Università degli Studi di Pavia

Department of Electrical, Computer and Biomedical Engineering
Ph.D. School in Microelectronics - XXXIV cycle

Development of IP blocks in 110 nm CMOS technology for monolithic pixel sensors

Advisors:
Prof. Lodovico Ratti
Prof. Gianluca Traversi

Author:
Matteo Pezzoli

Academic year 2020–2021

Contents

Introduction **1**

1 ARCADIA **3**

1.1 ARCADIA's groudwork - SEED 5

1.2 Project targets 6

2 Technology characterization **9**

2.1 Static characterization 9

2.1.1 Test setup 10

2.1.2 Static characteristics measurements 13

2.1.2.1 Characteristic current and subthreshold slope factor . 18

2.1.2.2 Intrinsic gain 25

2.1.2.3 Threshold voltage 27

2.2 Noise characterization 32

2.2.1 Test setup and methodology 39

2.2.2 Noise measurement data 44

3 Bandgap **53**

3.1 Bandgap reference working principle 53

3.1.1 Complementary to Absolute Temperature component 53

3.1.2 Proportional to Absolute Temperature component 56

3.1.3 Bandgap reference: voltage mode 57

3.1.4 Bandgap reference: current mode 59

3.2 Bandgap reference design in a 110 nm CMOS technology 60

3.2.1 Radiation damage mechanisms in semiconductor devices 60

3.2.2 MOSFET behaviour in the sub-threshold region 62

3.2.3 Implemented bandgap reference scheme 63

3.2.3.1 Startup circuit 64

3.2.3.2 Operational amplifier 65

3.2.3.3 Programmable resistors 66

3.2.4 Bandgap simulations 66
3.2.5 Reference current from bandgap core 71
3.2.6 Non-linear compensation approach 75
3.2.6.1 Linear (classic) compensation approach 75
3.2.6.2 Non-linear compensation approach 78
3.2.6.3 Compensation points position as a minimization problem 81
3.2.6.4 Basic implementation of the non-linear compensation 82

4 Custom LVDS Transmitter and Receiver 85
4.1 Problems and performance evaluation of a line driver/receiver 86
4.1.1 Transmission line and termination mismatch effects 87
4.1.2 Bandwidth related effects 89
4.1.3 Noise and jitter effects 92
4.1.4 Eye diagrams 95
4.2 Custom LVDS transmitter design and simulations 95
4.2.1 Common Mode Feedback amplifier 97
4.2.2 3-bit control of the driver current 98
4.2.3 Phase splitter 99
4.2.4 Transmitter simulation 99
4.3 Custom LVDS receiver design and simulations 101
4.4 Custom LVDS loopback structure measurements 106
4.4.1 Measurements results 107

5 Conclusions 111

Bibliography 113

Introduction

The research activity carried out in this thesis work concerns the design of analog circuits for monolithic CMOS pixel detectors. It has been carried out in the scope of ARCADIA, an INFN funded project with the goal of developing a full-scale demonstrator of a fully depleted, monolithic, pixel detector using a conventional CMOS fabrication process in 110 nm technology, in collaboration with an industrial partner. The developed technology is being targeted at very diverse field of applications, such as high-energy physics experiments at future colliders, space applications and medical imaging. The project will deliver a large 512x512 pixel matrix with a 25 µm pixel pitch to demonstrate the scalability of the sensor technology that has been the focus of previous research. This thesis will be focused on the characterization of the 110 nm technology used and on the development of two analog circuits that will be implemented in the aforementioned pixel matrix: a radiation hard bandgap voltage reference and a custom sub-LVDS transmitter and receiver pair.

The first chapter of the thesis will briefly cover the ARCADIA project, with the goal of providing the needed context to explain the choices taken during the design phase. Following, the second chapter will report on the work carried out to characterize the 110 nm CMOS technology used. Static analysis such as I_D - V_{GS} and I_D - V_{DS} curves, alongside characteristic normalized current, subthreshold slope, threshold voltage and intrinsic gain are described in the chapter, followed by noise measurements and analysis on the white and flicker components. The third chapter will focus on the design and simulation of a radiation hard bandgap voltage reference and an ancillary block used to generate a reference current for the analog frontend. A theoretical analysis of trimming strategies is reported at the end of the third chapter, comparing the implemented, widespread one with different approaches. Lastly, the fourth chapter will show the design and simulation of a custom sub-LVDS transmitter and receiver pair targeting a maximum frequency of 2 Gbit s^{-1}, followed by the results of measurements carried out on a test chip implementing both blocks in a loopback configuration.

1

ARCADIA

The work presented in this thesis has been carried out in the framework of the ARCADIA project, organized and funded by the Istituto Nazionale di Fisica Nucleare (INFN). ARCADIA starts as a proposal to the Scientific Committee of the INFN group 5, focused on technological research, progressing the work laid out by a previous INFN project, SEED. Both these projects will be briefly presented in this chapter to provide the reader with the context required to explore the topics covered in the following chapters. ARCADIA - Advanced Readout CMOS Architectures with Depleted Integrated sensor Arrays aims to develop a novel CMOS sensor platform based, as the name suggests, on monolithic, fully depleted pixel sensors.

Monolithic CMOS sensors have grown increasingly attractive in the context of physics research[1], due to lower cost and material budget when compared to the more widespread hybrid pixel sensors, in which the sensor and readout electronics are kept separate, to be connected later via bump-bonding or flip-chip technologies. Moreover, a monolithic, fully depleted CMOS sensor seems to be quite a novel idea in the literature, which features examples of both monolithic[2] and fully depleted[3] detectors, but not both at the same time. In non-fully depleted sensors, the signal is collected inside a relatively thin, around 20 µm, epitaxial layer; the charges freed by the incoming radiation then move towards the collecting electrode by diffusion effects only. This characteristic makes the sensor unsuited for timing applications with very strict time-tagging requirements, due to the slow collection time related to the diffusion of the charges, or for the detection of high energy X-rays, since the thin sensitive layer does not allow for enough charge to be generated. This is not an issue in fully-depleted CMOS sensors, in which the charge is collected by drift due to the electric field present in the depleted region, thus allowing a faster charge collection, coupled with a large sensitive area. The solutions found in literature are various, but unsurprisingly not without issues; the first approach reported here is the use of SOI (Silicon On Insulator) wafers[4], featuring two different substrates: a thin one, hosting

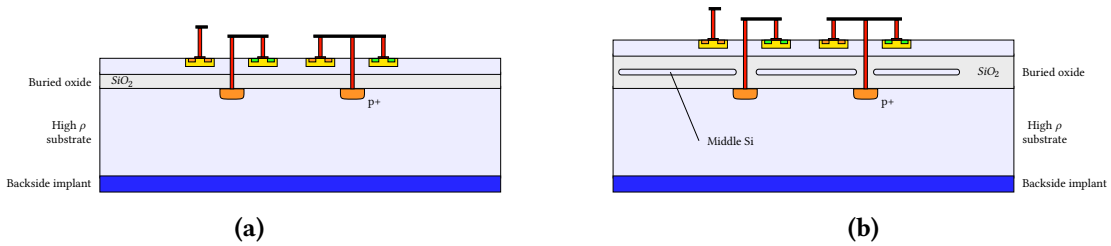


Figure 1.1 – Side-view of the monolithic sensor design based on SOI technology proposed by Arai et al. in [4] **(a)** and based on double-SOI technology proposed by Miyoshi et al. in [5] **(b)**. Drawing not to scale.

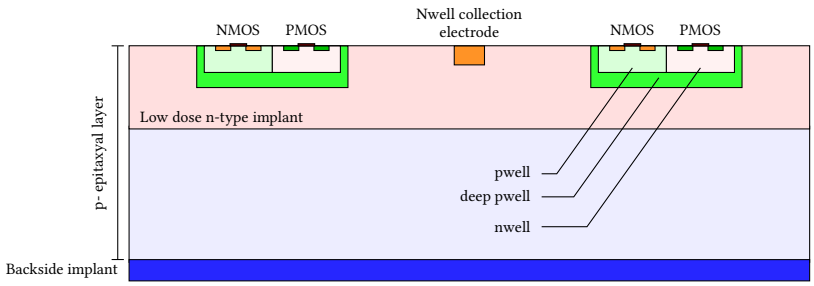


Figure 1.2 – Side-view of the monolithic sensor design for the ATLAS phase 2 upgrade proposed by Snoeys et al. in [6]. Drawing not to scale.

the front-end electronics, and a thick, high-resistivity one, acting as the sensor. A schematic view is reported in fig. 1.1a. The two substrates are separated by a buried oxide layer crossed only by the collecting electrode contacts. The problems linked with such an implementation are mainly two: in order to deplete the high resistivity sensor substrate, a high voltage is needed, which couples through the insulator changing the voltage of the front end substrate, ultimately changing the threshold of the front end devices. Solutions implementing a buried p-well shielding the front-end substrate from the high voltage are reported in the literature[4]. The second problem is linked to the insulator itself, which is implemented using silicon dioxide: this has the tendency to accumulate positive charges when subjected to radiation, affecting the performance of the front-end electronics and making the entire sensor very sensitive to TID (Total Ionizing Dose). Implementations counteracting this effect make use of a silicon layer buried between two insulator layers[5], allowing the biasing of the middle silicon layer in order to counterbalance the effects caused by the accumulated charge in the insulating oxide as shown in fig. 1.1b. These technologies are reported to work well, but they are the result of very specialized processes and thus quite expensive, due to both the additional process steps and the high cost of the SOI wafers. A second, more recent approach found in the literature has been proposed in the framework of the ATLAS pixel tracker upgrade[6] to ensure a charge collection fast enough to resolve hits from consecutive bunch crossings, separated by 25 ns, at the LHC. In the proposed solution,

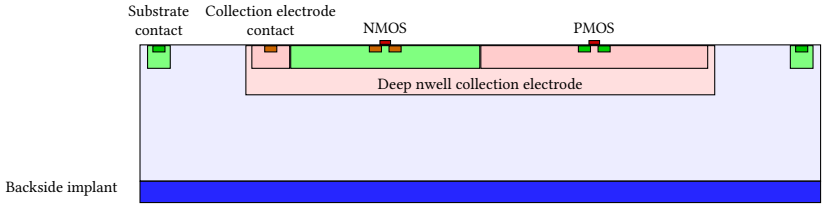


Figure 1.3 – Side-view of the monolithic sensor design proposed by Perić in [7]. Drawing not to scale.

an additional N- implant is added to allow the full depletion of the entire epitaxial layer, as shown in fig. 1.2. The epitaxial layer is reported to be between 15 µm and 30 µm, and its full depletion is reached by biasing the collecting electrode at 5 V and the substrate at −20 V, thus requiring the AC coupling of the electrode to the front-end electronics. The drawback of this solution lies in the limited thickness of the depletion zone, confined to the epitaxial layer. Lastly, a third approach found in the literature reports the use of front end-electronics embedded inside very large collecting electrodes, using deep n-wells[7], as shown in fig. 1.3. The use of a large collecting electrode allows for an uniform electric field in the substrate, granting better radiation hardness, and allows for the depletion of the bulk up to a thickness of 200 µm, but at the cost of a high parasitic capacitance, increasing the power consumption of the front-end electronics.

1.1 ARCADIA's groudwork - SEED

As mentioned, the work of the ARCADIA collaboration is the continuation of previous research in the scope of the SEED (Sensor with Embedded Electronics Development) project, an INFN-funded project carried out by a subset (INFN sections of Torino, Padova, Trento, Frascati and Perugia) of the current ARCADIA collaboration. The goal of the project was to research and develop the technology for a monolithic, fully depleted sensor alongside an industrial partner, LFoundry. The sensor was developed in a customized 110 nm CMOS technology, the same used in this thesis work, based on a high resistivity n substrate with a thickness of 300 µm. The p/n junction needed to deplete the sensor is formed using a p+ implantation manufactured with conventional backside processes, as seen in fig. 1.4. The front-end electronics are hosted inside deep p-wells to avoid competitive charge collection from PMOS n-wells[8]; a higher doping concentration layer, implemented as an n-type epitaxial layer, has been added to prevent the punch-through between the deep p-wells and the backside p+ contact when the sensor is fully depleted[9]. Guard rings have been added to increase the punch-through voltage and avoid side-current related effects; measures carried out on test structures indicate, for the maximum n-well bias voltage of 1.2 V, a full-depletion reached with a

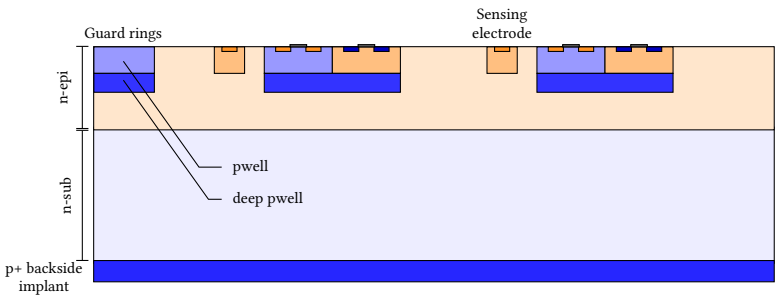


Figure 1.4 – Side-view of the SEED sensor. Drawing not to scale.

