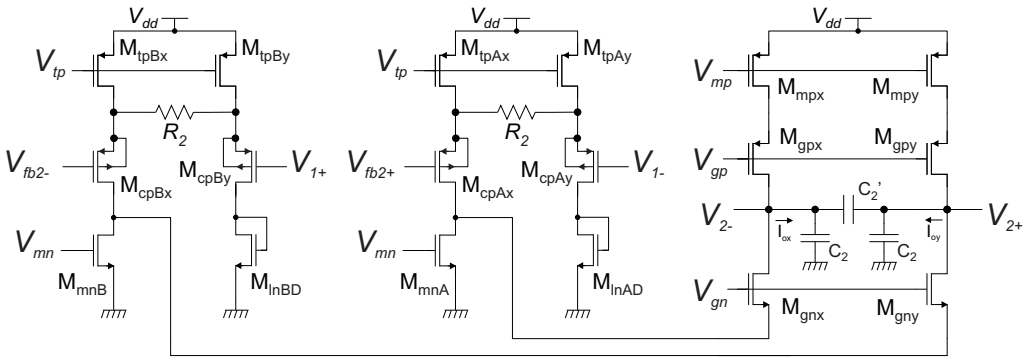


Figure 4.17: Schematic of INT2.

matters (they are responsible for the dominant pole compensation of INT2 CM feedback loop). It is easy to verify that $C_{1eq} = C_1 + 2C_1' = 100\,\text{pF}$.

Again from the schematic, and reminding that $v_{fb2D} = -\beta_2 v_{outD}$ and $v_{fb2CM} = \beta_{2CM} v_{outCM} = v_{outCM}$, it is straightforward to obtain the differential and CM input-output relationships, as follows:

$$I_{out2D} = I_{oy} - I_{ox} = G_{m2}(v_{1D} + v_{fb2D}) \tag{4.73}$$

$$I_{out2CM} = I_{ox} + I_{oy} = G_{m2}(v_{1CM} - v_{fb2CM}) \tag{4.74}$$

In these relationships, G_{m2} is the transconductance of each resistor-degenerated pair, that has the following well known expression:

$$G_{m2} = \frac{g_{mc}}{2 + g_{mc} R_{d2}} = 628\,\text{nA/V} \tag{4.75}$$

in our case, where R_{d2} is the degeneration resistor.

It is simply verified that both differential and CM feedbacks are negative. Furthermore, since $v_{2CM} = R_{out2} I_{out2CM}$, we have a CM gain of

$$A_{cc} = G_{m2} R_{out2} \gg 1 \tag{4.76}$$

since R_{out2} again is a cascode output resistance, that can be as high as a few GΩ.

Since noise costraints, as said, are not critical for this stage, topology and the principal design choices derived from input and output ranges considerations.

4.6.2 Input ranges

Input CM range As said, for low frequencies we have a virtual short circuit between the input differential voltages v_{1D} and v_{fb2D}, each of which can have an amplitude up to 800 mV. With our design, input differential range is not a problem, since each pair works with almost equal voltages (granted that both signals share the same common mode, as expected). With our design, on the other hand, input CM range is critical, since it is expected to cover a range from 0.5 to 1.3 V. For p pairs, chosen in our case for range considerations, the critical region is the one nearest to the positive supply: we have

$$V_{inCMmax} = V_{DD} - V_{DSat-t} - V_{GSc} = V_{DD} - V_{ODt} - V_{th} - V_{ODc} \quad (4.77)$$

To make it equal to 1.3 V, we had to set the following values:

- $V_{ODt} = 100\,\text{mV}$, that is the minimum for strong inversion region;
- $V_{th} = 500\,\text{mV}$ with no body effect, to keep it at minimum: for this reason, pairs based on pMOS were used, with insulated n-well connected to source terminals;
- $V_{ODc} = -100\,\text{mV}$, so in deep subthreshold region.

Input differential range As said, input differential range is not critical for low frequencies signals, thanks to virtual short circuit. But again, if signals have a frequency near the cutoff, a phase shift can be present between v_{1D} and $v_{fb2D} = -\beta v_{outD}$. In this case, for example, a phase shift of 90° is seen at the cutoff frequency: here, then, when input voltage v_{1D} is at a maximum, feedback voltage will be null. So, each pair will undergo an input differential voltage of 400 mV, if a 1 kHz, 8 mV input signal were presented.

Now, since in classical differential pairs input differential range is around $2\sqrt{2}V_{TEc}$, it would be impossible to reach our target with MOSFETs operating in sub-threshold region. Thus, **resistor degenerated pairs** were used. A current of 220 nA was used - easy to get, simply mirroring it from OTA1 bias stage - to guarantee the needed range. With this value of drain current and the overdrive of $V_{ODc} = -100\,\text{mV}$, pMOSFETs of the pairs showed a transconductance of $g_{mc} = 4.7\,\mu\text{A/V}$: thus, from eq. (4.75), we find the needed value for the degeneration resistance

$$R_{d2} = \frac{1}{G_{m2}} - \frac{2}{g_{mc}} \approx 1.17\,\text{M}\Omega \quad (4.78)$$

Output range Output range costraints are the same as those seen for OTA1: since the output stage has the same cascode structure, overdrive voltages were set at similar values.

4.6.3 Final dimensions

The main electrical parameters of the relevant devices have been already discussed and evaluated in the previous paragraphs. In table 4.4, all the devices' dimensions (bias stage excluded) are reported.

Device	W/L [µm/µm]	Device	W/L [µm/µm]
M_{cp}	60/3	M_{mp}	2/25
M_{tp}	15/27	M_{gp}	3/5.5
M_{mn}	$2 \cdot 1.5/75$	M_{gn}	2/15

Table 4.4: INT2 devices dimensions.

4.7 INT3

4.7.1 Topology

Topology of INT3 is directly derived from INT2, since costraints about input ranges, differential tranconductance and CM gain are substantially the same. The main difference is that this time we need a low output impedance, since INT3 must be able to drive properly the resistive feedback network. So, a Gm-OpAmp topology was chosen, that is shown in figure 4.18.

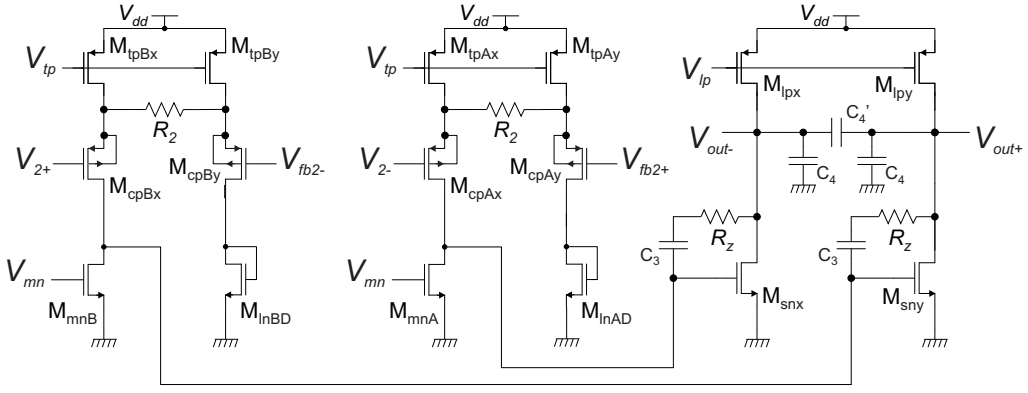


Figure 4.18: Schematic of INT3.

As we can see, the fully differential OpAmp is simply made with two class A common source nMOSFETs: indeed this output stage, while extremely simple, is sufficient to achieve the main characteristics of a fully differential OpAmp, that are high input impedance and high negative differential gain. With this topology, Miller capacitors are used to integrate the first stage's output current. Since their function and role in the INT3 transfer function is equivalent to the common mode capacitors of INT1-2, a relatively low value of capacitance of $C_3 = 50\,\text{pF}$ was chosen to avoid excessive areas: this time, indeed, it is not possible to introduce the equivalent of the differential mode capacitors used in INT1,2, which in terms of area would have been four times more effective than common mode ones. The consequent value of transconductance needed is

$$G_{m3} = \omega_{03} C_3 = 1.257\,\mu\text{A/V} \tag{4.79}$$

In this case as well, it is easy to obtain the main electrical parameters of this block. As transconductance, like for INT2 it is given by the transconductance of the input resistor degenerated pair, whose expression was given in eq. (4.75). As common mode gain, however, this time the expression is different: from the schematic, it is easy to understand that

$$A_{cc} = \frac{G_{m3}}{2} R'_{out3} \cdot g_{ms} R_{out3} \gg 1 \tag{4.80}$$

since both R'_{out3} (output resistance of the pairs) and R_{out3} are proportional to the output resistance r_d of a MOSFET in saturation region. This high CM gain, as for INT2, is necessary for the correct operation of INT3 CM feedback loop.

As said for INT2, the design of INT3 as well was mainly dictated by ranges considerations.

4.7.2 Input ranges

Input CM range The same considerations made for INT2 are valid: so, in this case as well, pair transistors work in deep sub-threshold regions, and the resulting overdrive voltage for pairs, tails and mirrors are the same as the ones seen in INT2.

Input differential range For this stage as well, input differential range only matters when large, high frequency signal are presented. Actually, since only one integrator separates v_2 from output, phase shift will be even less, in particular it will be only 45° at f_c. Anyway, we decided to mantain

the same input differential range for INT3 pairs, for a faster design: since we have a G_{m3} double compared to G_{m2}, a double drain current was chosen ($I_{Dc} = 440\,\text{nA}$).

The previous considerations allowed us the complete design of the pairs. In particular:

- a double drain current will cause a sub-threshold operating MOSFET to have a double transconductance: thus, we have for pairs' transistors a $g_{mc} \approx 9.4\,\mu\text{A/V}$;
- In order to obtain the desired G_{m3}, a new value for the degeneration resistance was found again using eq. (4.78) (the value obtained from this equation had to be slightly corrected after simulations). We have then

$$R_{deg3} = 576\,\text{k}\Omega \tag{4.81}$$

4.7.3 Output ranges

For a class A common source, output range limits are given by the following relationships.

$$V_{out-max} = V_{DD} - V_{DSat-l} = V_{DD} - V_{ODl} \tag{4.82}$$
$$V_{out-min} = V_{DSat-s} = V_{ODs} \tag{4.83}$$
$$\tag{4.84}$$

Since we want a R2R output, overdrive voltages should be at minimum. In particular:

- V_{ODl} was set to 100 mV to guarantee strong inversion;
- V_{ODs} was slightly less, around 20 mV, since we needed a higher g_{ms} (this MOSFETs are the equivalent of an OpAmp, so they should have a gain as high as possible)

Furthermore, we need to consider the maximum output current, that must be enough to drive the resistive nework for all output voltages. Now, the values of the feedback resistances were chosen as a trade-off between thermal noise and current consumption: thus, for the main feedback network, two pairs of 1 kΩ and 199 kΩ were chosen, while for the secondary feedback network four 100 kΩ resistors were used. Hence, the overall resistive load turns out to be

$$R_{L-tot} = [2(1+199)] \parallel [2(100+100)]\,\mathrm{k\Omega} = 200\,\mathrm{k\Omega} \tag{4.85}$$

So, for a maximum output voltage of $\pm 1.6\,\mathrm{V}$, we obtain a value for the maximum output current needed:

$$I_{out-max} = \frac{V_{out-max}}{R_{L-tot}} = 8\,\mu\mathrm{A} \tag{4.86}$$

Thus, a current $I_{Dl} = 11\,\mu\mathrm{A}$ was set as quiescent current for the loads (and common sources as well), obtained thourgh mirrors from INT2 bias stage with a 5 multiplying factor.

4.7.4 Noise

While noise did not influence the design of this stage, it had to be monitored more than the one from INT2: in fact, as said, high frequency noise from INT3 will be the dominant contribution in total output noise at frequencies higher than f_c. However, simulations proved that this contribution was not critical, compared to the total output noise power, so the original design described above was confirmed.

4.7.5 Common mode loop stabilization

A brief digression about CM feedback is needed at this point. For INT3 as well a dominant pole compensation was implemented, through Miller capacitances C_3 that play the same role of the CM capacitances C_2 for INT2. Unlike them, however, C_3 do not introduce only the dominant pole, but also a zero with positive real part. That is, these capacitors cause an inversion of the signal transfer function at high frequencies: in fact, while the low frequency signals are inverted by the common sources when passing from gate to drain, at high frequencies Miller capacitors are short circuits that bypass the common sources, that usually invert the signal. This way, feedback can turn from negative to positive, causing potential instability: it can be proved that the positive zero introduced by C_3 is at a frequency

$$f_z = \frac{1}{2\pi C_3 \left(\dfrac{1}{g_{ms}} - R_z\right)} \tag{4.87}$$

Now, without R_z and with the simulated value of $g_{ms} = 206\,\mu\mathrm{A/V}$, we would get a value of the zero frequency $f_z \approx 30\,\mathrm{kHz}$, that is relatively low: at this frequency, A_{cc} may not have been decreased enough by the dominant

pole, and there may be a frequency at which the Barkhausen condition is satisfied, causing oscillations in the CM output. Thus, a common solution, employed in this work as well, is to use a resistor R_z to change the zero frequency: in this case, a value of

$$R_z \approx \frac{1}{g_{ms}} \tag{4.88}$$

is used, thus shifting the zero to very high frequency (ideally to infinity). For this reason, in this case R_z is called *zero nulling resistor*. From the previous relationship, and with some corrections believed appropriate after simulations, a value of $R_z = 6\,\mathrm{k}\Omega$ was set.

4.7.6 Final considerations and dimensions

As we can see from the schematic, additional capacitors were used at the output, although they do not change the frequency response of the integrator (since they are driven by a low impedance ouput stage). Common mode capacitors $C_4 = 1\,\mathrm{pF}$ were used to reduce the amplitude of output common mode voltage spikes caused by charge injection effects; differential capacitor $C_4' = 10\,\mathrm{pF}$, instead, was used as an additional filtering element for high frequency differential voltage noise.

After all these consideration, the final dimensions of INT3 devices (bias stage excluded) are reported in table 4.5.

Device	W/L [µm/µm]	Device	W/L [µm/µm]
M_{cp}	60/1.5	M_{sn}	20/1
M_{tp}	$2 \cdot 10/18$	M_{lp}	$10 \cdot 2.5/1$
M_{mn}	$2 \cdot 1/50$		

Table 4.5: INT3 devices dimensions.

4.8 CMDA

4.8.1 Topology

The Common Mode Difference Amplifier is a device with two input ports *in* and *fb*, whose open loop operation is described by the following relationship:

$$V_H = A_{CDMA}(v_{inCM} - v_{fbCM}) \tag{4.89}$$

where $A_{CMDA} \gg 1$.

Such operation can be achieved by using two differential pairs A and B (and a consequent summing stage), where non inverting inputs v_{A+} and v_{B+} will be connected to v_{in+} and v_{in-}, while inverting inputs of each pair will be connected to the feedback positive and negative signals.

Now, our aim is to set $v_{fbCM} \approx v_{inCM}$: since the latter is expected to vary within the whole range V_{DD} to ground, and differential signals are negligible (they can reach 8 mV maximum), it is evident that we need **Rail-To-Rail CM input range** for both differential pairs.

The best way to design this stage, thus, is to employ two differential pairs (one p and one n type) for each input port, exactly like we did in the preamplifier: a folded cascode summing stage is then the most straightforward way to deal with the four pairs' output currents. Compared to the preamplifier, however, CMDA presents the following relevant differences.

- Noise will affect only the output common mode, so it is not critical. This allowed us to use lower quiescient current for the pairs' transistors: a nominal current of 220 nA was used.

- A high gain was needed, and achievable thanks to the high output resistance of the cascode summing stage; however, in order to drive the resistive load we need a low output impedance, achievable only with a two stage architecture.

- We need a single-ended output: thus, the cascode stage can operate the differential-to-single-ended conversion.

- Since the current consumption was negligible compared to the preamplifier's, there was no need of self-biasing pairs. In addition, since a current mirror is used to operate the differential-to-single-ended conversion, the excess current caused by a pair turning off is mirrored and autocompensated, thus not causing any additional output CM offset. Therefore, to reduce schematic complexity, traditional pairs were used.

- It can be proved that, for common mode signals, the equivalent resistive load for CMDA is $R_L = 100\,\mathrm{k\Omega}$ (see figure 4.1 at page 70). Now, since $V_H \approx v_{inCM}$ can swing from gound to V_{DD}, we find that the maximum output current for CMDA will be

$$I_{out-max} = \left| \frac{V_{Hmax} - V_{outCM}}{R_L} \right| = \left| \frac{V_{Hmin} - V_{outCM}}{R_L} \right| = 9\,\mu\mathrm{A} \qquad (4.90)$$

To save current consumption, instead of a class A topology we can use a class AB output stage. As known, this topology needs a level shifter (to drive properly both nMOSFET and pMOSFET), that for a R2R output can be a so called Monticelli cell [21].

According to the previous considerations, here are the schematics of a n pair (figure 4.19a), a p pair (figure 4.19b) and the output stage (figure 4.19c).

First of all, as said for the preamplifier, these n and p pairs are relative to port A: obviously, an exact copy of these pairs is implemented for port B. Secondly, a closer look at the output stage allows distinguishing several relevant subcircuits, that are the cascode summing stage and the actual class AB output stage. In the cascode summing stage, in turn, we can distinguish the two Monticelli cells and a wide range precision cascode current mirror, used to operate the differential-to-single-ended conversion. Finally, it should be noticed that Miller capacitors C_c are used for dominant pole compensation, while R_z again are zero nulling resistors.

4.8.2 Noise

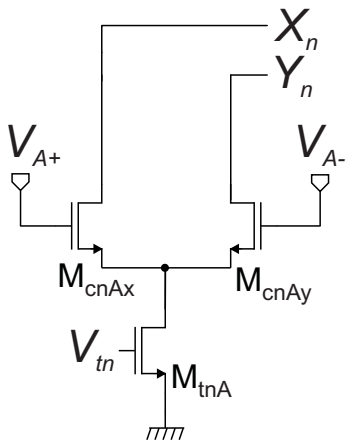
Since, as said, noise from this stage affects only the output common mode, device sizing was almost completely dictated by range considerations, as it will discussed in detail in the following paragraphs.

4.8.3 Input differential and CM ranges

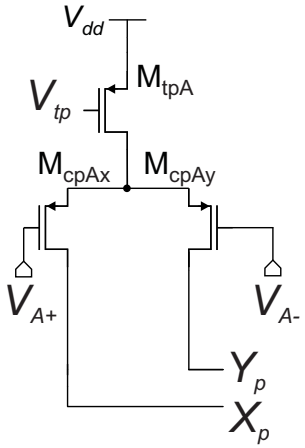
If the CMEQ technique works properly, we have $v_{inCM} \approx v_{fbCM}$; furthermore, at low frequency, virtual short circuit sets $v_{inD} \approx v_{fbD}$. As a consequence, if positive signals are input of one pair and negative signals are input of the other one, then input differential voltage for each pair will be negligible (even at high frequency it will be only a few mV), while input CM voltage can swing from ground to V_{DD}.

This costraints are basically the same that we assumed for the design of the preamplifier's input pairs. Since the topology of the input stage (differential pairs) as well is substantially the same as in the preamplifier, the same relationships led to the design of the pairs' transistors: this time, being noise contributions almost irrelevant, design aimed to maximaze ranges. In particular:

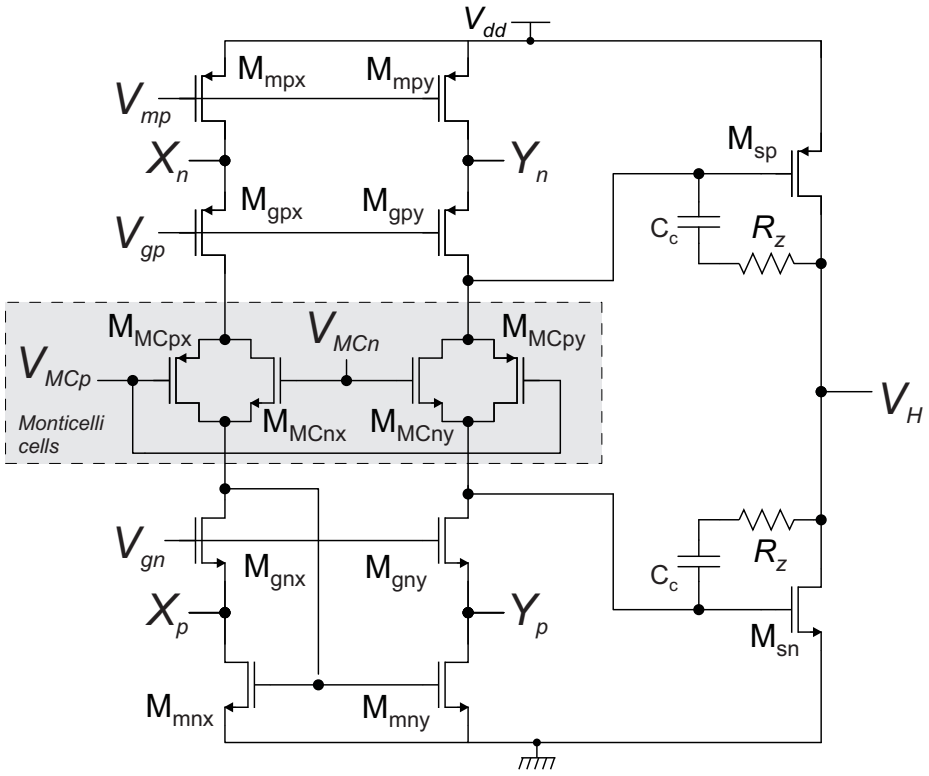
- $V_{ODt} \approx 100\,\text{mV}$, the minimum to guarantee strong inversion



(a) CMDA n pair.



(b) CMDA p pair.



(c) Output stage: folded cascode stage with differential-to-single-ended conversion through a wide range precision current mirror, Monticelli cells (bias omitted) and the class AB output stage.

Figure 4.19: Schematic of the CMDA. Only the differential pairs relative to port are shown, for simplicity.

- M_c work in sub-threshold region, to maximize g_m efficiency, since differential range is not an issue for these pairs.
- $V_{ODm} \approx 100\,\text{mV}$
- $V_{DSm} \approx 150\,\text{mV}$, to guarantee saturation of transistors M_m

This way, input CM range includes ground (p pairs working) and positive rail (n pairs working), and in a significant portion of the range both pairs are working. It must be noticed that each transistor M_m should mirror a current for two pair transistors I_{Dc}, plus a current for the common gate stage, that is $I_{Dg} = 2I_{Dc}$ for current ranges considerations (see preamplifier). Thus, this sets a value for $I_{Dm} = 4I_{Dc} \approx 1\,\mu\text{A}$ (since the effective simulated I_{Dc} ins nearly 250 nA, slightly more than the nominal value).

4.8.4 Output ranges

The output of CMDA must be R2R, to correctly track the input CM that can swing from ground to the positive supply rail. It is easy to understand that the output range, like in INT3 output stage, depends on common sources' overdrive voltages, that for this reason were set to the minimum that guarantees stong inversion, that is $V_{ODs} \approx 100\,\text{mV}$. Furthermore, a quiescent current $I_{Ds} \approx 1\,\mu\text{A}$ was set for this stage.

4.8.5 Cascode summing stage and Monticelli cells

This stage is responsible for the generation of the output stage's MOSFETs gate voltages: high gain is achieved through the high cascode output resistance, while conversion from differential to single-ended is made through a wide range precision cascode current mirror. Since voltage range is to be maximized, common gates' overdrive voltages were set to the minimum, that is $V_{ODg} \approx 100\,\text{mV}$.

In the middle of the cascode stage, two Monticelli cells operate as level shifter, in order to drive properly the output common source MOSFETs. The design of the Monticelli cells was made according to the following considerations.

- In their normal operation, that is for not extremely large voltages, both of them are in saturation region: so we chose to let an equal current flow in both n and p transistors: $I_{D-MCp} = I_{D-MCn} = I_{Dg}/2 \approx 250\,\text{nA}$.

- The following relationship is valid:

$$V_{MCn} = V_{OD-MCn} + V_{ODsn} + 2V_{th} \tag{4.91}$$

  Now, V_{OD-MCn} was set as the value that made the currents equal (found after simulations), so, being known threshold voltages, V_{MCn} set the common source overdrive voltages. This voltage is set by a dedicated bias network, not shown in figure. The resulting value after simulations was $V_{MCn} \approx 1\,\text{V}$.

- An equivalent relationship can be found for transistors M_{MCp}, that leads to a value of voltage $V_{MCp} \approx 430\,\text{mV}$.