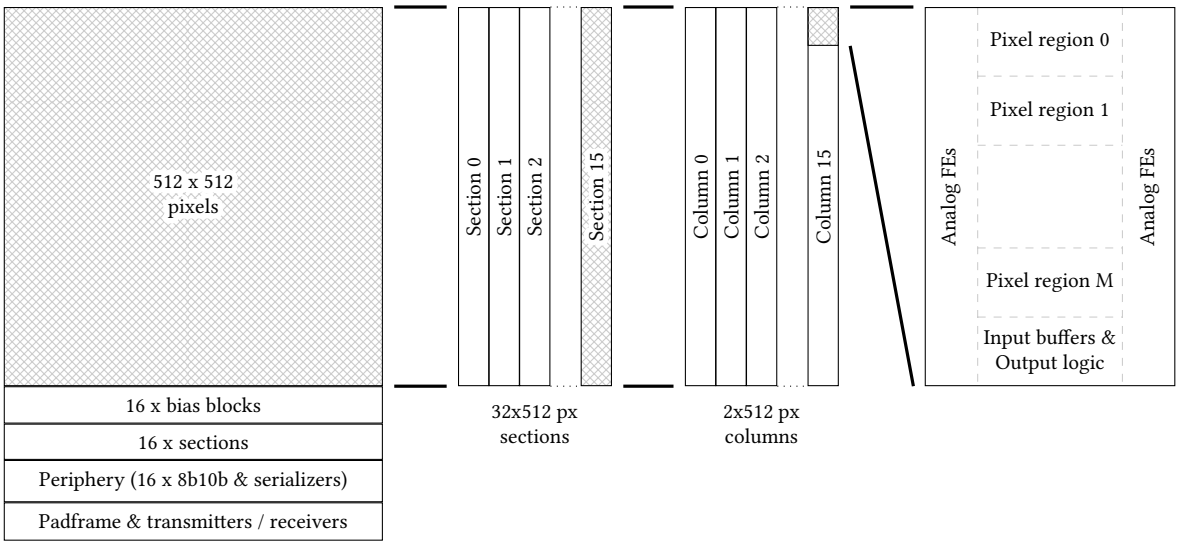


Figure 1.5 – Block diagram view of the ARCADIA Main Demonstrator 1.

backside bias of ∼−175 V and a punch-through at ∼−250 V. SEED has been tested using a small (2 mm×2 mm) ASIC, MATISSE (Monolithic AcTIve pixel SenSor Electronics)[10], proving the technology apt for X-ray imaging applications and promising for particle physics once the design has been ported to substrates with smaller thicknesses.

1.2 Project targets

ARCADIA aims to develop a large-scale demonstrator to progress the work carried out by SEED. To this end, the project deliverable is the Main Demonstrator 1 (MD1), a 512x512 pixel matrix with a pixel pitch of 25 µm; the matrix is divided into 16 identical sections, each one having dedicated bias, readout blocks and an I/O stage for data transmission up to 320 Mbit s^{-1}; the transmitter and receiver pair used in the data transmission will be discussed in chapter 4. A block-diagram showing the different blocks is reported in fig. 1.5. Each section, 32 pixels wide and 512 pixel tall, is further divided into 16 double columns which, for the sake of addressing, are then divided

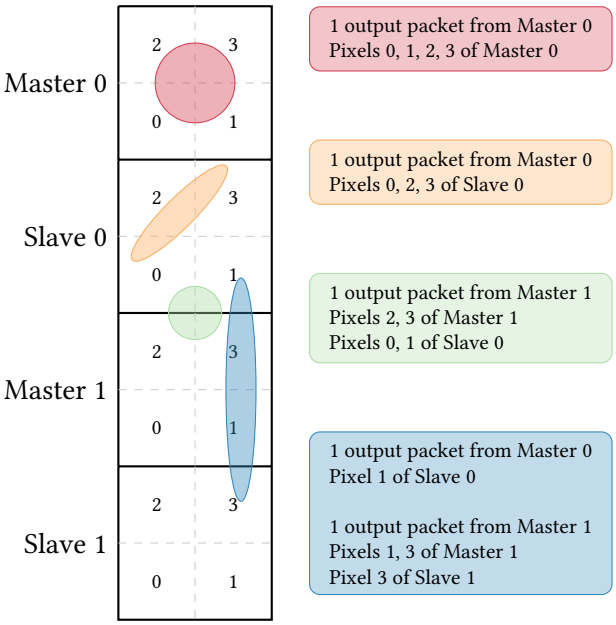


Figure 1.6 – Examples of the clustering scheme inside the ARCADIA MD1 matrix. Master pixels transmit packets when they have the transmission token, then wait for an acknowledge from the periphery logic.

again into 2x4 pixel cores, constituted by a 2x2 master and a 2x2 slave. Data packets are sent to the periphery by masters using a tokenized system: a token is passed down the double column from master to master, which will send data on pixel hits, collecting them from both adjacent slaves to minimize data sent towards the readout. Considering a master_0-slave_0-master_1-slave_1 2x8 pixel region as an example, a slave tries to aggregate its data to the corresponding master (e.g. master_0 for slave_0), but can opt to aggregate to the other master if the latter is hit, while its slave is not; fig. 1.6 shows some examples of the described clustering scheme. The data transmission from the periphery is done by the means of a serializer and a sub-LVDS driver. The data output is driven in two different modes: in the high-rate mode each of the sections will use their own sub-LVDS transmitter to output data after serialization; data is encoded according to a 8b10b scheme. In the slow-rate mode, also called space mode, all transmitters with the exception of the one in the first section are disabled, and all data packets are conveyed to the enabled transmitter; packets are encoded in a 32b40b scheme in this case. The space mode has been added, as the name suggests, to account for the very strict power consumption requirements of space experiments, by taking advantage of the usually fairly low particle rate in such applications[11, 12, 13, 14, 15].

2

Technology characterization

This chapter will explore the static and noise characterization of the 110 nm CMOS technology used in the ARCADIA project. Each section will explore the rationale and setup for the measurements, followed by the presentation of the experimental results and the relevant analysis.

2.1 Static characterization

A static characterization of the devices from a particular technology may seem, to an initial analysis, a bit redundant; the device models used in the PDK and included in the technology documentation are possibly a product of measurement campaigns conducted by the foundry on a number of devices orders of magnitude bigger than the one presented in this work, in addition to being carried out with an underlying level of understanding made available by the foundry technologist themselves. Nevertheless, it is the author's opinion that a new set of measurements is still going to provide useful information for the designs using this technology. Even without considering the verification of the device models, the interest in a new and somewhat limited characterization campaign can be twofold: for starters, the results from the static characterization of the devices are needed for their noise analysis, pursuing the goal of comparing the measurement results with theoretical models, since the small number of devices will permit the use of per-device data, as opposed to aggregated data used in larger characterization campaigns. As an additional asset, the measurements performed will lay the groundwork for comparison when analyzing the performance degrade during future irradiation campaigns; a set of complete, pre-irradiation data is mandatory to ensure the possibility of having relative degradation data, in addition to absolute values for the post-irradiation device characteristics.

2.1.1 Test setup

The measurements have been performed on a total of 30 different devices, reported in table 2.1. As it can be seen from the table, core devices are present for both N and

NMOS		
W [µm]	L [nm]	Notes
100	120	Core, $N_f = 40$
	240	Core, $N_f = 40$
	360	Core, $N_f = 40$
	720	Core, $N_f = 40$
200	120	Core, $N_f = 40$
	240	Core, $N_f = 40$
	360	Core, $N_f = 40$
	720	Core, $N_f = 40$
600	120	Core, $N_f = 40$
	240	Core, $N_f = 20$
		Core, $N_f = 40$
		Core, $N_f = 60$
		ELT1
		ELT3
	360	Core, $N_f = 40$
		IO
	720	Core, $N_f = 40$

PMOS		
W [µm]	L [nm]	Notes
100	120	Core, $N_f = 40$
	240	Core, $N_f = 40$
	360	Core, $N_f = 40$
	720	Core, $N_f = 40$
200	120	Core, $N_f = 40$
	240	Core, $N_f = 40$
	360	Core, $N_f = 40$
	720	Core, $N_f = 40$
600	120	Core, $N_f = 40$
	240	Core, $N_f = 40$
	360	Core, $N_f = 40$
		IO
	720	Core, $N_f = 40$

Table 2.1 – List of NMOS and PMOS devices available for the characterization. W and L refer to the width and length of the device, respectively. Notes indicate whether the device is a core, and ELT or a thick oxide (IO) one. N_f is the number of fingers in the device; the number of fingers for IO and ELT devices is 40.

P channels, in every combination of the 120 nm, 240 nm, 360 nm and 720 nm channel lengths with the 100 µm, 200 µm and 600 µm channel widths. Other devices have been added with the goal of studying the effect of ionizing radiation on a different number of fingers, on IO (thick oxide, 3.3 V) and Enclosed Layout Transistors (ELTs)[16]. An additional comment may be needed regarding the ELTs; alongside the technology scaling, there has been a need to prevent leakage currents from flowing from a device to an adjacent one. This has been achieved by isolating the channel using the so-called Shallow Trench Isolations (STI) but, although this has indeed obtained the effect of preventing unwanted currents flowing between devices, it has the tendency of accumulating positive charge when subjected to ionizing radiation [17, 18]. This charge build-up can lead to the development of a parasitic field: the polarity of the accumulated

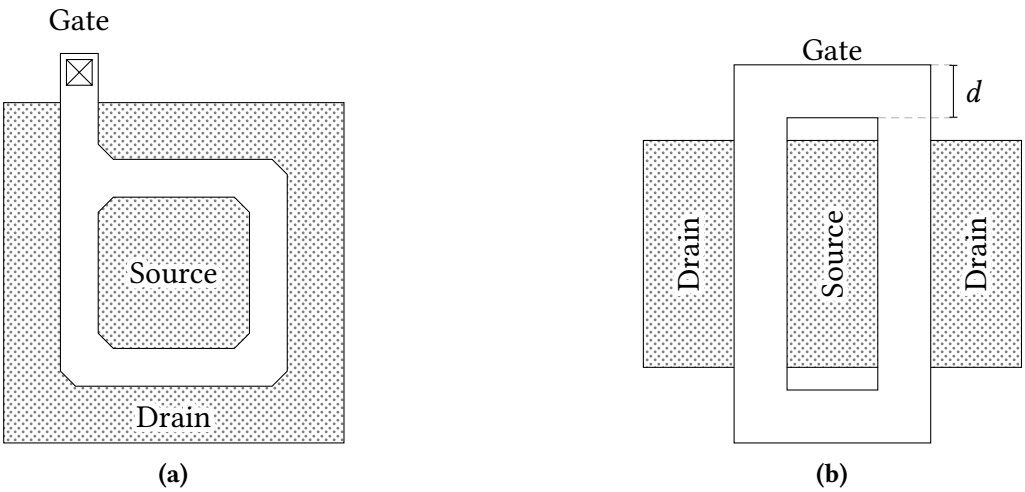


Figure 2.1 – Schematic representation of a classic ELT (anular) design **(a)** and a less area-demanding approach aimed at achieving the same effect **(b)**. The dimension marked by d is the extension of the gate poly which has been set to 1 µm and 3 µm respectively in two different version of ELTs implemented in this design.