4.8.6 Final dimensions

The previous considerations were sufficient to the complete design of CMDA: the main devices' dimensions are reported in table 4.6.

Device	W/L [µm/µm]	Device	W/L [µm/µm]
M_{cn}	4/10	M_{gp}	2/3.2
M_{tn}	$2 \cdot 1/7$	M_{mp}	$4 \cdot 2/3.2$
M_{cp}	8/6	M_{MCn}	1/7
M_{tp}	$2 \cdot 2/2.8$	M_{MCp}	2/3.2
M_{mn}	$4 \cdot 1/7$	M_{sn}	$4 \cdot 1/7.3$
M_{gn}	1/7	M_{sp}	$4 \cdot 2/3$

Table 4.6: CMDA devices dimensions.

4.9 Switches and clock generator

4.9.1 Deviator

The switching blocks are in charge to steer the signals flow according to a clock signal. Now, a deviator is actually made with two switches, each one consisting in a pass transistor or a pass gate. As known, for voltages near the positive rail, pMOSFETs are more suitable, while nMOSFETs are best used with voltages near ground; pass gates instead are used when source and drain voltages can be any voltage within the range ground-V_DD, so that the switch in ON state do not cause any voltage drop in the signal path.

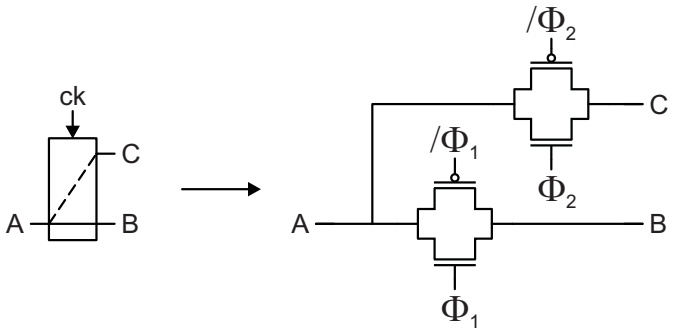


Figure 4.20: Electrical implementation of a switch.

Figure 4.20 shows the actual implementation of a simple deviator with pass gates, highlighting the control signals.

Suppose that node A should be connected to node B during phase 1 (Φ_1, solid line) and with node C during phase 2 (Φ_2, dashed line). So, the following considerations can be made:

- $/\Phi_1$ is the complementary voltage of Φ_1 (and the same for Φ_2), since nMOS and pMOS of the same pass gate must be ON or OFF at the same time and so must be driven by complementary gate voltages;
- Φ_1 and Φ_2 must never be high at the same time, otherwise nodes A, B and C would be shorted together. So, a **non overlapping clock** is strongly recommended.
- Similarly, $/\Phi_1$ and $/\Phi_2$ must never be low at the same time, for the same reason. Again, a non overlapping clock generator is suitable to produce these signals.

4.9.2 Non overlapping clock generator

Figure 4.21 shows the digital block representation of a non overlapping clock generator. Input signal *ck* is a reference clock signal with accurate 50% duty cycle, while output signals are indicated by the labels.

4.9.3 Final dimensions

Digital NAND and NOT gates present in non overlapping clock generator were both realized with their classical CMOS implementation, where minimal transistors were used (dimensions $W/L = 280\,\mu m/180\,\mu m$).

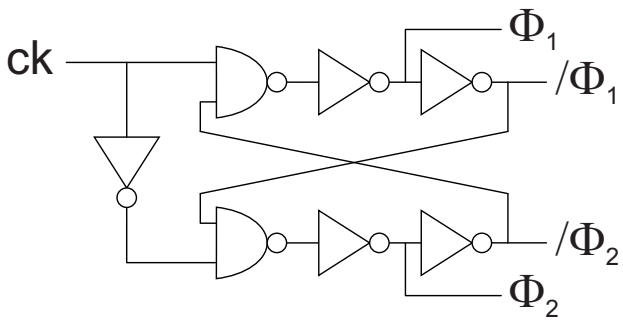


Figure 4.21: Gate diagram of the Non Overlapping Clock Generator.

Furthermore, switching blocks S_{on} and S_{op}, shown in figure 4.7, were implemented with n pass transistors and p pass transistors, respectively, all with minimal dimensions, since the ON resistance R_{ON} was negligible in their case.

On the other hand, blocks S_{ps} (ps stands for Port Swapping) and S_{fb} were implemented with pass gates, since they are expected to deal with voltages from ground to the positive rail. In addition, in this case R_{ON} was to be kept as low as possible. For S_{ps}, an high R_{ON} would have caused low currents to charge the input capacitors, thus limiting the chopper frequency (they, together with input large parasitic capacitances C_{in}, would form a relatively low frequency pole); a high R_{ON}, in addition, would add a considerable amount of thermal noise to the small input signal. For S_{fb} the problem is even more critical: in fact, their R_{ON} will be in series with the feedback resistors, significantly altering factor β and so DC gain A_0.

The best trade-off between low R_{ON} and dimensions were to set $R_{ON} \approx 500\,\Omega$ when $V_S = 0.9\,\mathrm{V}$, that is in the worst condition for a pass gate: this would cause a gain error, that if needed could be corrected by resizing the resistors in the feedback network from 199 to $198.5\,\mathrm{k}\Omega$.

In the end, the final dimensions for pass gate transistors resulted, after simulations: $W/L = 3\,\mu\mathrm{m}/0.18\,\mu\mathrm{m}$ for nMOSFETs and $W/L = 15\,\mu\mathrm{m}/0.18\,\mu\mathrm{m}$ for pMOSFETs.

Chapter 5

Simulations and results

In this chapter, we are going to discuss the most relevant simulations, in order to extract the principal electrical characteristics of the proposed InAmp and to compare them to the original goals. All analyses were made with the simulator Eldo by Mentor Graphics.

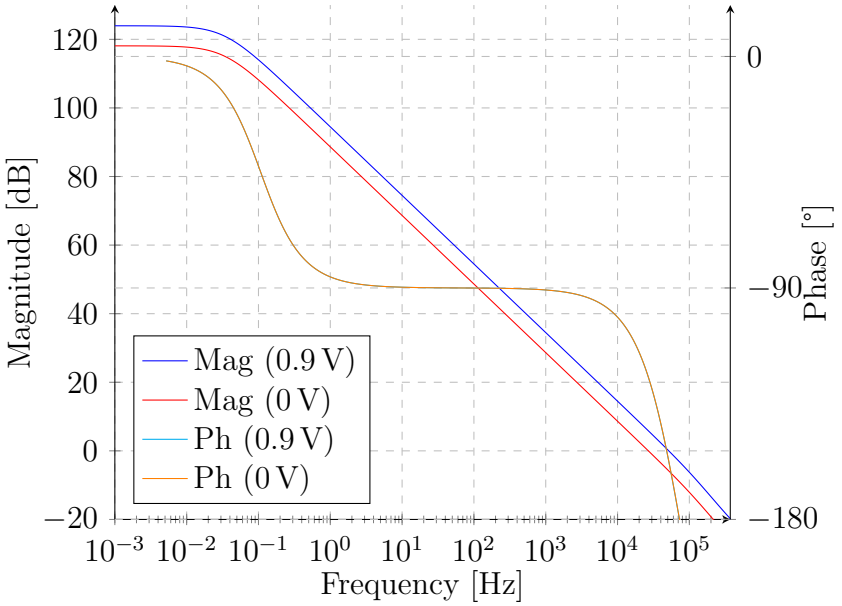
5.1 Frequency domain

5.1.1 Intergrators' open loop transfer functions

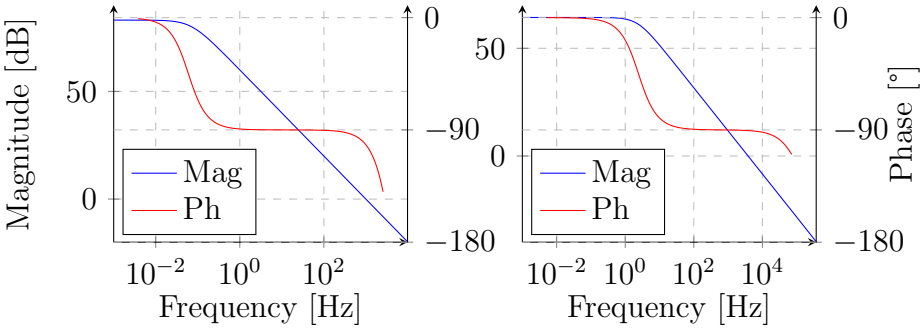
Figure 5.1 shows the open loop frequency responses of the three integrators, the basic elements for the filtering behaviour of the InAmp. From the plots, it is possible to evaluate the finite (but very high) DC gain of the non ideal integrators, and above all the unity-gain frequencies: for INT1, responses with different input CM voltages (0.9 V and 0 V) are reported, in order to show how its unity-gain frequency depends on the input CM. The plots for $V_{inCM} = 1.8\,\mathrm{V}$ are not reported, since they are very similar to the one for 0 V. It is worth to remind that the filter was designed to be a third order Butterworth LPF for an input CM voltage of 0.9 V, while for other values the variation in ω_{01} is expected to alter the filter's frequency response. Table 5.1 shows the unity-gain frequencies ($f_{0i} = \omega_{0i}/2\pi$), pole frequencies f_p and DC gains that can be evaluated from the AC simulations:

	INT1 (0.9 V)	INT1 (0 V)	INT2	INT3
A_{DC}	124 dB	118 dB	83 dB	64.3 dB
f_p	33 mHz	33 mHz	70 mHz	2.33 Hz
f_0	51.5 kHz	26.8 kHz	997 Hz	3.84 kHz

Table 5.1: The three integrators' main parameters.



(a) INT1, with different input CM voltages



(b) INT2

(c) INT3

Figure 5.1: Open loop frequency response of the three integrators.

5.1.2 Signal Transfer Function

Then, figure 5.2 shows the full system's frequency response, obtained with the three integrators in closed loop to form the State Variable Filter. The plot shows also how the input CM voltage affects the frequency response, since it is reported for $v_{inCM} = 0\,\text{V}$ and $0.9\,\text{V}$: both responses are compared with the ideal ones, obtained with the Python functions.

As we can see, cutoff frequencies are slightly higher than expected. The following values can be found for cutoff frequencies ($-3\,\text{dB}$):

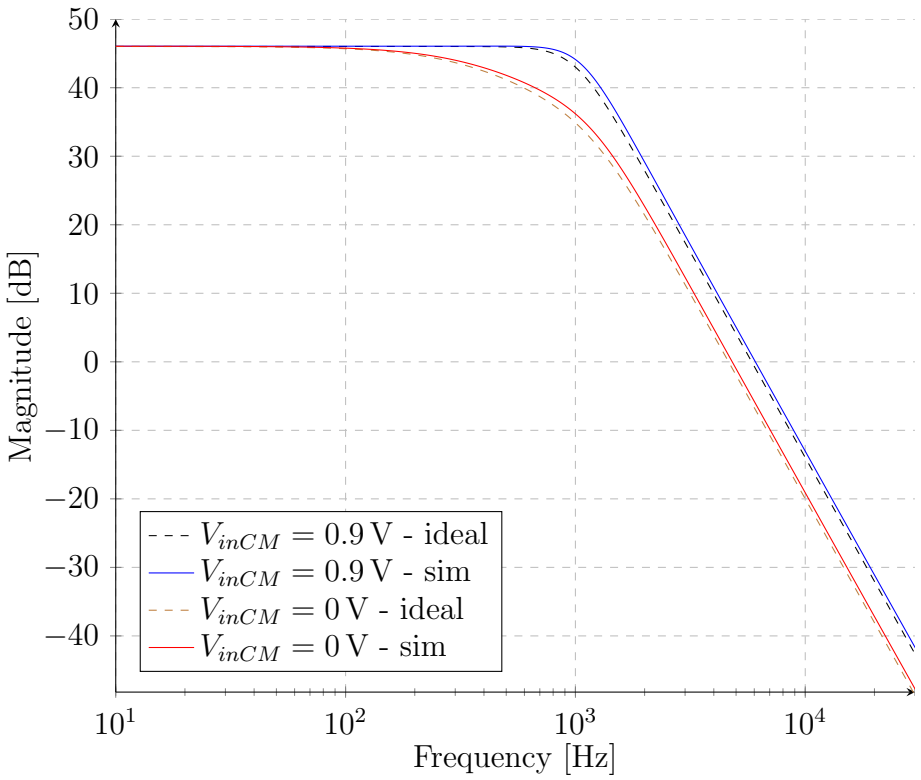


Figure 5.2: Frequency response.

$$v_{inCM} = 0.9\,\mathrm{V}: \qquad f_c = 1096\,\mathrm{Hz} \tag{5.1}$$

$$v_{inCM} = 0\,\mathrm{V}: \qquad f_c = 390\,\mathrm{Hz} \tag{5.2}$$

For both curves, DC gain can be evaluated in 200.864=46.058 dB (typical). Then, we must notice that, with v_{inCM} significantly different from 0.9 V, only one of the three unity-gain frequencies (ω_{01} in particular) changes, so the filter is no more of a Butterworth kind: thus, not only the cutoff frequency changes, but also the shape of the frequency response is different.

5.2 Time domain

5.2.1 Step response

A step response was evaluated and compared to the ideal one, found with the Python functions. The input signal was a 1 mV differential step over a 0.9 V common mode input voltage. Figure 5.3 shows the differential output, compared to the ideal response.

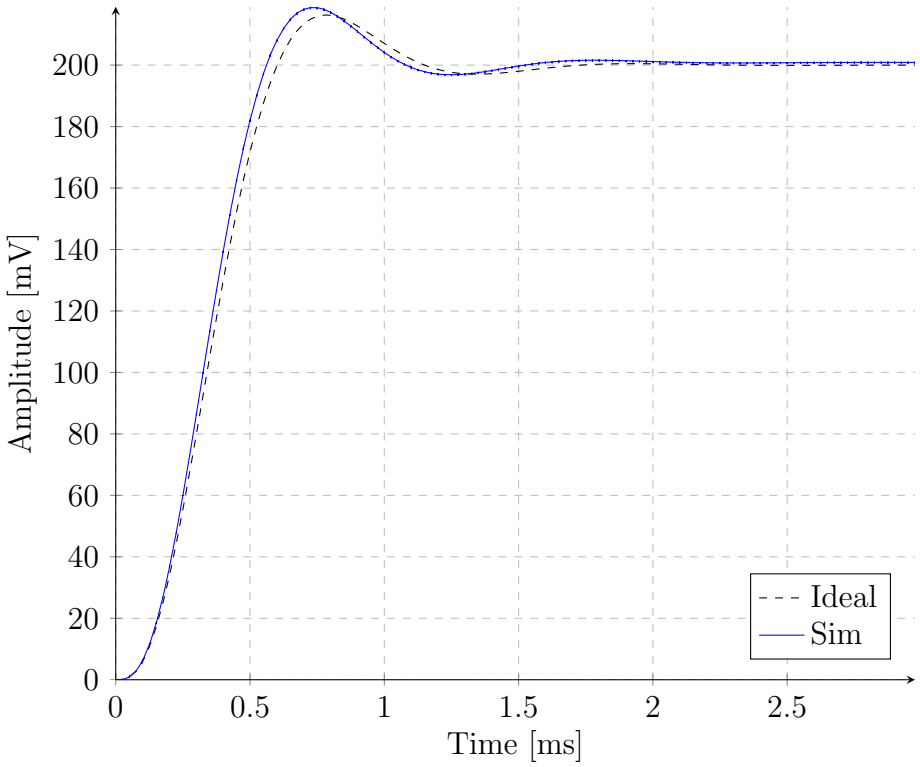


Figure 5.3: Step response.

As we can see, simulated response is slightly faster than expected: this is because the cutoff frequency resulted higher than the expected value of 1 kHz. However, the shape of the output signal (overshoot, damped oscillations) are very close to the ideal one, testifying the correct implementation of the Butterworth filter.

Figure 5.4, instead, shows the internal stages' signals: in particular, it is worth noticing INT1 and INT2 outputs, with half the amplitude of the output voltage, and the chopped output of the preamplifier. The latter, in particular, shows how the preamplifier produces large output signals only

in transients (that is, at high frequencies), while at steady-state the virtual short circuit tends to null its input and output voltages.

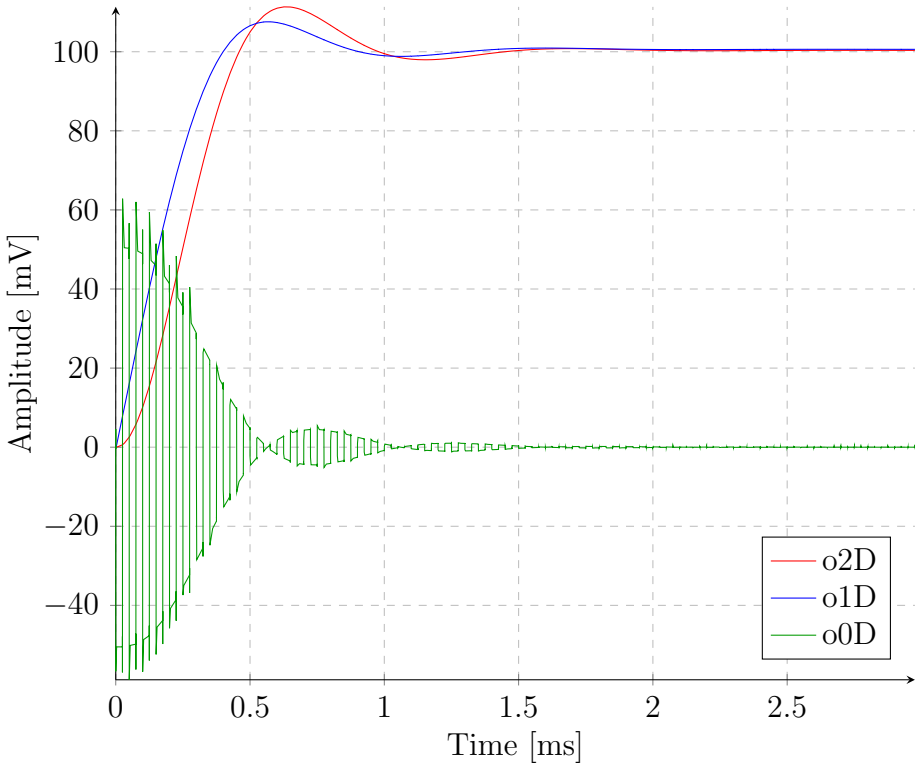


Figure 5.4: Inner stages' signals for a step response.

A particular on the steady-state output differential signal is shown in figure 5.5. Indeed, figure 5.3 could not highlight the small voltage spikes present on the output, that however are worth noticing.

The positive spike is caused by the charge injection of the feedback modulator S_{fb}: when its transistors turn off, part of the channel charge flows into the output capacitances and cause the positive spikes. The larger negative spike, instead, is due to the fast inversion of the feedback signal: the feedback resistors, indeed, have a relevant parasitic capacitance towards ground, that must be charged almost instantaneously at each clock transition. The finite output resistance of INT3 then causes this negative spikes. The introduction of capacitors C_4 and C_4' on the output has mitigated this effect, since they are a natural charge storage and can supply impulsive currents, that do not have to be supplied completely by the output stage of INT3.

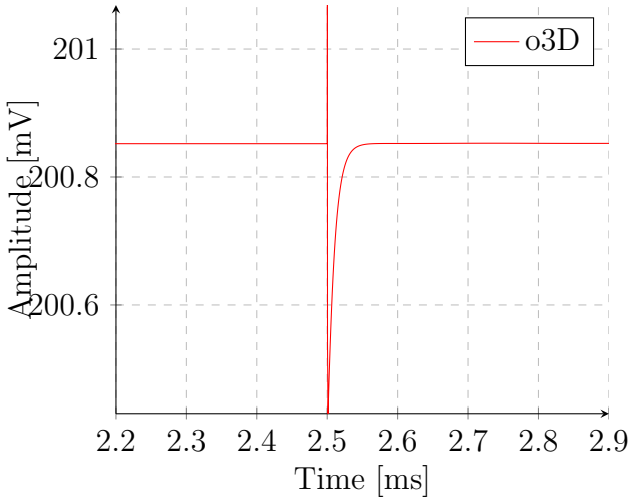


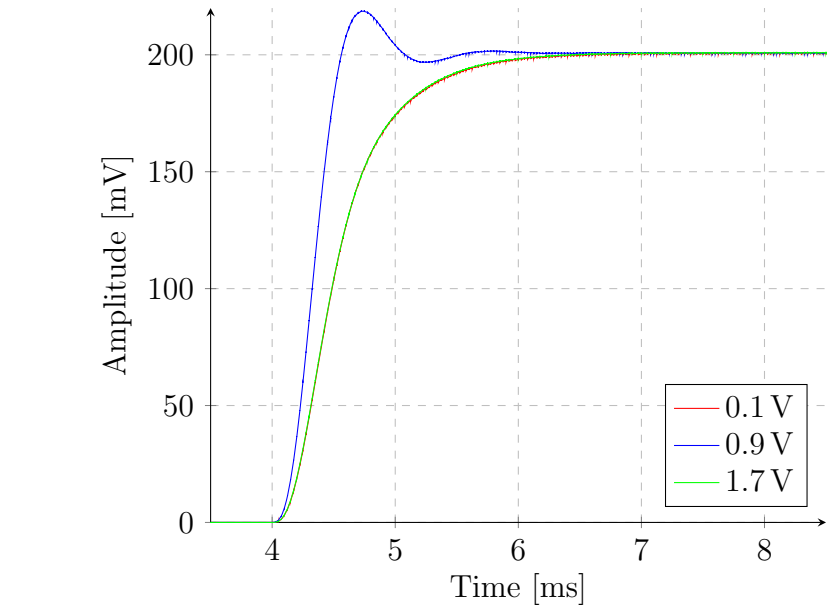
Figure 5.5: Voltage spikes in the differential output.

5.2.2 Step response vs input CM and supply voltage

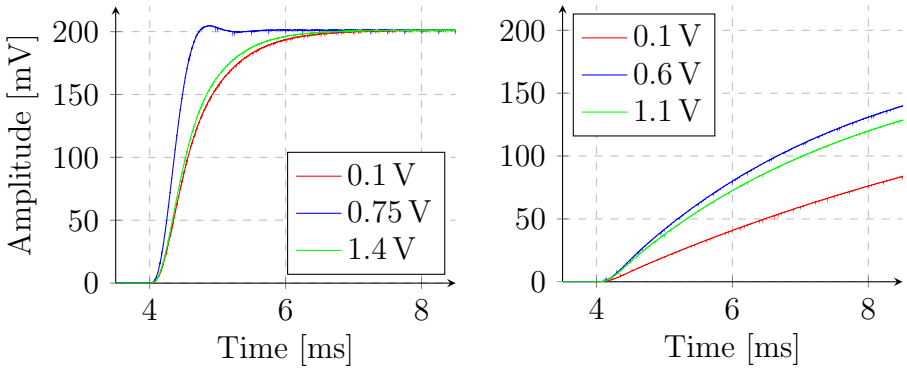
The purpose of the following plots is to show how different input CM voltages affect the operation of the proposed InAmp, and to test it with different supply voltages. As we can see from figure 5.6a ($V_{DD} = 1.8\,\text{V}$), step response changes with input CM voltages different from $V_{DD}/2$, since INT1 unity-gain frequency changes and the filter is no more of a Butterworth kind (see also figure 5.2). However, it is evident that the settling time is still comparable and the performaces are not significantly changed.

The same can be said for a 1.5 V supply voltage (figure 5.6b), that is the voltage of a standard alkaline AA batteries. This time, also the step response simulated with $v_{inCM} = V_{DD}/2 = 0.75\,\text{V}$ is slightly different from the classical Butterworth response, but it is evident that the performance is not compromised for neither of the three signals.

Unfortunately, the proposed InAmp did not pass the test with a $V_{DD} = 1.2\,\text{V}$ (the voltage of standard AA rechargeable batteries), as tesified by the bad looking responses in figure 5.6c. However, this had to be expected: since we designed the input pseudo-differential pair of OTA1 to work with a $V_{GS} = 1\,\text{V}$, this means that its input CM voltage with this supply voltage would be only 0.2 V. This value is low enough to turn completely off, for instance, the output summing stage of the preamplifier, thus compromising the performances of the whole system.



(a) $V_{DD} = 1.8\,\mathrm{V}$



(b) $V_{DD} = 1.5\,\mathrm{V}$ (c) $V_{DD} = 1.2\,\mathrm{V}$

Figure 5.6: Step responses with different supply voltages and V_{inCM} (values in the legends).