charge poses an issue mainly in N-channel devices, where the parasitic field polarity is the same as the field generated by the gate potential when the channel is formed. This effect results in a lower threshold voltage in N-channel devices, or in a higher threshold voltage for P-channel ones; the effect of the STI-induced threshold variation is quite undesirable but, on the bright side, its effects are relevant only for narrow devices, becoming less impactful for large W transistors[19]. A second, more prominent effect caused by the charge accumulated in the STIs affects only N-channel devices: this concerns the creation of parasitic paths for the drain current, increasing the device leakage[20]. Furthermore, the same charge accumulation can be seen in the gate oxide, leading to similar results: the threshold initially gets lower for N-channel devices and higher for P-channel ones, a trend that is true up to relatively low values of total ionizing dose (around a few Mrad[21]). This effect is then outweighed at higher doses, for which the formation of traps near the oxide-silicon interface and the resulting decrease of the carrier mobility leads to an increase in the threshold for both NMOS and PMOS, effectively compensating for the variation due to lower doses in NMOS devices. It must be noted that the effects of the accumulated charge in the gate oxide are becoming less impactful due to the increasingly thin gate oxide in modern technologies, allowing the trapped charge to escape due to tunneling effects[22]. Enclosed Layout Transistors are designed to counteract this effect: the classic approach, shown in fig. 2.1a is to design the gate with an annular shape, surrounding the source (or drain) diffusion; the drain (or source) diffusion is then wrapped around the gate. This approach allows the STI to be quite far from the actual channel, mitigating the effect of the accumulated

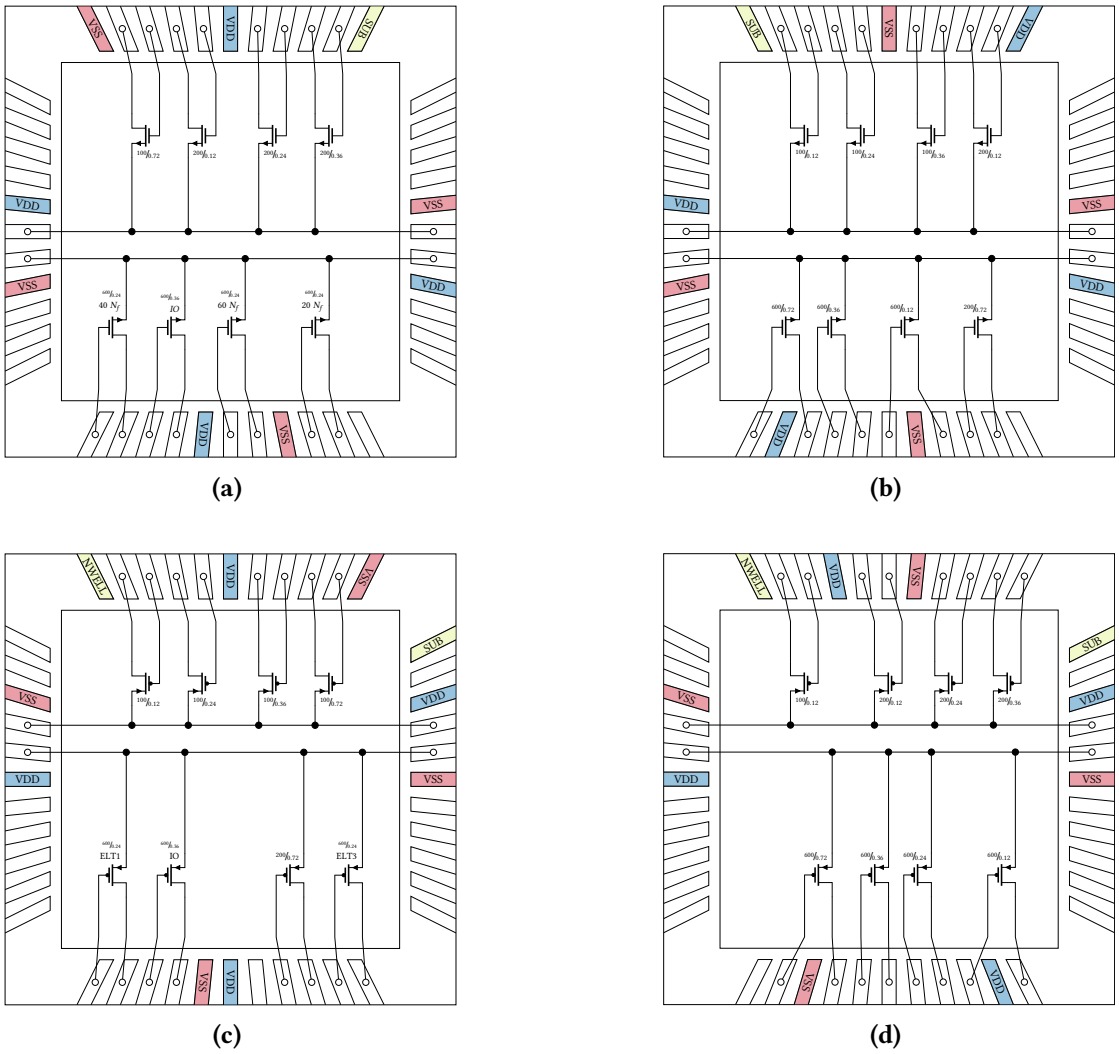


Figure 2.2 – Bonding diagrams for LCC44 packages containing single devices for the 110 nm technology characterization. **(a)** and **(b)** are bonding diagrams for NMOS devices, **(c)** and **(d)** for PMOS ones.

charge, but is usually quite demanding in terms of area. In order to ensure a similar effect without sacrificing much area, the devices can be designed using the approach schematically reported in fig. 2.1b: by stretching the gate poly farther, the STIs will be moved accordingly, in a position that is more distant from the channel. Each device has been measured on two different chips to confirm the measurement results. The ELT1 and ELT3 devices mentioned in table 2.1 refer to the stretching of the gate poly 1 µm and 3 µm beyond the actual gate area, respectively.

The **D**evices **U**nder **T**est (from now on referred to as DUTs) have been provided on naked dice, then bonded to a readily available LCC44 package. The use of the LCC package, coupled with the constraints on the bonding wires maximum angle of attack, has led to the need of two different bonding diagrams, represented schematically in fig. 2.2.

The presence of the IO devices (thick oxide, 3.3 V core devices) has forced the use of the corresponding clamping diodes in the padring, but this should be of no influence to the device characterization. As can be noted from the diagram reported in fig. 2.2, each device exposes the drain and gate, whereas the sources are shorted together to form two separate groups (top and bottom); the use of a common source connection leads to a smaller padring, but also to a higher than needed source series resistance for the devices more distant from the source pads, estimated to be around 2 Ω to 3 Ω[I]. This contribution can, and has been, reduced by connecting both available source contacts. A very simple board has been designed to host a LCC44 socket and 1 mm banana connectors, in order to attach the devices to the needed instrumentation which, in this case, was an Agilent Technology E5270B Parameter Analyzer; given the quasi-static nature of the measurements, no particular measures are required for the board design, apart from a minimization of the connection resistance. The instrument is composed of various Source-Measurement Units (or SMUs for short) synchronized and controlled by a single mainframe, which can be managed by the provided software running on a desktop PC connected to the instrument via its GPIB port.

2.1.2 Static characteristics measurements

Using the described setup, static device characteristics can be readily extracted; in particular, the measurements performed on the DUTs are as follows:

- I_D - V_{GS} curve, obtained by measuring the drain current while sweeping the gate to source voltage from 0 V to 1.2 V. This measurement has been repeated with different drain to source voltages, 0 V to 1.2 V (extremes included) with 200 mV steps, to create a family of I_D - V_{GS} curves.

- I_D - V_{DS} curve, obtained by measuring the drain current while sweeping the drain to source voltage from 0 V to 1.2 V. This measurement has been repeated with different gate to source voltages, 0 V to 1.2 V (extremes included) with 200 mV steps, to create a family of I_D - V_{DS} curves.

Curves showing some result from the aforementioned measurements, for the sake of example, are reported in figs. 2.3 to 2.7.

The data gathered from the I_D - V_{GS} and I_D - V_{DS} curves can then be used to extract the characteristic parameters for device modeling; first and foremost, the first order derivative of the previously mentioned curves can be calculated, resulting by definition

[I] Measured from the most distant device (i.e. the one in the center of the chip) to the left or right source connection.

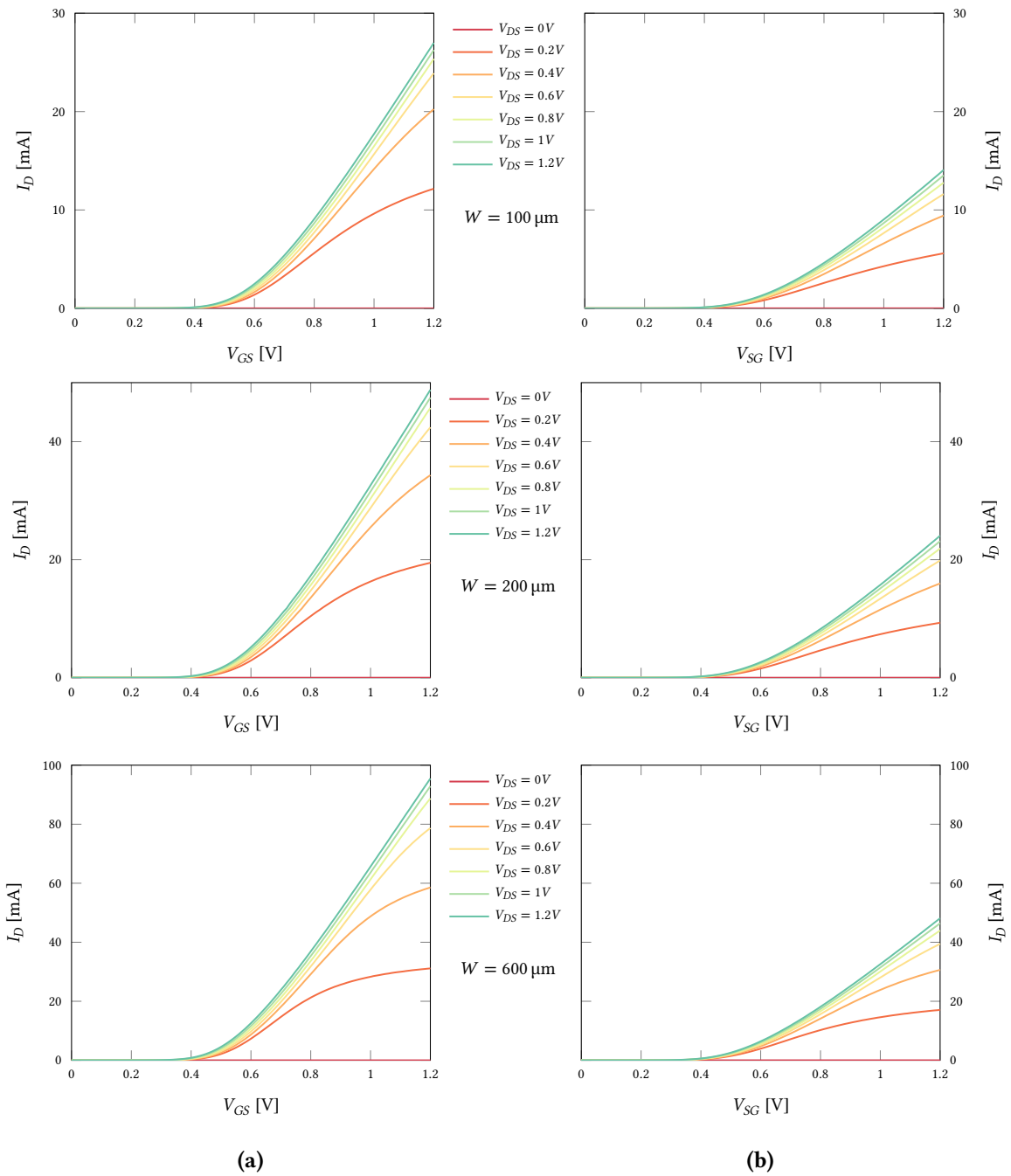


Figure 2.3 – I_D - V_{GS} curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =120 nm and channel widths (from the top down) of 100 µm, 200 µm and 600 µm.