5.2.3 Input impedance

It is interesting to obtain an estimation of the input impedance, and appreciate the effectiveness of the Port Swapping technique.

Now, since the input modulator is responsible for chopping and port swapping at once, the effects of the two techniques cannot be evaluated one at a time. However, reminding the relation for the input impedance in a chopped amplifier

$$R_{inD}^{(ch)} = \frac{1}{4C_{inD}f_{ch}} \tag{5.3}$$

we could find a value for the capacitance $C_{inD} = C_{in}/2$ with a simple AC analysis (with chopping disabled) and find a value for $R_{inD}^{(ch)}$: this value, then, can be compared with the value obtained with a transient analysis, that takes into account both chopping and port swapping. With port swapping, as said in paragraph 3.1, input resistance has the following expression, for frequency that tends to 0 (with the necessary introduction of β_2, not present in the reference design):

$$R_{inD}^{(ch-ps)} = \frac{\frac{\beta}{\beta_2}A_{DC1}}{2C_{inD}f_{ch}} = 2\frac{\beta}{\beta_2}A_{DC1}R_{in}^{(ch)} \tag{5.4}$$

We can evaluate this expression as well and compare it with the simulation result. From an AC simulation of the input current with chopping disabled, we can obtain the value of the input differential capacitance:

$$C_{inD} = \frac{C_{in}}{2} = 10.5\,\mathrm{pF} \tag{5.5}$$

This value, with chopping enabled, would give an input resistance of

$$R_{inD}^{(ch)} = \frac{1}{4C_{inD}f_{ch}} = 1.2\,\mathrm{M\Omega} \tag{5.6}$$

It is now possible to estimate, according to eq (5.4), the equivalent input resistance with port swapping enabled, reminding from paragraph that $A_{DC1} = 124\,\mathrm{dB} \approx 1.6 \times 10^6$. We have then

$$R_{inD}^{(ch-ps)} = 2\frac{\beta}{\beta_2}A_{DC1}R_{in}^{(ch)} = 38.4\,\mathrm{G\Omega} \tag{5.7}$$

It is interesting now to find directly the differential input resistance from a simulation and compare these results. One way to evaluate it is to set a transient analysis, set a certain V_{inD} and evaluate the input currents; after an integration of the differential current $I_{inD} = (I_{in-x} - I_{in-y})$ over a complete clock period, it is easy to obtain a mean value for the differential current and so a value for $R_{inD}^{(ch-ps)}$.

Figure 5.7 shows a particular of the differential current during one clock period: it's evident that I_{inD} is not null only during clock switches, while its steady-state value is, logically, zero.

From the plot points, an integration of I_{inD} over one clock period resulted

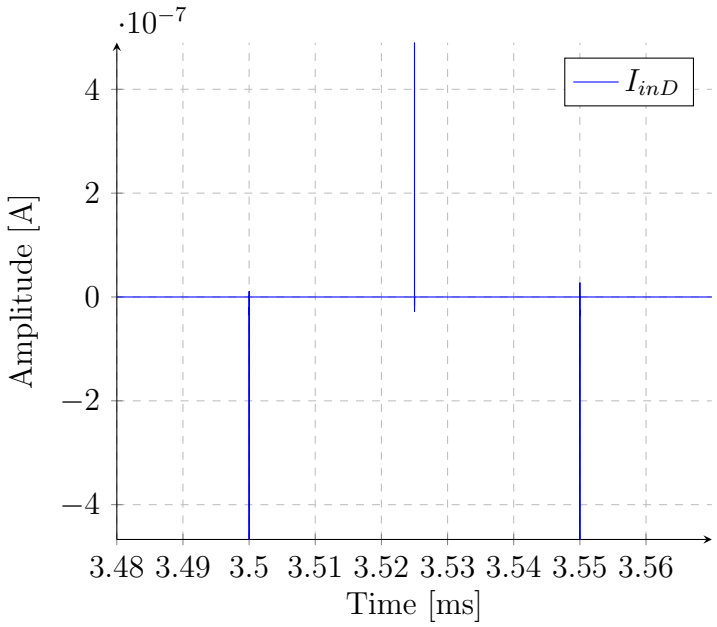


Figure 5.7: Input differential current with 1 mV input.

in a mean value of $< I_{inD} >= 6.43 \times 10^{-14}$ A, for a resulting equivalent input resistance of

$$R_{inD}^{(ch-ps)} = 15.5\,\mathrm{G\Omega} \tag{5.8}$$

It is worth noticing that, with such small values of the quantities in play, this value can be affected by a huge error; however, it is not too far from the value expected from eq. (5.4) (at least the order of magnitude is the same). In any case, it has been proved that port swapping is extremely effective in increasing input resistance, that is 4 order of magnitude higher than it would be with simple chopper modulation.

In addition, this test allowed us to evaluate also the common mode input current I_{inCM}: thus, knowing that $V_{inCM} = 0.9$ V, it is possible to estimate the common mode input resistances, also known as isolation resistances (R_{IS}). After integration over one complete clock period, I_{inCM} showed a mean value of 5.68×10^{-15} A, for a resulting value of insulation resistances of

$$R_{IS} = 5.7 \times 10^{15}\,\Omega \tag{5.9}$$

The accuracy of this value, again, is most probably compromised by

numerical error: anyway, the important target was to verify that it had an extremely high value.

5.2.4 High frequency distortion

A transient analysis was used to evaluate qualitatively the distortion of high frequency, large signals that was one of the most critical issues of the reference design. More than one time, this problem was mentioned in chapter 4: thus, a differential input sinewave was set at 700 Hz frequency and 8 mV amplitude over a 0.9 V common mode (this is one of the worst cases). Figure 5.8 shows the output signal, compared with an ideal sine waveform (generated with a Python function) with same amplitude, frequency and phase. The two waveforms are evidently very similar: of course, an accurate analysis of distortions would require a Fourier transform and the study of THD. For our purpose, a visual comparison between the two waveforms can be sufficient to state that no significant distortions are present, even for critical signals like this one.

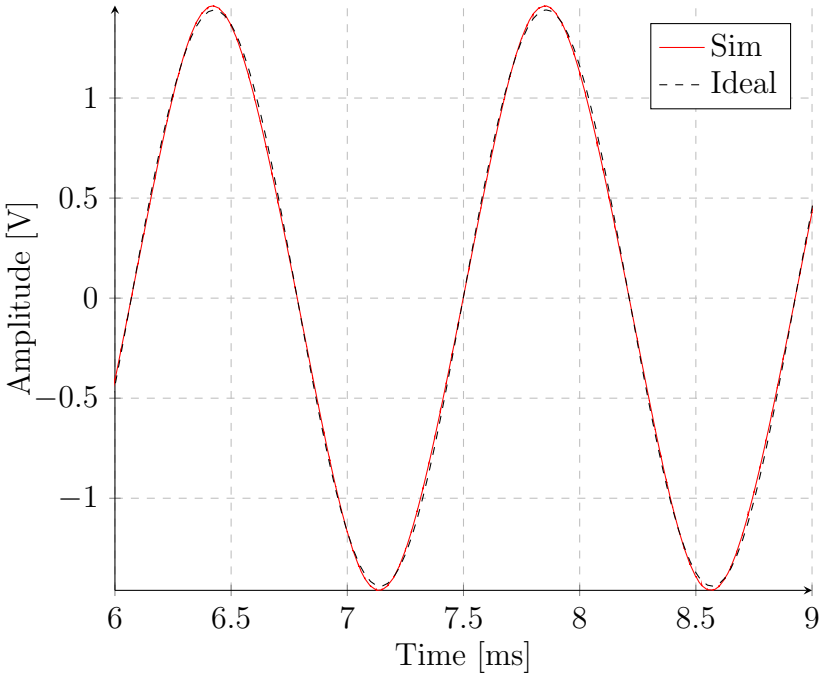


Figure 5.8: Comparison between output and an ideal sine waveform.

It is interesting, finally, to observe the signals produced by the inner stages, that are preamplifier, OTA1 and INT2: the phase shift between the

signals is evident from figure 5.9 (and also between these and the output, in figure 5.8). It must be noticed that, at these frequencies, the preamplifier's output - and input as well - is not null, since virtual short circuit is not valid when phase shift is not negligible. However, as expected, all this signals are always within the range of their own stages, thus minimal distortion is introduced even at high frequencies and large amplitudes.

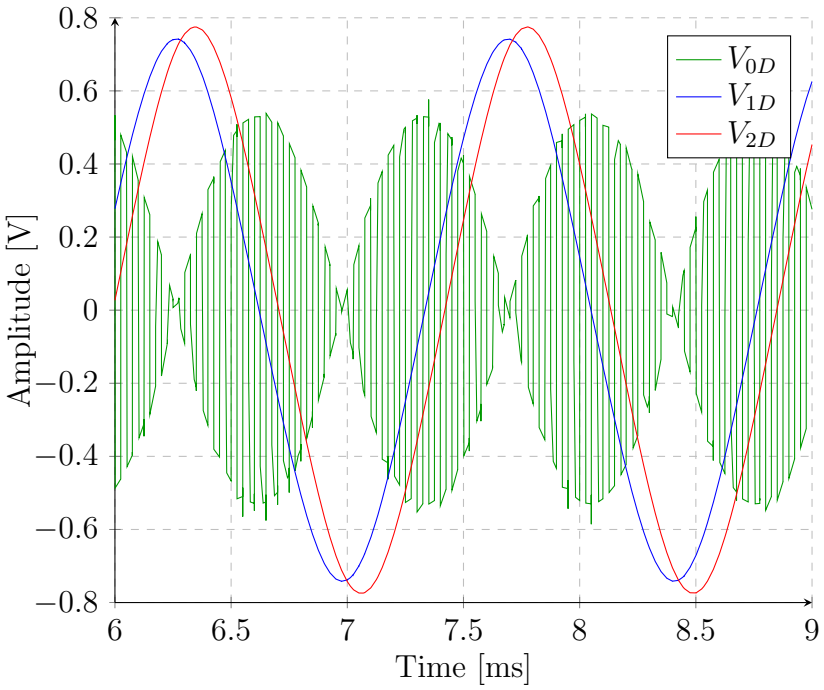


Figure 5.9: Inner stages signals.

5.2.5 CM-differential intermodulation

As expected from paragraph 3.2.3 and proved in paragraph 5.1.2, input CM affects the InAmp's frequency response for frequencies near f_c (actually from around 300 Hz up is perceptible). Now, suppose that we are dealing with a signal at relatively high frequency, for intance at $f_w = 700$ Hz, and the input common mode is affected by a large disturbance: the classical example can be the 50 Hz mains voltage. It is easy to understand that, since $STF(f_w)$ depends strongly on the input CM voltage (figure 5.2), the input wave will be amplified by a different factor depending on the instantaneous amplitude of the CM disturbance, resulting in a sort of intermodulation between CM and differential paths.

Figure 5.10 represents the situation in which a 1 mV, 700 Hz sine wave is presented at the differential input, while a 500 mV, 50 Hz sine wave (disturbance) is superimposed on the DC value of the input CM voltage, that is 0.9 V, as usual. We can see clearly from the simulation that the differential output is modulated in amplitude by the disturbance.

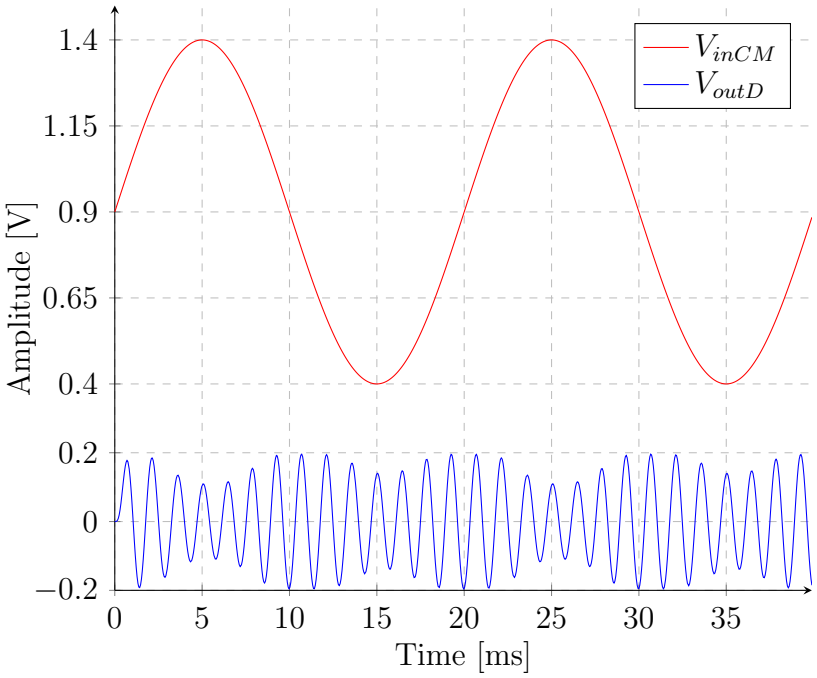


Figure 5.10: Intermodulation between common mode and differential paths.