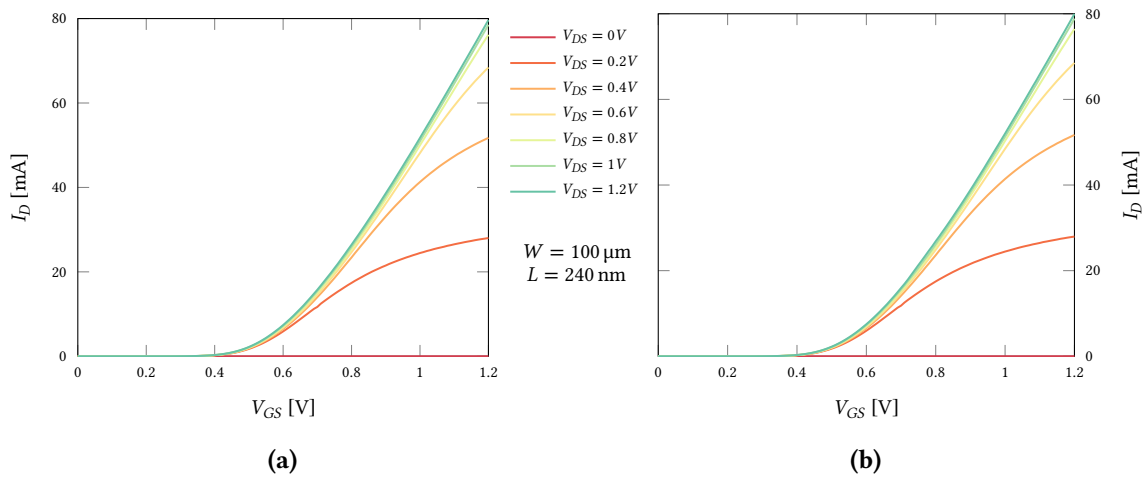


Figure 2.4 – I_D - V_{GS} curves for N channel devices with an enclosed layout design (ELT devices). These devices have a channel width of W =600 µm and a channel length of L =240 nm. **(a)** shows the curves for an ELT with d=1 µm and **(b)** shows the curves for d=3 µm.

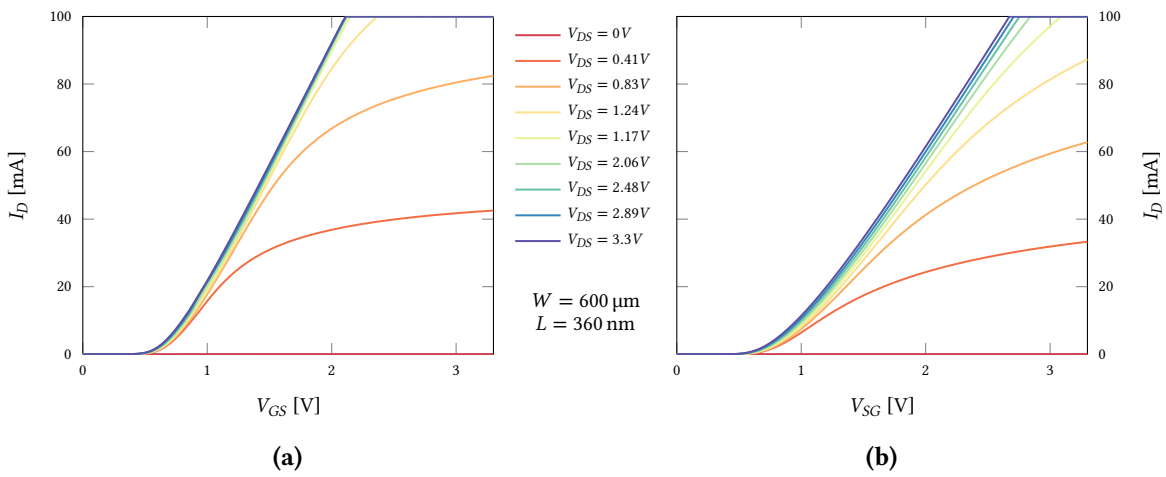


Figure 2.5 – I_D - V_{GS} curves for N channel thick oxide devices **(a)** and P channel thick oxide devices **(b)** with channel length of L =360 nm and channel width of 600 µm. Note that the current compliance used in these measurements, which is also the maximum compliance allowed by the semiconductor parameter analyzer, is 100 mA.

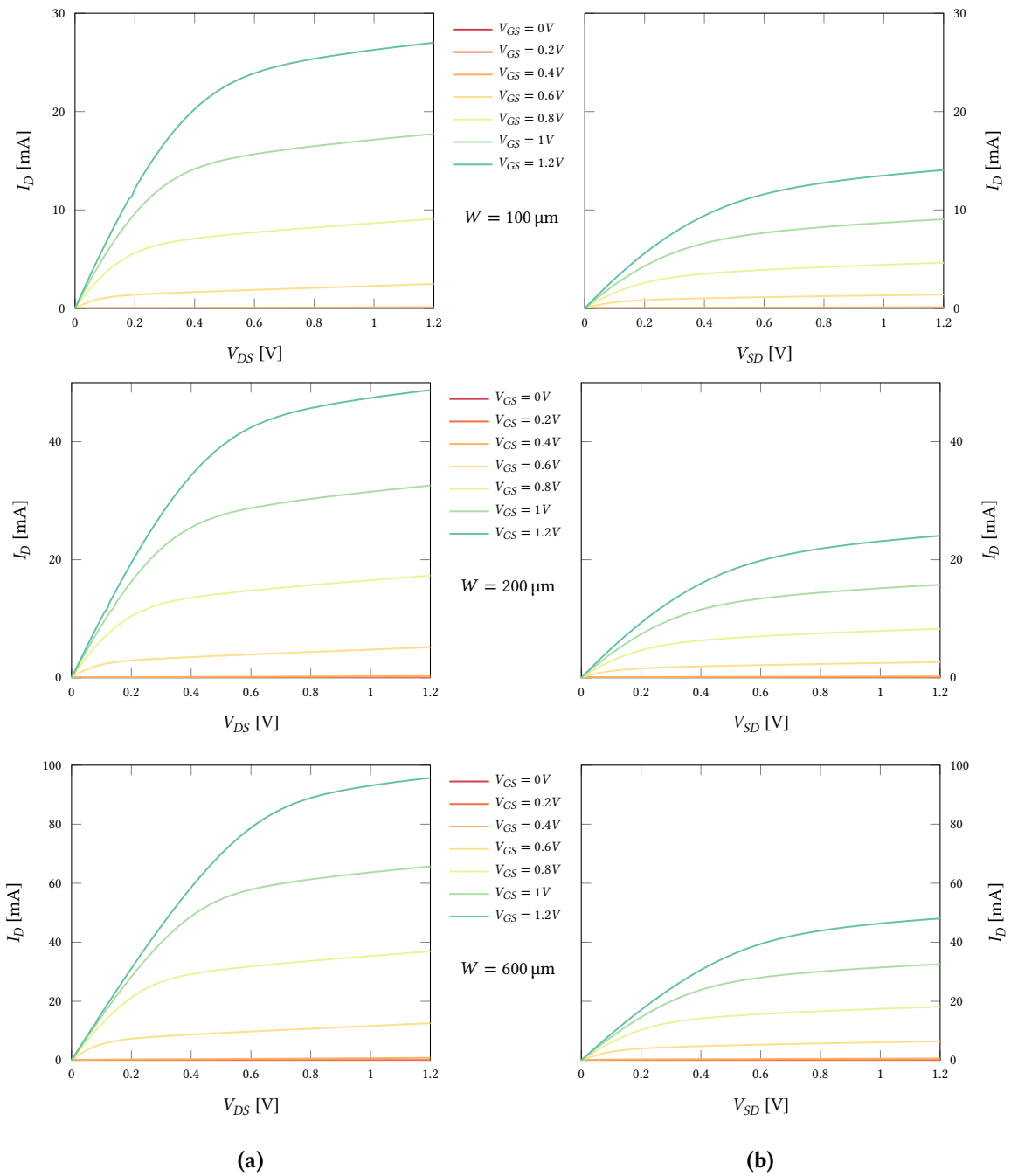


Figure 2.6 – I_D - V_{DS} curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =120 nm and channel widths (from the top down) of 100 µm, 200 µm and 600 µm.

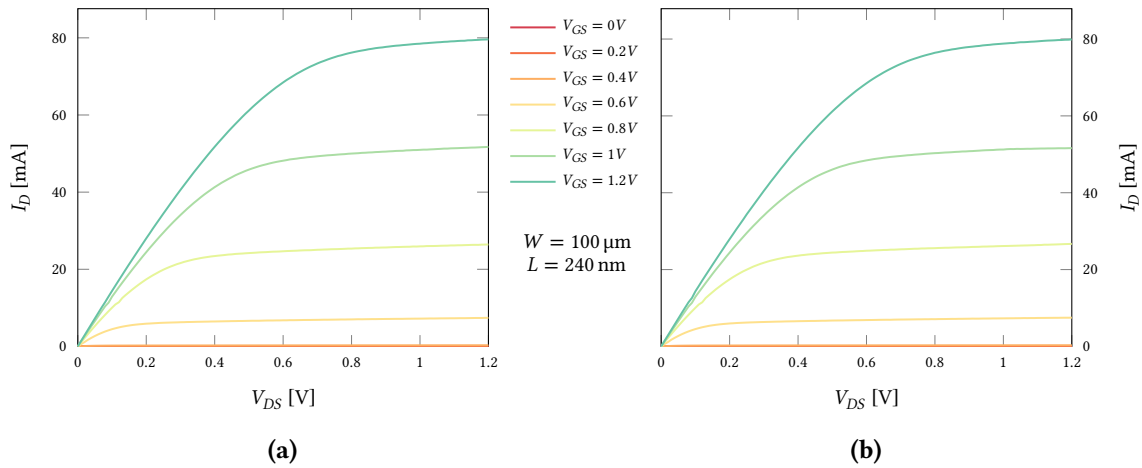


Figure 2.7 – I_D - V_{DS} curves for N channel devices with an enclosed layout design (ELT devices). These devices have a channel width of W =600 µm and a channel length of L =240 nm. **(a)** shows the curves for an ELT with d=1 µm and **(b)** shows the curves for d=3 µm.

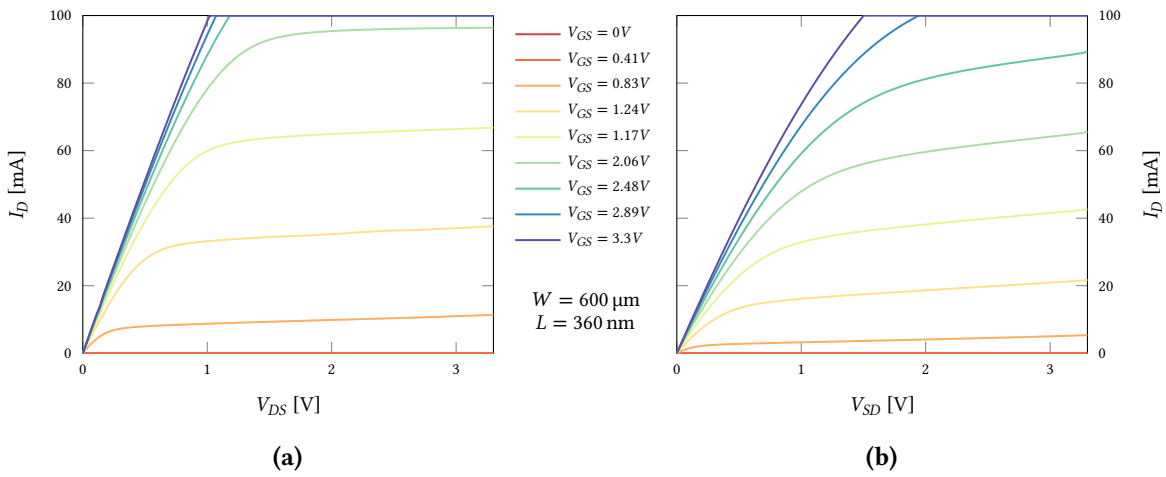


Figure 2.8 – I_D - V_{DS} curves for N channel thick oxide devices **(a)** and P channel thick oxide devices **(b)** with channel length of L =360 nm and channel width of 600 µm. Note that the current compliance used in these measurements, which is also the maximum compliance allowed by the semiconductor parameter analyzer, is 100 mA.

in the device transconductance g_m and the output conductance g_{ds}, reported in figs. 2.9 to 2.11 and figs. 2.12 to 2.14, respectively.

2.1.2.1 Characteristic current and subthreshold slope factor

Other, less directly obtainable information, can be extracted from the measurement data; one example of such information, useful for the design of analog circuits, is the indication of the devices region of operation as a function of its bias point.

A characteristic current I_Z can be defined as the drain current of the device in the middle of moderate inversion [23, 24]

$$I_Z = 2\mu \frac{W}{L} C_{ox} n_{sub} V_t^2 \tag{2.1}$$

where μ is the mobility of the majority carriers, C_{ox} the gate-oxide capacitance per unit area, V_t is the thermal voltage and n_{sub} the subthreshold slope coefficient [25]. The I_Z expression and a method of extracting it from the measured data can be deduced by considering that in weak inversion, the transconductance can be expressed as:

$$g_m\Big|_{\substack{Weak\\Inversion}} = \frac{I_D}{n_{sub} V_t} \tag{2.2}$$

From (2.2) the transconductance efficiency $\frac{g_m}{I_D}$ in weak inversion is

$$\frac{g_m}{I_D}\Bigg|_{\substack{Weak\\Inversion}} = \frac{1}{n_{sub} V_t} \tag{2.3}$$

meaning the transconductance efficiency is constant in weak inversion. Similarly, the transconductance and its efficiency can be expressed for a device in strong inversion as

$$g_m\Big|_{\substack{Strong\\Inversion}} = \sqrt{2\frac{\mu \frac{W}{L} C_{ox} I_D}{n_{sub}}} \tag{2.4}$$

$$\frac{g_m}{I_D}\Bigg|_{\substack{Strong\\Inversion}} = \sqrt{2\frac{\mu \frac{W}{L} C_{ox}}{n_{sub} I_D}} \tag{2.5}$$

meaning the transconductance efficiency is inversely proportional to the square root of the drain current. Equating (2.3) and (2.5) and solving for I_D leads to the characteristic current expression I_Z presented in (2.1). This shows quite clearly a method of extracting I_Z from the measurements: as can be seen from the example shown in fig. 2.15, plotting a $\frac{g_m}{I_D}$-I_D curve in a logarithmic scale shows a weak inversion (constant) asymptote and

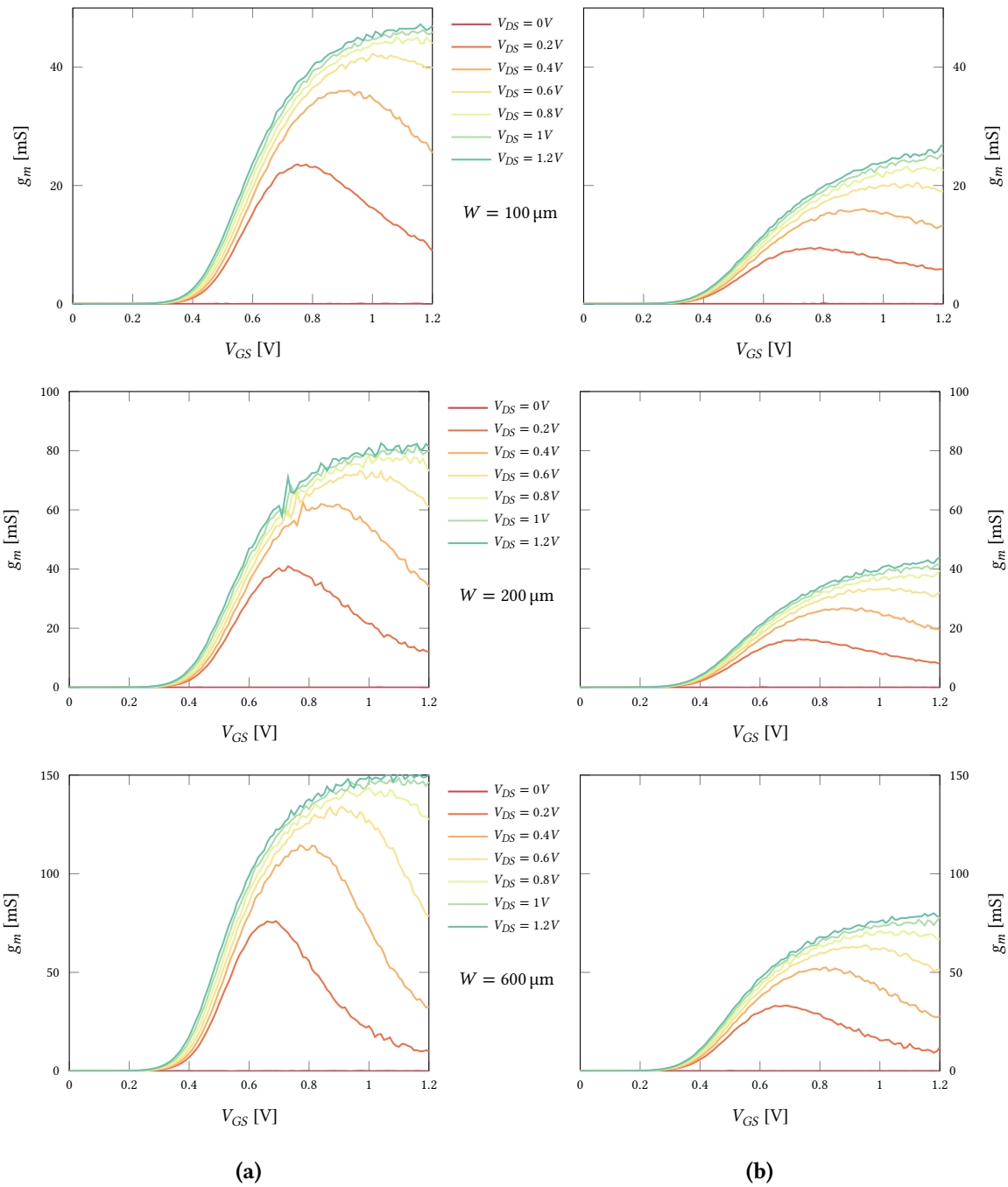


Figure 2.9 – g_m - V_{GS} curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =120 nm and channel widths (from the top down) of 100 µm, 200 µm and 600 µm.

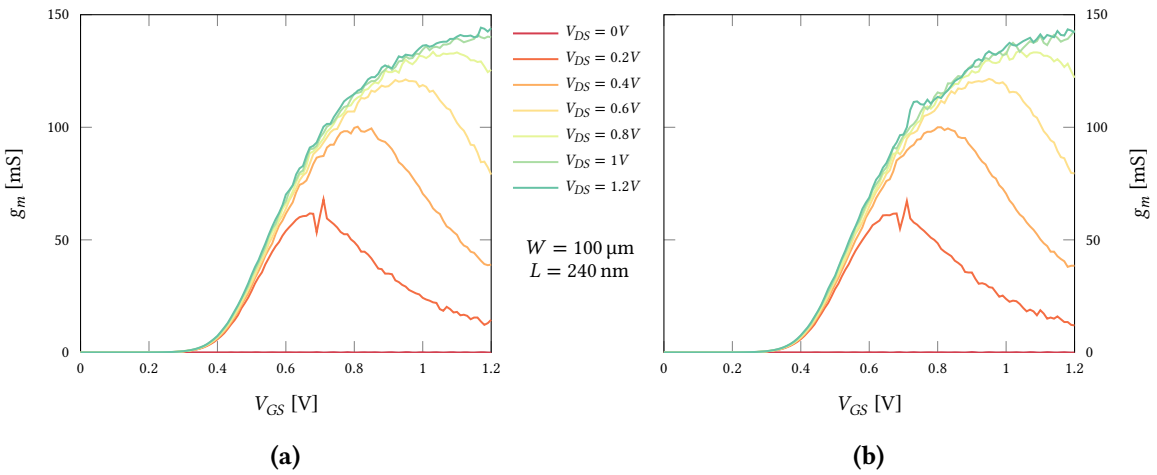


Figure 2.10 – g_m - V_{GS} curves for N channel devices with an enclosed layout design (ELT devices). These devices have a channel width of W =600 µm and a channel length of L =240 nm. **(a)** shows the curves for an ELT with d=1 µm and **(b)** shows the curves for d=3 µm.

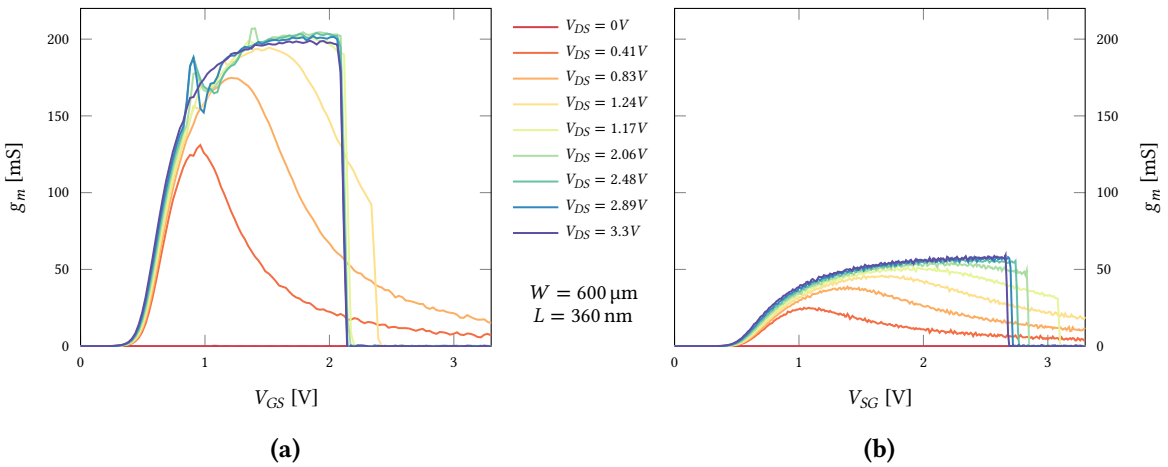


Figure 2.11 – g_m - V_{GS} curves for N channel thick oxide devices **(a)** and P channel thick oxide devices **(b)** with channel length of L =360 nm and channel width of 600 µm. Note that the current compliance used in these measurements, which is also the maximum compliance allowed by the semiconductor parameter analyzer, is 100 mA hence the transconductance immediately drops to zero once the compliance is reached.

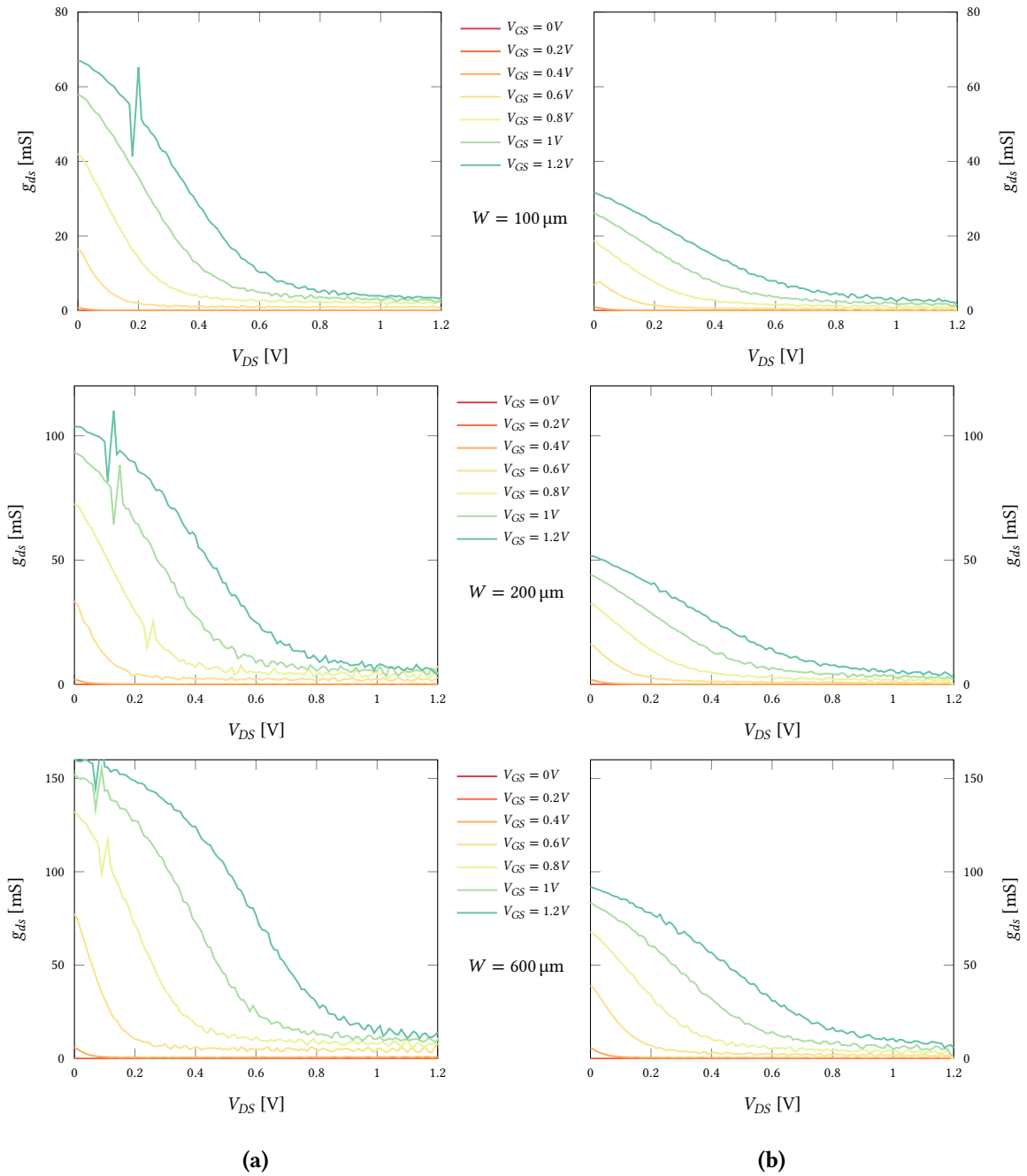


Figure 2.12 – g_{ds} - V_{GS} curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =120 nm and channel widths (from the top down) of 100 µm, 200 µm and 600 µm.