5.3 Differential ranges

The most straightforward way to estimate output differential range would be a DC sweep on the input voltage source, in order to obtain a relation $V_{outD} = f(V_{inD})$ and verify its linearity over a certain range. However, due to convergence problems during the simulation, a different approach was used: a series of DC bias points were simulated, for input differential voltages from −12 mV to 12 mV with 25 µV steps. These points were then plotted to obtain the transcharacteristic function, shown in figure 5.11.

As we can see, the output saturates at around ±1.71 V: therefore, a good linearity is provided evidently at least for an output differential range of ±1.6 V, that is for input signals between ±8 mV.

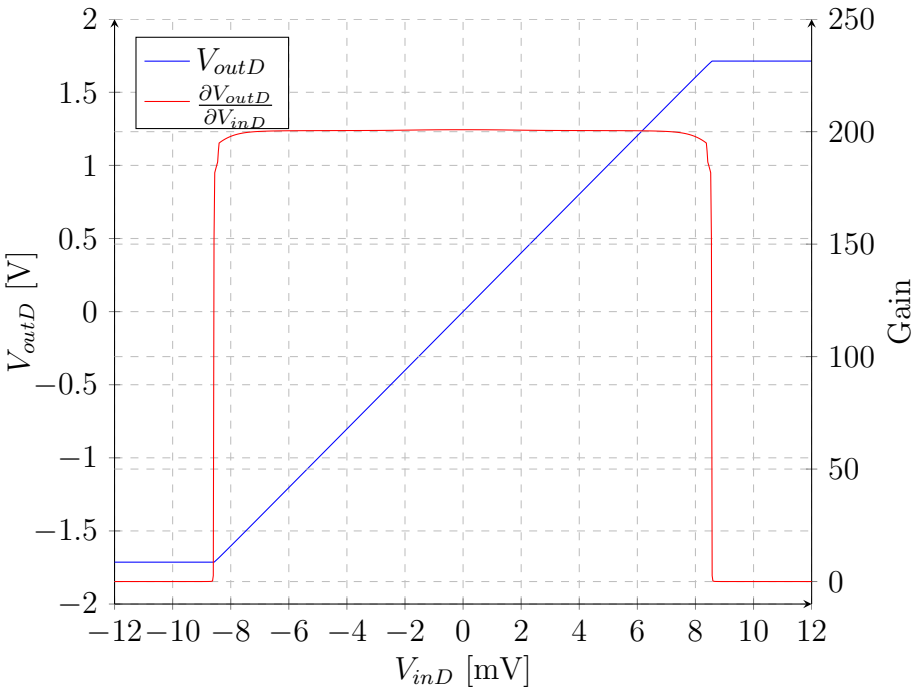


Figure 5.11: Voltage spikes in the differential output.

5.4 Noise

5.4.1 Noise density, power and dynamic range

Figure 5.12 shows the output noise density vs. frequency: the solid lines were obtained from a PSS simulation, to verify the effect of the chopper modulation on the output noise, for different corners. The dashed line, instead, represents the output noise density that would be present without chopping modulation (then, a simple AC simulation was used for it): as evident from the plots, chopper modulation is extremely effective in flicker noise rejection.

For the TYP noise density, the following parameters can be measured:

$$v_{n-floor} = 3.91\,\mu\text{V}/\sqrt{\text{Hz}}: \qquad f_k = 5.5\,\text{mHz} \tag{5.10}$$

Notice that this extremely low flicker corner frequency was obtained by enlarging OTA1 common gates, the only devices of the first integrator whose noise contribution is not modulated. However, since noise from common

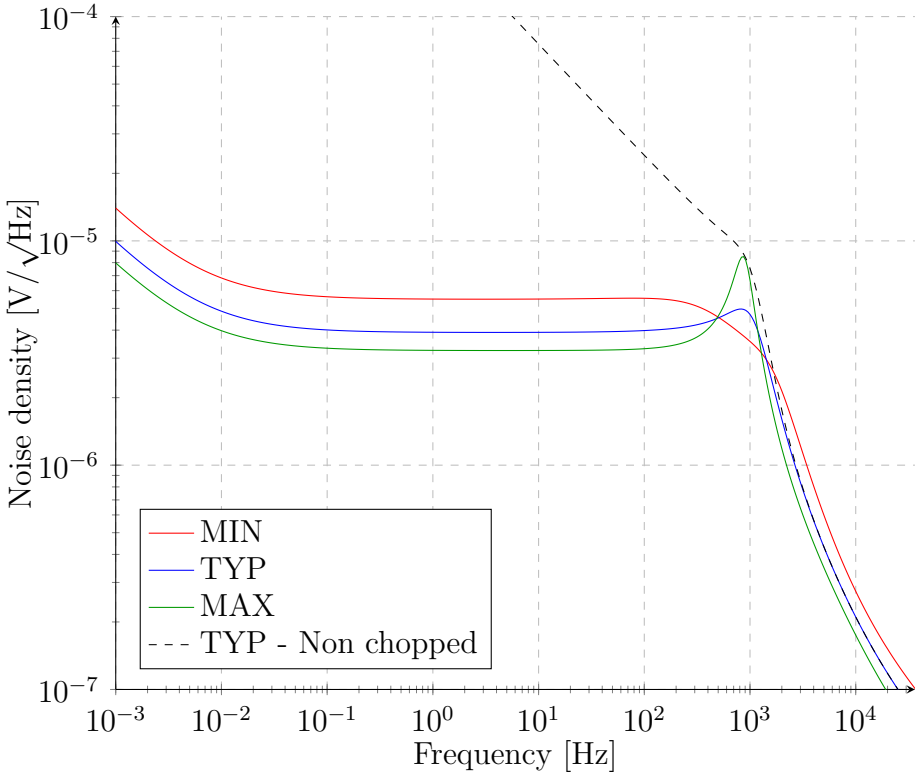


Figure 5.12: Output noise density vs corners.

gates is already strongly reduced around a hundred times compared to the other devices, not extremely large areas were needed.

Regarding the baseband noise floor, it is interesting to evaluate the equivalent RTI noise density. Dividing it by the nominal A_0 found by the previous simulations, we obtain

$$v_{nRTI} = 19.5\,\text{nV}/\sqrt{\text{Hz}} \tag{5.11}$$

We have then reached our original goal of keeping RTI noise density lower than $20\,\text{nV}/\sqrt{\text{Hz}}$ (typical).

Figure 5.13, finally, shows how the input CM voltage affects the noise. When it is different from $V_{DD}/2$, only one type of pairs is active: the other is off and it will not give noise contributions, but since G_{m0} is now half the original value, the overall RTI noise density will be increased. For $v_{inCM} = 0\,\text{V}$ and $1.8\,\text{V}$ it reaches a value of around $6.4\,\mu\text{V}/\sqrt{\text{Hz}}$, that is a RTI value of $31.8\,\text{nV}/\sqrt{\text{Hz}}$.

It is interesting now to evaluate the total output noise power, in order to

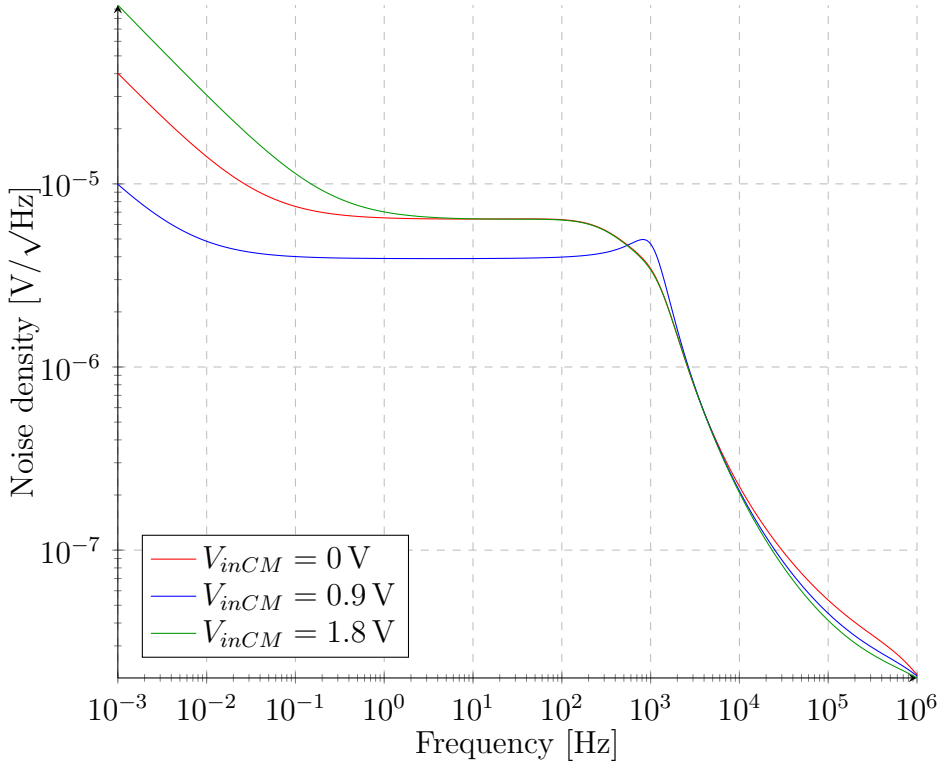


Figure 5.13: Output noise density vs input CM voltage.

estimate the Dynamic Range and the Equivalent Number Of Bits. Suppose that a traditional Nyquist ADC is employed to sample a signal, whose bandwidth is equal or less than 1 kHz. Then, noise at higher frequency will undergo foldover, so all the output noise power will be superimposed on the sampled signal. Thus, integrating the simulated output noise PSD $S_n(f)$ (that is noise density squared) from 1 mHz to 1 MHz, we obtain a total output noise power of

$$P_{n-out} = \int_{1\,\mathrm{mHz}}^{1\,\mathrm{MHz}} S_n(f)df = 33.34\,\mathrm{nW} \tag{5.12}$$

that, divided by the DC gain squared, gives the total equivalent input noise power

$$P_{nRTI} = \frac{P_{n-out}}{A_0^2} = 0.826\,\mathrm{pW} \tag{5.13}$$

The square root of this value represents the RMS value of RTI noise:

$$v_{nRTI} = \sqrt{P_{n-out}} = 0.91\,\mu\text{V} \tag{5.14}$$

Then, if we consider a peak-to-peak signal equal to 4 times its RMS value (that is, a noise band of $\pm 2\sigma$), we can compare it to the maximum input differential range and obtain a value for the Dynamic Range. We have then

$$DR = \frac{2v_{Dmax}}{4v_{nRTI}} = 4400 \tag{5.15}$$

that gives us an Equivalent Number of Bits of

$$ENOB = \lfloor \log_2 DR \rfloor = \lfloor 12.1 \rfloor = 12 \tag{5.16}$$

Actually, if a higher resolution is needed, a high rate oversampling ADC (like a $\Delta\Sigma$) can be used instead of a classical Nyquist converter. Oversampling reduces foldover effect, then digital filtering eliminates noise from outside the band of interest. Suppose we are dealing with a 200 Hz bandwidth input signal (this was the bandwidth of the reference design) and we use a $\Delta\Sigma$ ADC with oversampling ratio $OSR = 256$. Then we have an effective sampling frequency of $f_{os} = 51.2\,\text{kHz}$: thus, only noise at frequencies higher than f_{os} will be shifted to the baseband. It is evident that this contribution will be negligible, so we can estimate the total equivalent RTI noise power as

$$P_{nRTI} = \frac{1}{A_0^2} \int_{1\,\text{mHz}}^{200\,\text{Hz}} S_n(f) df = 74.9\,\text{fW} \tag{5.17}$$

that gives an equivalent RTI RMS noise of

$$v_{nRTI} = \sqrt{P_{n-out}} = 0.273\,\mu\text{V} \tag{5.18}$$

This means values of DR and ENOB of

$$DR = 14615 \qquad \rightarrow \qquad ENOB = \lfloor 13.8 \rfloor = 13 \tag{5.19}$$

5.4.2 Noise Equivalent Factor (NEF)

As introduced in chapter 1.3, NEF is a figure of merit that is a quantitative parameter of the noise efficiency: basically, it is the ratio between the total RTI noise of an amplifier and the total RTI noise of a bipolar transistor with equal bandwidth and current consumption (i.e., collector current).