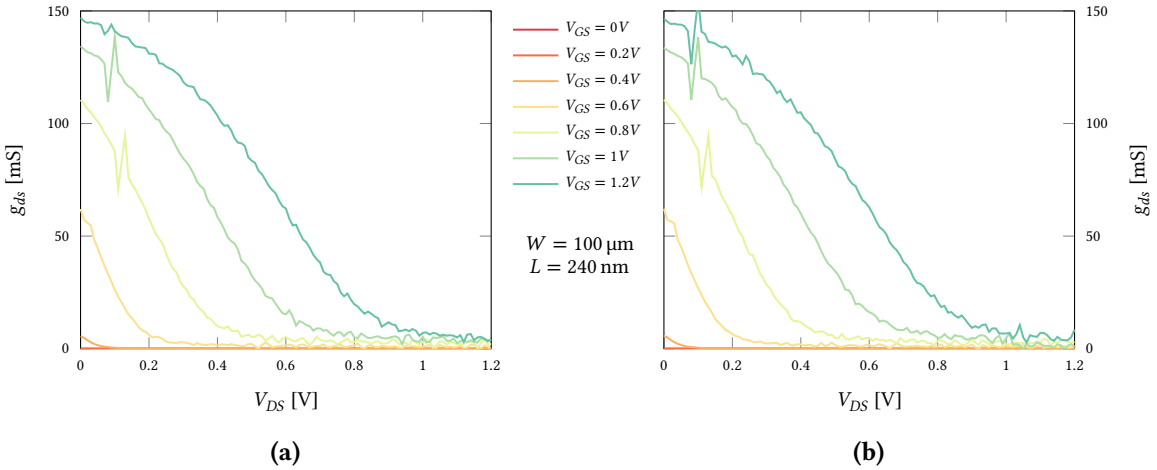


Figure 2.13 – g_{ds} - V_{GS} curves for N channel devices with an enclosed layout design (ELT devices). These devices have a channel width of W =600 µm and a channel length of L =240 nm. **(a)** shows the curves for an ELT with d=1 µm and **(b)** shows the curves for d=3 µm.

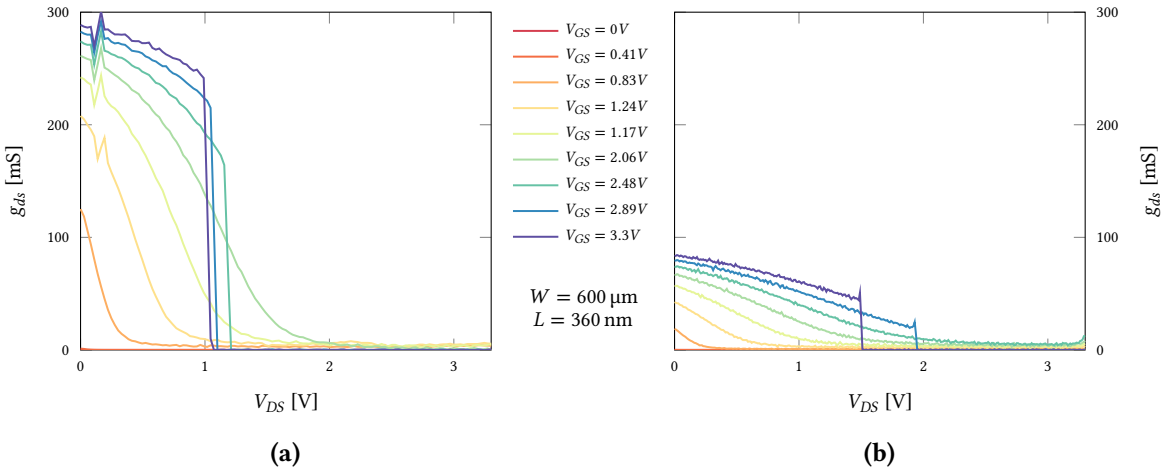


Figure 2.14 – g_{ds} - V_{DS} curves for N channel thick oxide devices **(a)** and P channel thick oxide devices **(b)** with channel length of L =360 nm and channel width of 600 µm. Note that the current compliance used in these measurements, which is also the maximum compliance allowed by the semiconductor parameter analyzer, is 100 mA hence the output conductance immediately drops to zero once the compliance is reached.

a strong inversion (inverted square law) one. If the low current data are interpolated according to the model in (2.3) (line with a first derivative equal to zero) and the high current data are interpolated according to the model in (2.5) (line with a first derivative equal to −½, given the log scale plot), then computing the intersection point of the two fits yields the value of the characteristic current I_Z. The data also shows an increase in the curve slope for higher drain currents, deviating from the initial −½ slope. This effect is given by the carrier velocity saturation: for low-intensity field, the carrier velocity increases linearly with the field. This is accounted for, in the MOSFET equations, by means of the low-field mobility μ. For higher intensity fields, the carrier velocity begins to saturate due to phonon scattering. When considering the carrier velocity saturation, the drain current loses its square-law dependence on V_{GS} [26]; this can be shown by adding a "degrading factor" to the low-field mobility

$$\mu_{vs} = \frac{\mu}{1 + \frac{V_{DS}}{L \cdot E_c}} \tag{2.6}$$

where E_c is the critical field, above which the velocity saturation occurs. The drain current becomes then simply

$$I_{D,vs} \simeq \frac{I_D}{1 + \frac{V_{DS}}{L \cdot E_c}} \tag{2.7}$$

The saturation current can be derived, as usual, by considering the equation in triode, then differentiating it to find the V_{DS} at which the current is at its maximum:

$$I_{D,vs}\Big|_{Triode} = \mu C_{ox} \frac{W}{L} \frac{1}{1 + \frac{V_{DS}}{L \cdot E_c}} \left[\left(V_{GS} - V_{th} \right) V_{DS} - \frac{V_{DS}^2}{2} \right] \tag{2.8}$$

By assuming that $L \cdot E_c \ll V_{DS}$ and ignoring the quadratic term, an approximation of the current in saturation can be obtained:

$$I_{D,sat} \simeq \mu C_{ox} W E_c \left(V_{GS} - V_{th} \right) \tag{2.9}$$

Equation (2.9) shows indeed a loss of the square-law dependence on V_{GS}, meaning that the transconductance loses its dependence on I_D, hence the transconductance efficiency slope in the $\frac{g_m}{I_D}$ - I_D plot becomes -1; this effect can be seen in the high current density region of fig. 2.15. As previously mentioned, I_Z is located, by definition, in the middle of the moderate inversion; a commonly accepted definition has the device enter weak inversion for currents one decade lower than I_Z, and strong inversion for I_D one decade higher than I_Z [27].

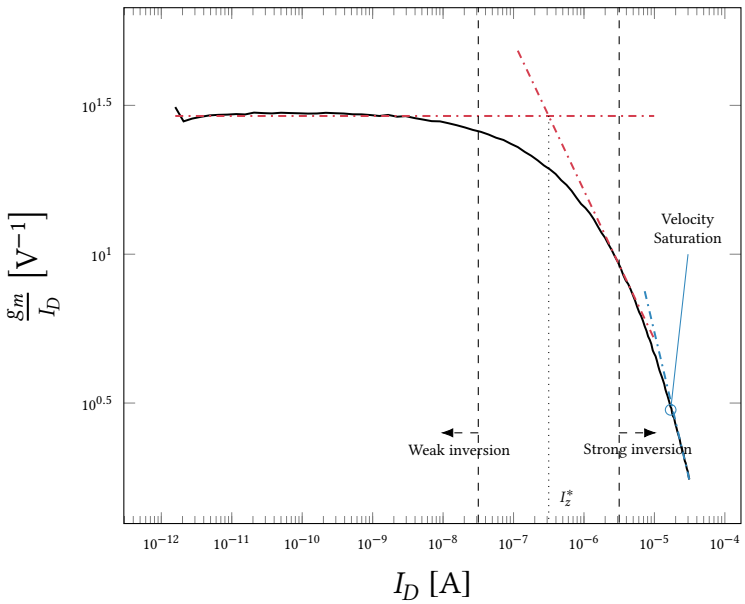


Figure 2.15 – Transconductance efficiency $\frac{g_m}{I_D}$ versus I_D plots for a NMOS device with W =100 µm and L =120 nm. The red linear fits (in logarithmic scale) show the regions of weak (left) and strong (right) inversion; the meeting point of the two fits falls exactly in the mid point of the moderate inversion, where the device has a drain current equal to I_Z. The blue linear fit in the high current region shows a slope of -1, indicating that carrier velocity saturation has occurred.

These definitions can be simplified into a single *Inversion coefficient*, I_{C0}, defined as:

$$I_{C0} = \frac{I_D}{I_Z} \tag{2.10}$$

leading to the weak inversion region being defined as the one where $I_{C0} \leq 0.1$ and the strong inversion region the one where $I_{C0} \geq 10$. The inversion coefficient is, by definition, equal to 1 in the middle of the moderate inversion. Table 2.2 reports the subthreshold slope factor n_{sub}, extracted from the transconductance data in weak inversion, as per (2.2), and the normalized characteristic current I_Z^*, for all the devices types in the characterization campaign. The normalized characteristic current is defined as

$$I_Z^* = I_Z \cdot \frac{L}{W} \tag{2.11}$$

and denotes the characteristic current for a unit area ($W = 1$ and $L = 1$) device, in order to ensure a meaningful comparison between differently sized devices. For all the devices but the thick oxide ones, the number reported is a mean value of the results extracted from that particular device type (core or ELT).

NMOS		
Type	I_Z^* [nA]	n_{sub}
Core	300	1.25
ELT	250	1.22
IO	250	0.56

PMOS		
Type	I_Z^* [nA]	n_{sub}
Core	110	1.3
IO	25	0.78

Table 2.2 – Values of the normalized characteristic current I_Z^* and the sub-threshold slope factor n_{sub}.

2.1.2.2 Intrinsic gain

Another useful figure in analog design, especially when dealing with gain stages, can be found in the intrinsic gain of a MOSFET; said parameter, A_{Vi}, is defined as the small-signal, gate-to-drain voltage gain when the device is biased by an ideal current generator:

$$A_{Vi} = g_m \cdot r_o = \frac{g_m}{g_{ds}} \tag{2.12}$$

Plotting A_{Vi} - I_{C0} curves allows meaningful comparison of different devices, since it grants the ability to correlate the intrinsic gain of two devices in the same region of operation. In weak inversion:

$$g_{ds} = \lambda_{wi} I_D \tag{2.13}$$

where λ_{wi} is the channel length modulation coefficient. Since, as seen in (2.2), g_m in weak inversion is independent of the channel length L and λ_{wi} in (2.13) is inversely proportional to the channel length L, the intrinsic gain in weak inversion is proportional to L; furthermore, it can be deduced from (2.2) and (2.13) that the intrinsic gain in weak inversion is independent of I_D. Regarding strong inversion, the output conductance can be expressed as

$$g_{ds} = \lambda_{si} I_D \tag{2.14}$$

where λ_{si} can be roughly approximated [28] as

$$\lambda_{si} \simeq \frac{10^7}{\sqrt{N} \cdot L} \tag{2.15}$$

with N the doping concentration. Given (2.4) for the transconductance in strong inversion, A_{Vi} in the same inversion region is proportional to $I_D^{-1/2}$. Combining the two results in a bi-linear approximation, the intrinsic gain as a function of I_{C0} can be shown to be constant (and increasing alongside L) during weak inversion and, in a log scale, linear with slope of $-1/2$; this behavior can be seen in the plots reported in 2.16.

The intrinsic gain is also a good indicator for the scaling impact on the technology:

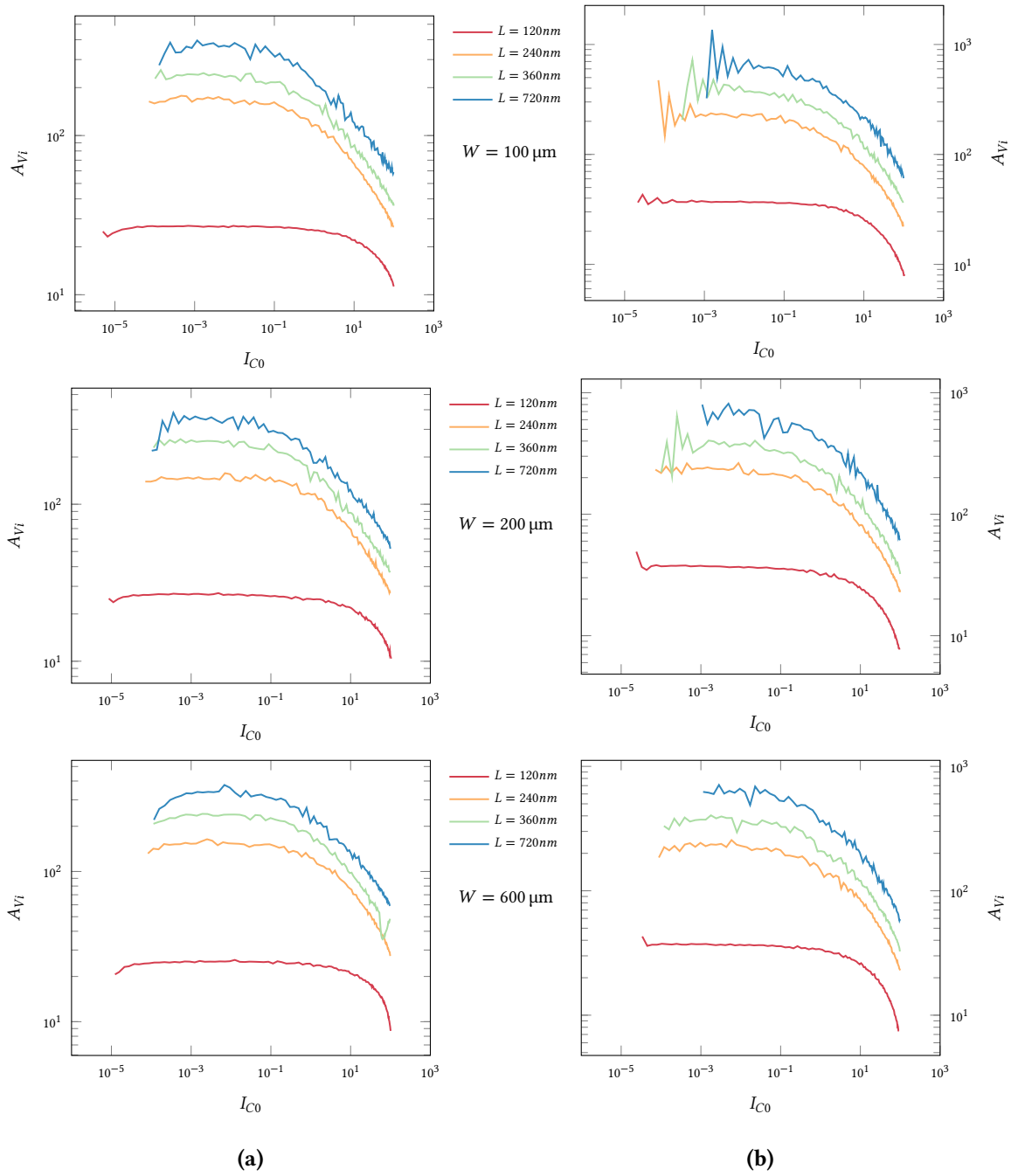


Figure 2.16 – Intrinsic gain A_{Vi} versus inversion coefficient I_{C0} plots, comparing different channel lengths L; plots **(a)** show N-channel devices, **(b)** show P-channel ones. All devices have been biased at the same drain-source voltage $V_{DS} = 1V$.

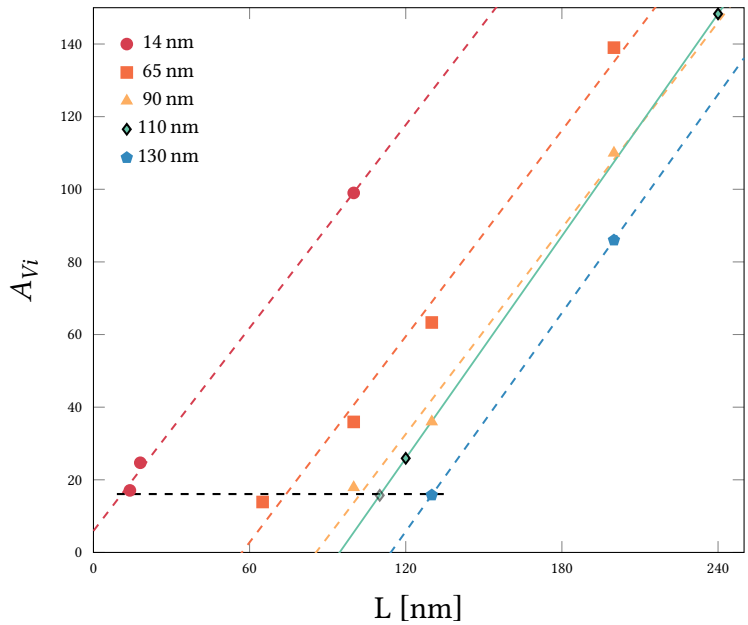


Figure 2.17 – Intrinsic gain A_{Vi} as a function of channel length, for different technologies. All data reported has been taken from devices biased on the onset of moderate inversion, with $I_{C0} = 0.1$. The solid line indicates the 110 nm technology analyzed in this work. The dashed black line indicates the mean value of A_{Vi} alongside the shortest channel devices in the analyzed technologies.

fig. 2.17 shows a comparison of different technology nodes, including the devices under test. By biasing every device to ensure the same inversion coefficient, in particular at $I_{C0} = 0.1$ where the device lies between the weak and moderate inversion and the A_{Vi} value is close to its horizontal asymptote, a meaningful comparison can be assured. It can be shown [29] that

$$A_{Vi} \propto \alpha \cdot L \tag{2.16}$$

with α the scaling factor, meaning that the A_{Vi} remains constant for the minimum gate lengths of different technologies, assuming a scaling factor of α on the process lithography.

2.1.2.3 Threshold voltage

The threshold voltage V_{th} is usually defined as the V_{GB} (gate-bulk voltage) needed to "invert" the channel, a situation in which the minority carriers presence is equal to the majority carriers one. The use of this definition for the extraction of the threshold voltage poses some difficulties, consequently other methods are used, mainly relying on the analysis of I_D - V_{GS} curves (or other curves constructed starting from them). The literature proposes a plethora of methods [30] meant to estimate V_{th}; this work will present a comparison of three different methods: the Ratio Method (RM), the Transconductance

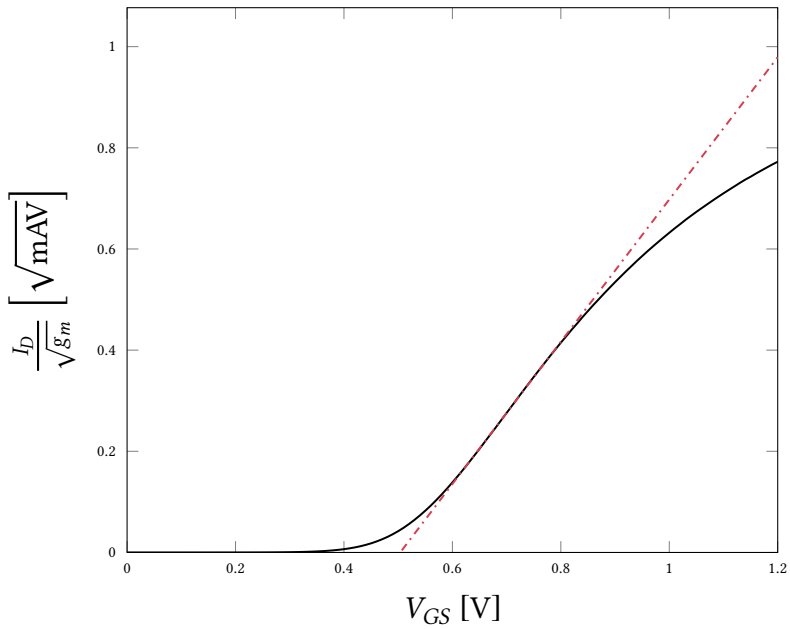


Figure 2.18 – $I_D/\sqrt{g_m}$ - V_{GS} curve used in the Ratio Method to extract the threshold voltage. The device has been biased with $V_{DS} = 10mV$; the x-axis intercept is the value of the estimated threshold voltage V_{th}. Results from a device with W =100 µm and L =120 nm have been shown here as an example.

Change Method (TCM) and the Second Derivative Logarithmic Method (SDLM). The Ratio Method has been proposed by Jain [31] and Ghibaudo [32] and is based on the assumption that the ratio of the drain current to the square root of the transconductance is linear with respect to the gate voltage, and the intercept onto the gate-voltage axis is in fact the threshold voltage, in particular:

$$\frac{I_D}{\sqrt{g_m}} = \frac{1}{\sqrt{\frac{L-\left(\Delta L_{eff}-\frac{\mu V_d}{v_{sat}}\right)}{W\mu C_{ox}V_d}}} \cdot \left(V_g - V_{th}\right) \tag{2.17}$$

where ΔL_{eff} is the decrease in the channel length due to pinch-off and v_{sat} the V_{DS} voltage at which carrier velocity saturation starts occurring. In this model, $V_g = V_{th}$ when $\frac{I_D}{\sqrt{g_m}} = 0$ or, as previously introduced, the threshold voltage is at the intercept onto the gate voltage axis. An example of this method is reported in fig. 2.18. The second method used in this work is the Transconductance Change Method, proposed by Wong et al. [33], the concept of which is to find the maximum point of the transconductance derivative with respect to V_{GS}, which also is the band-bending point ($\phi_s = 2\phi_F - V_{SB}$, where ϕ_s is the surface potential and ϕ_F the flat-band potential), meaning the point where the inversion layer carrier volume density is equal to the substrate doping volume density; as mentioned previously, this is the definition of threshold. An example of this

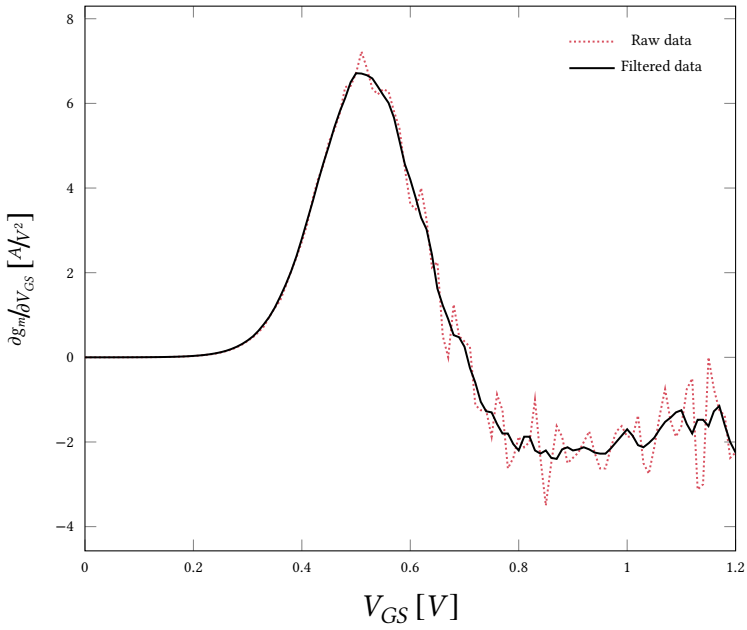


Figure 2.19 – $\partial g_m/\partial V_{GS}$ - V_{GS} curve used in the Transconductance Change Method to extract the threshold voltage. The device has been biased with $V_{DS} = 10mV$; the gate voltage at which the transconductance change is at its maximum is the value of the estimated threshold voltage V_{th}. Due to the derivative operation, errors are amplified; a set of filtered (moving average) data has therefore been used. Results from a device with W =100 µm and L =120 nm have been shown here as an example.