According to [1], we remind the original definition of NEF:

$$NEF = \frac{v_{n,ampRTI}}{v_{n,bjt}} = \frac{v_{nRTI}}{\sqrt{B\frac{\pi}{2}\frac{4kTV_T}{I_{DD}}}} \tag{5.20}$$

where B is the bandwidth (we found it to be equal to 1096 Hz), V_T is the thermal voltage and I_{DD} is the quiescent supply current. A DC simulation showed a value for quiescent supply current of $I_{DD} = 345\,\mu A$: thus, we obtain

$$NEF = \frac{0.91\,\mu V}{46.2\,nV} = 19.7 \tag{5.21}$$

Now, this value of NEF can appear relatively high, if compare to the values of the state of the art solutions, whose NEF is usually around 8 or 9. However, it must be noticed that InAmps designed to have a R2R input CM range are usually not particularly efficient: indeed, there are no works of this category in the literature that reports a value for NEF. In fact, a R2R input CM range forced us to design dual input pairs (both n and p) instead of one, and a folded cascode topology that is much less noise-efficient than, for instance, a telescopic cascode topology.

5.5 CMFB loop stability

5.5.1 CMFB open loop frequency response

As discussed in paragraph 4.5.4, great attention was paid in the design of the CMFB loop, to guarantee its stability at all frequencies. As said, a dominant pole compensation has been employed, together with a secondary feedback loop in charge of decreasing the DC gain of the loop: stability was then studied (and achieved) after the analysis of the loop gain frequency response.

In order to avoid the Barkhausen conditions to be satisfied, it is necessary to provide an adequate phase margin: this parameter is equal to the phase of the frequency response, evaluated at the frequency at which gain is 0 dB. In other terms, it is a measure of how much the phase can drop (due to process corners, temperature, etc.) before the rise of oscillations.

Figure 5.14 shows the frequency response of the dominant pole compensated CMFB open loop $\beta A(f) = [A_{kc}A_{cc}A_{CMFB}](f)$, as found in paragraph 4.5.4, obtained with an AC simulation on the cut loop (that is, with a test generator that cuts the loop).

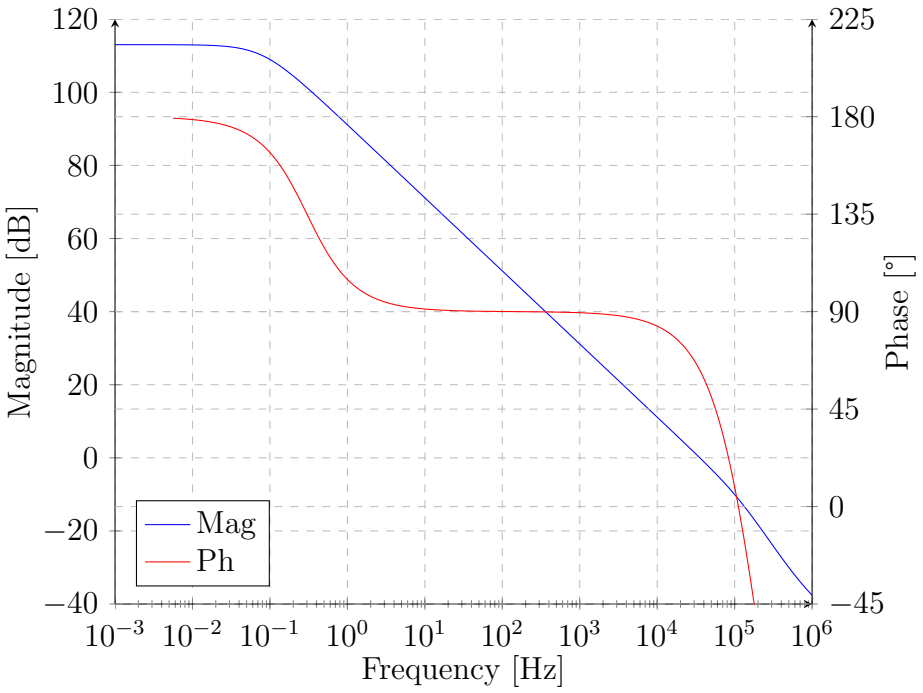


Figure 5.14: CMFB loop frequency response.

This simulation shows the main parameters of the loop. First of all, we have a DC gain of 113 dB: since this value is, in first approximation, inversely proportional to the relative error on the output common mode voltage, it is important that it is as high as possible. Far more important is the phase margin (PM): from the simulation we get

$$PM = 59.5^{\circ} \tag{5.22}$$

that is a good value to guarantee stability. Also the gain margin can be evaluated from this plot, and it turns out to be equal to 11.3 dB.

5.5.2 CMFB step response

As an additional test for stability, a step response can be analyzed. In our case, we imposed a 100 mV voltage step on V_{refCM}, and simulated the output v_{1CM}. As we can see from figure 5.15, the steady-state condition is reached in few tens of µs and no oscillation rises after the stimulus. This, for instance, allows V_{refCM} to be set manually from outside the chip, and even changed during the InAmp's operation, if needed.

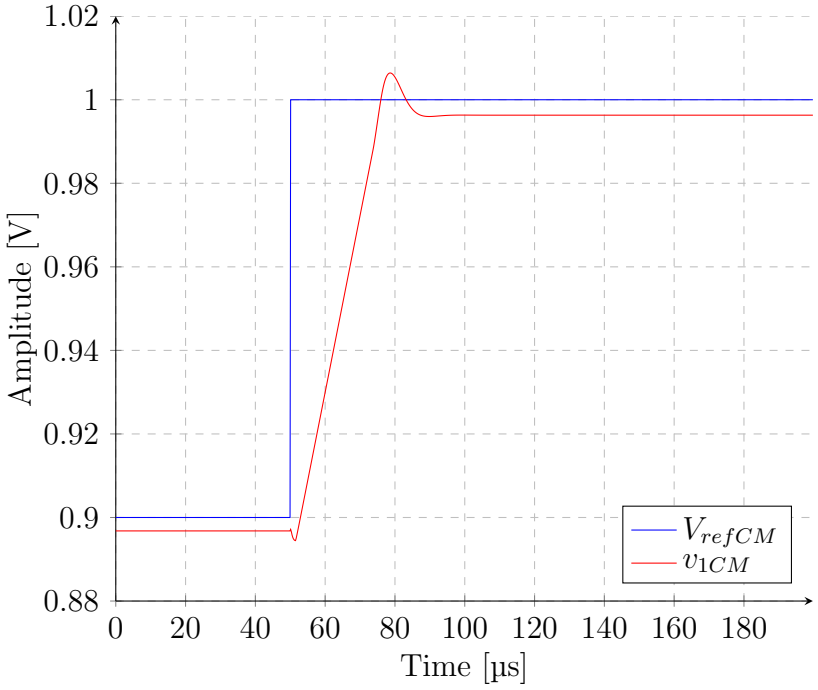


Figure 5.15: Step response of the CMFB loop.

5.6 Montecarlo

Montecarlo simulations are used to evaluate the effect of process errors and local variations (mismatch) of the devices' parameters on the global performances of electronic circuits. Briefly, a Montecarlo simulation is the series of N simulations, where N is the number of the desired samples to test: in each one of them, all parameters of the circuit's devices are varied according to a certain statistical distribution (usually gaussian), with different standard deviations for global and local errors and for the different devices, depending on the process.

Then, each version of the circuit is simulated and a certain output quantity can be described statistically: this is the case, in particular, of offset and gain error, that will be analyzed in the following paragraphs.

For all Montecarlo simulations, a number of samples $N = 100$ was used. Under some points of view, this number may appear insufficient for an accurate statistical analysis. Indeed, typically we evaluate the tail of the gaussian distribution, that is usually set at 3σ: since it represents the value that includes 99.7% of the samples, it is logical that at least a thousand samples are necessary to evaluate it with good precision. On the other hand,

since most simulations (transient in particular) are extremely complex and long, a value of 100 has been chosen as a trade-off between accuracy and simulation time.

5.6.1 Offset

A Montecarlo DC simulation was used to evaluate the distribution of differential output offset voltage V_{OS-out}, with chopper modulation disabled. This value is interesting both to have an idea of the usual offset that affects CMOS circuits without low frequency errors reduction techniques, and above all to compare it with the value obtained from the chopped system and appreciate the effectivness of this technique.

Figure 5.16 shows the distribution of the values of output offset obtained with this Montecarlo simulation.

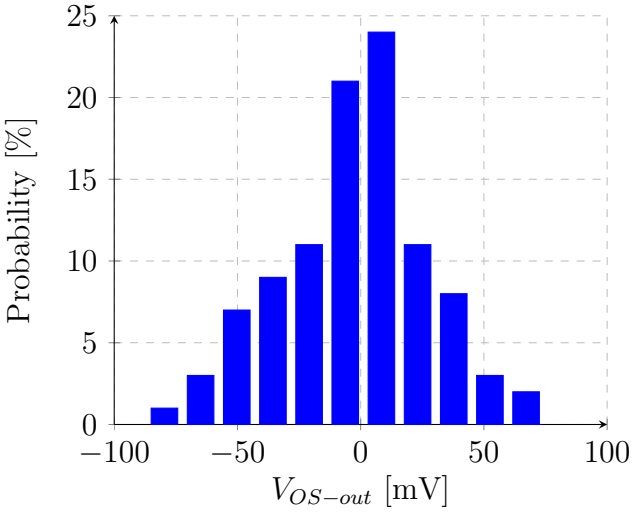


Figure 5.16: Distribution of output offset (chopper disabled).

Then, the output offset was evaluated after enabling chopper modulation. To appreciate the effects of this dynamic technique, a transient analysis was necessary: thus, an input differential voltage $V_{inD} = 0\,\mathrm{V}$ was set and the output voltage was measured after a few ms, to be sure to have reached the steady-state. The output residual offset for each Montecarlo iteration was evaluated as the average value of the output voltage over a complete clock cycle: this way we can take into account the effect of the voltage spikes and analyze the effective voltage that is being processed by

the system (that, being a LPF, is sensible to mean values). Again 100 samples were tested: the distribution of their output offset is showed in figure 5.17

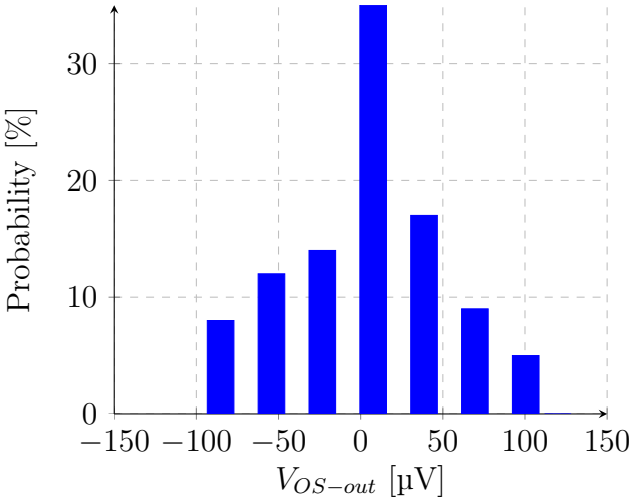


Figure 5.17: Distribution of output offset (chopper enabled).

As evident from the histograms, chopper modulation is extremely effective, and reduces offset by almost three orders of magnitude. Table 5.2 reports the standard deviations of these distibutions, together with the one relative to the equivalent input offset, obtained dividing V_{OS-out} by the nominal gain $A_0^{(nom)} = 200.864$.

	No chopper	Chopper
$\sigma_{V_{OS-out}}$	30 mV	47 µV
$\sigma_{V_{IO}}$	150 µV	0.23 µV

Table 5.2: Offset σ with and without chopper modulation.

It can be noticed that, according to this statistics, 99.7% (3σ) of the actual chips will have an RTI offset less than 0.7 µV.

5.6.2 Gain error

Similarly, 100 samples were tested and their gains were evaluated, both with and without chopper modulation and port swapping. In this case, a simple AC analysis was set to evaluate gain with chopper disabled, and the values

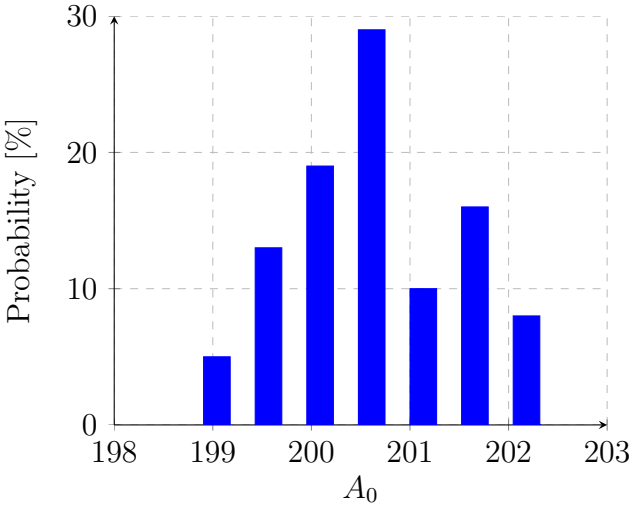


Figure 5.18: Distribution of gain values (chopper and port swapping disabled).

for $f = 1\,\text{mHz}$ were used as DC gain. Figure 5.18 shows the distribution of gain values among the 100 samples.

Gain distribution in the chopped amplifier's samples was evaluated in the same way as for offset: indeed, a transient analysis was employed to study the output differential voltage of the sample, having 1 mV as input differential voltage. An interval of time of 3.5 ms was considered in order to let the system reach the steady-state condition. After that, the output signals were averaged over one clock cycle to extract the mean values of the output voltages, that divided by the input voltage gave the effective gain of each one of the samples. Figure 5.19 shows the distibution of the values obtained by this simulation.

The standard deviations of these distributions are again reported in table 5.3: here, the relative gain error is obtained dividing the gain errors by the nominal gain $A_0^{(nom)} = 200.864$.

	No chopper	Chopper
$\sigma_{\Delta A_0}$	0.85	0.88
$\sigma_{\frac{\Delta A_0}{A_0}}$	0.42%	0.44%

Table 5.3: Gain σ with and without chopper modulation.

As evident from the table, gain error does not change significantly whether chopper and port swapping are working or not. This can suggest us that, in

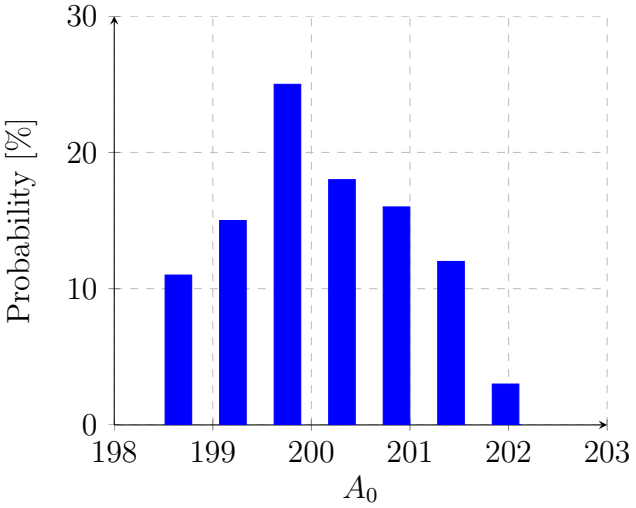


Figure 5.19: Distribution of gain values (chopper and port swapping enabled).

our case, gain error is mostly due to mismatch between the feedback resistors: this one, indeed, was not excpected to be eliminated by port swapping. In case we need higher gain accuracy, a solution could be to increase the area of the resistors, keeping equal their proportions (and thus the values of resistance).

5.7 Electrical parameters summary

In conclusion, all the results of the simulations are summarized in table 5.4. Unless otherwise noted, the test condition for the simulations were $V_{DD} = 1.8\,\mathrm{V}$ and $V_{inCM} = 0.9\,\mathrm{V}$. All values are typical.

Symbol	**Parameter**	**Test cond.**	**Value**	**Unit**
A_0	DC Gain	$f = 1\,\text{mHz}$	200.864	
f_c	Cutoff freq. ($-3\,\text{dB}$)	$V_{inCM} = 0.9\,\text{V}$	1096	Hz
f_c	Cutoff freq. ($-3\,\text{dB}$)	$V_{inCM} = 0\,\text{V}$	390	Hz
V_{DD}	Supply voltage	$V_{inCM} = 0$ to $1.8\,\text{V}$	1.5 to 1.8	V
R_{inD}	Diff. input resistance		15.5	GΩ
R_{IS}	Isolation resistance		5.7×10^{15}	Ω
V_{outD}	Diff. output swing		±1.7	V
v_n	RTI noise density	$f = 1\,\text{Hz}$	19.5	$\text{nV}/\sqrt{\text{Hz}}$
f_k	Corner frequency		5.5	mHz
v_n	RTI noise RMS	1 mHz to 200 Hz	0.273	µV
DR	Dynamic Range	1 mHz to 200 Hz	14600	
ENOB	Equivalent Nb. Of Bits	1 mHz to 200 Hz	13	
I_{DD}	Supply current	$v_{inD} = 0\,\text{V}$	345	µA
NEF	Noise Efficiency Factor		19.7	
PM	CMFB Phase Margin		59.5	°
V_{IO}	Input offset	σ	0.23	µV
$\frac{\Delta A_0}{A_0}$	Gain error	σ	0.44	%
	Total area	(active areas)	$< (250)^2$	µm

Table 5.4: Simulated electrical parameters summary.

Conclusions and future developments

In this thesis, we proposed an original design of a current feedback intrumentation amplifier, optimized for the readout of integrated thermal sensors. After an overview on the state of the art, a high level analysis allowed the choice of the topology and a first sizing of the blocks, taking into account the main issues of CMOS technology and of the current feedback architecture. After that, an accurate study of the transistor level schematics led to the dimensions of the single devices.

As testified by the simulations discussed in chapter 5, the proposed design proved to meet the expected specifications. fFicker noise and offset were successfully rejected by chopper modulation, and the values of residual offset, in-band noise floor and corner frequency resulted adequate to the requirements of precision applications. Port swapping, together with common mode equalization, achieved a good level of gain accuracy, resulting as well an excellent means to increase input impedance. Even current consumption resulted relatively low, considered that Rail-To-Rail topologies are rarely optimized for low power, thus making this solution suitable even for mobile applications.

One of the few problems left in this work is the variation of INT1 transonductance depending on input common mode voltage: this variation, while not compromising gain accuracy thanks to common mode equalization, affects noise floor and the frequency response at the band's limit. A future development of this project could take this issue into account, and try to mitigate it, for example, with the use of constant-g_m input transconductors.

The next step in the design of this system would be an accurate design, that would have to pay attention to symmetries in the devices' place and route, in order to minimize the effects of local errors (i.e., mismatch) of the process.

Finally, the last step would be the tapeout and the full characterization of the prototype chips produced by the foundry.

Appendix A

Acronyms

AC	alternate current
ADC	analog to digital converter
AFE	analog front end
AMS	analog-mixed signal
AZ	auto-zeroing
BPF	band pass filter
CDS	correlated double sampling
CFIA	current feedback instrumentation amplifier
CHS	chopper stabilization
CM	common mode
CMDA	common mode difference amplifier
CMEQ	common mode equalization
CMFB	common mode feedback
CMRR	common mode rejection ratio
CT	continuous time
DAC	digital to analog converter
DAS	data acquisition system
DC	direct current
DEM	dynamic element matching
DoF	degree of freedom
DR	dynamic range
DT	discrete time
ENOB	equivalent number of bits
GBW	gain-bandwidth (product)
GERL	gain error reduction loop
HPF	high pass filter
IA	instrumentation amplifier
LPF	low pass filter

MEMS	micro-electro-mechanical system
NEF	noise efficiency factor
NTF	noise transfer function
OA	operational amplifier
ORL	offset reduction loop
OTA	operational transconductance amplifier
PM	phase margin
PS	port swapping
PSD	power spectral density
R2R	rail-to-rail
RMS	root main square
RRL	ripple reduction loop
RTI	referred to input
SC	switched capacitor
SoC	system on a chip
STF	signal transfer function
SVF	state variable filter
THD	total harmonic distortion

Bibliography

[1] Steyaert MSJ, Sansen WMC, and Chang Z. A micropower low-noise monolithic instrumentation amplifier for medical purpose. *IEEE Journal of Solid State Circuits*, SC-22(6):1163–1168, 1987.

[2] Wu R, Huijsing JH, and Makinwa KAA. *Precision Instrumentation Amplifiers and Read-Out Integrated Circuits*. Springer Science, 2013.

[3] Martin K, Ozcolak L, Lee YS, and Temes GC. A differential switched-capacitor amplifier. *IEEE Journal of Solid-State Circuits*, SC-22(1):104–106, 1987.

[4] Fan Q, Huijsing JH, and Makinwa KAA. A 1.8uw 1uV-offset capacitively-coupled chopper instrumentation amplifier in 65nm CMOS. *IEEE ESSCIRC, digital technical papers*, pages 170–173, 2010.

[5] Schaffer V, Snoeij MF, Ivanov MV, and Trifonov DT. A 36 v programmable instrumentation amplifier with sub-20 uV offset and a CMRR in excess of 120 dB at all gain settings. *IEEE Journal of Solid-State Circuits*, 44(7):2036–2046, 2009.

[6] Krabbe H. A high-performance monolithic instrumentation amplifier. *IEEE ISSCC, digital technical papers*, pages 186–187, 1971.

[7] Bruschi P. Notes on mixed signal design. Available at `http://docenti.ing.unipi.it/~a008309/mat_stud/MIXED/2014/Dispense/`.

[8] Enz CC and Themes GC. Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization. *IEEE vol 84(11)*, 84(11):1584–1614, 1996.

[9] Pertijs MAP and Kindt WJ. A 140 dB-CMRR current-feedback instrumentation amplifier employing ping-pong auto-zeroing and chopping. *IEEE Journal of Solid State Circuits*, 45(10):2044–2056, 2010.

[10] Wu R, Makinwa KAA, and Huijsing JH. The design of a chopped current-feedback intrumentation amplifier. *IEEE ISCAS*, pages 2466–2469, 2008.

[11] Butti F, Bruschi P, and Piotto M. A chopper modulated low noise instrumentation amplifier for MEMS thermal sensors interfacing. In *Proceedings of Ph.D. Research in Microelectronics and Electronics (PRIME)*, pages 133–136, 2011.

[12] Burt R and Zhang J. A micropower chopper-stabilized operational amplifier using a SC notch filter with synchronous integration inside the continuous-time signal path. *IEEE Journal of Solid State Circuits*, 41(12):2729–2736, 2006.

[13] Luff GF. Chopper stabilized amplifier. (US7724080 B2), 2010. US Patent.

[14] Kusuda Y. Auto correction feedback for ripple suppression in a chopper amplifier. *IEEE Journal of Solid State Circuits*, 45(8):1436–1445, 2010.

[15] Wu R, Makinwa KAA, and Huijsing JH. A chopper current-feedback intrumentation amplifier with a 1 mHz 1/f noise corner and an AC -coupled ripple reduction loop. *IEEE Journal of Solid State Circuits*, 44(12):3232–3243, 2009.

[16] Witte JF, Huijsing JH, and Makinwa KAA. A current-feedback instrumentation amplifier with 5 uV offset for bidirectional high-side current-sensing. *IEEE ISSCC, digital technical papers*, pages 74–75, 2008.

[17] Wu R, Huijsing JH, and Makinwa KAA. A current-feedback instrumentation amplifier with a gain error reduction loop and 0.06% untrimmed gain error. *IEEE Journal of Solid State Circuits*, 46(12):2794–2806, 2011.

[18] Butti F, Bruschi P, and Piotto M. A method to increase the input resistance of current feedback chopper instrumentation amplifiers based on input port alternation. pages 1–10, 2015. unpublished data.

[19] Ferri G and Sansen W. A rail-to-rail constant-g_m low-voltage CMOS operational transconductance amplifier. *IEEE Journal of Solid State Circuits*, 32(10):1563–1567, 1997.

[20] Del Cesta F, Longhitano AN, Bruschi P, and Piotto M. A novel architecture for current-feedback instrumentation amplifiers with rail-to-rail input range. In *Proceedings of Ph.D. Research in Microelectronics and Electronics (PRIME)*, pages 1–4, 2014.

[21] Monticelli DM. Class ab output circuit with large swing. (US4570128 A), 1986. US Patent.