method is reported in fig. 2.19. Lastly, another method used to estimate the threshold voltage is the Second Derivative Logarithmic Method, proposed by Aoyama [34]. This method defines V_{th} as the voltage at which $\frac{\partial^2 \log I_D}{\partial V_{GS}^2} = 0$. The reason can be found in the equations modeling I_D in the various inversion levels of the channel: when in weak inversion, the drain current follows an exponential law

$$I_D = \mu C_{ox} \frac{W}{L} \left(n_{sub} V_t\right)^2 e^{\frac{V_G - V_{th} - n_{sub} V_t}{n_{sub} V_t}} \cdot \left(1 - e^{\frac{V_{DS}}{n_{sub} V_t}}\right) \tag{2.18}$$

while it follows a polynomial law in strong inversion, depending on the region of operation of the MOSFET:

$$I_D\Big|_{Linear} = \mu C_{ox} \frac{W}{L} \left[\left(V_{GS} - V_{th}\right) V_{DS} - \frac{V_{DS}^2}{2}\right] \tag{2.19}$$

$$I_D\Big|_{Saturation} = \frac{1}{2} \mu C_{ox} \frac{W}{L} \left(V_{GS} - V_{th}\right)^2 \tag{2.20}$$

The transition point between the exponential law and the polynomial one is the threshold, and this transition will be accompanied by a sharp change in the curve derivative,

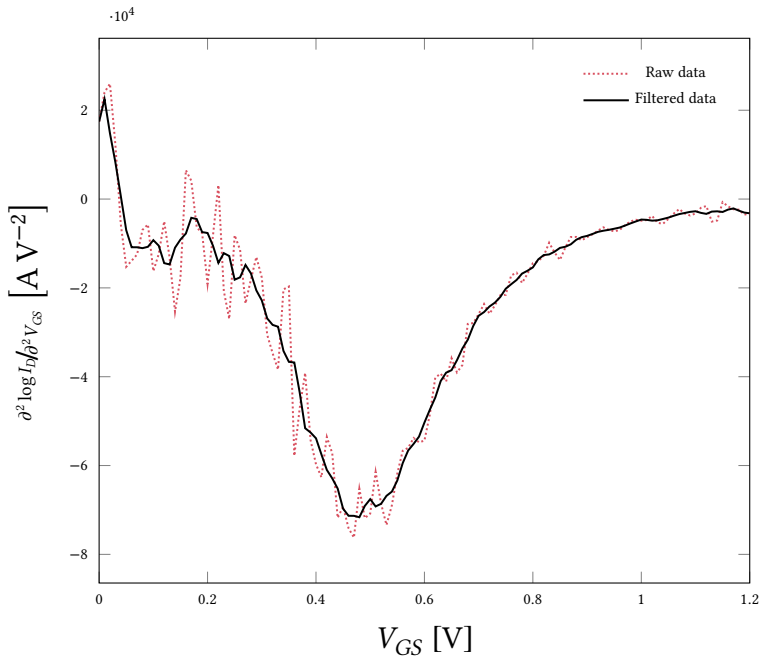
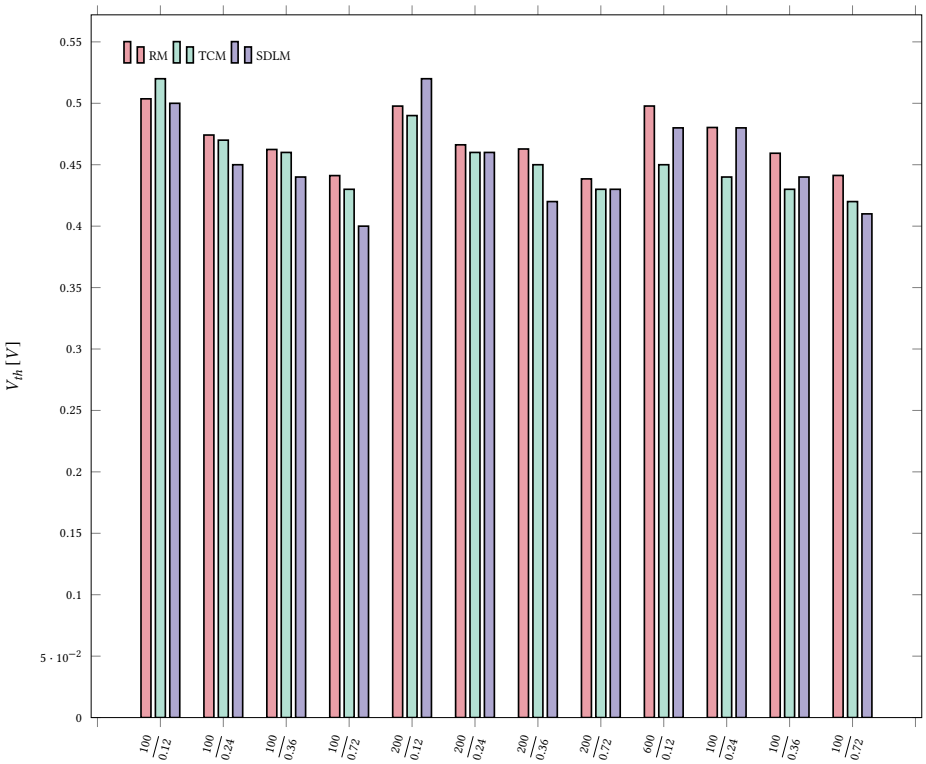


Figure 2.20 – $\partial^2 \log I_D / \partial^2 V_{GS}$ - V_{GS} curve used in the Second Derivative Logarithmic Method to extract the threshold voltage. The device has been biased with $V_{DS} = 10mV$; the gate voltage at which the second derivative of the drain current logarithm is at its minimum is the value of the estimated threshold voltage V_{th}. Due to the derivative operation, errors are amplified; a set of filtered (moving average) data has therefore been used.

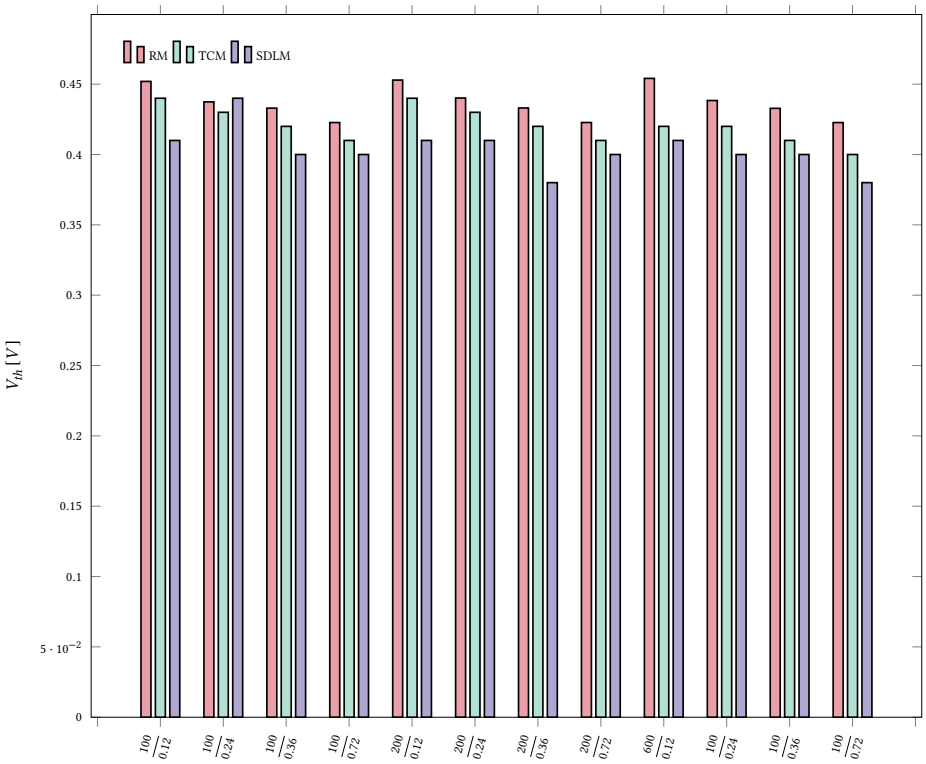
hence the method searches for the minimum point in the second derivative of the drain current, as shown in the example reported in fig. 2.20. A comparison between the highlighted methods is show in the histogram of fig. 2.21. Aggregated values summarizing the extracted thresholds are reported in table 2.3; as it can be seen, there is good agreement between the three methods, and each shows limited standard deviations, especially for PMOS devices.

NMOS			PMOS		
Method	Mean [V]	Std. dev. [mV]	Method	Mean [V]	Std. dev. [mV]
RM	0.469	22.7	RM	0.437	11.4
TCM	0.454	28.7	TCM	0.421	12.4
SDLM	0.453	36.7	SDLM	0.403	15.6

Table 2.3 – Summary of the extracted thresholds using three different methods: Ratio Method (RM), Transconductance Change Method (TCM) and Second Derivative Logarithmic Method (SDLM).



(a)



(b)

Figure 2.21 – Threshold extraction results for NMOS **(a)** and PMOS **(b)** devices and for each of the presented extraction methods (Ratio Method, Transconductance Change Method and Second Derivative of the drain current Logarithm).

2.2 Noise characterization

Noise in semiconductor devices is a stochastic process generated by physical phenomenons in the semiconductor structure, such as the thermal excitation of the charge carriers or the trapping and de-trapping in defects present at the interface region between different materials. Noise, as opposed to environmental disturbance, cannot be removed by design precautions such as shielding or filtering, thus it is of paramount importance to know the noise characteristics of a microelectronic technology for a noise-aware design of analog circuits in such a technology. This section will introduce various representations of noise and its sources, and subsequently report on the noise characterization performed during this thesis work. As a first intuition, one may suppose to use the temporal mean as a feature to represent the features of a noise process; due to the ergodicity of the process, the temporal mean would indeed be a meaningful description, but a simple mean would probably mask the signal fluctuations. As a basic example, a noise signal could show huge variations centered around 0, while featuring a null mean. The mean square value of a signal is therefore used as an indicator of how noisy and fluctuating the signal is. The same amount of information is provided by the root mean square of the signal, which is defined as the square root of the mean square value:

$$v_{N,rms} = \sqrt{\lim_{T\to\infty} \frac{1}{2\cdot T} \cdot \int_{-T}^{T} v_N(t)^2 \, dt} = \sqrt{\overline{v_N(t)^2}} \tag{2.21}$$

The value of $v_N(t)$ is unknown due to its randomness, but it is possible to deduce the value of $v_N^2(t)$ thanks to Fourier analysis, in case the power spectral density of the noise is known. Given a time T, consider a truncated version of the signal $v_N(t)$ as:

$$v_{N,T}(t) = \begin{cases} v_N(t) & |t| \leq T \\ 0 & |t| > T \end{cases} \tag{2.22}$$

It can be noted that, for the truncated function $v_{N,T}(t)$

$$\overline{v_{N,T}(t)^2} = \lim_{L\to\infty} \frac{1}{2\cdot L} \cdot \int_{-L}^{+L} v_{N,T}(t)^2 \, dt \tag{2.23}$$

$$= \frac{1}{2\cdot T} \cdot \int_{-T}^{+T} v_{N,T}(t)^2 \, dt \tag{2.24}$$

since $v_{N,T}(t)$ is zero for $|t| > T$. Furthermore,

$$v_N(t) = \lim_{T\to\infty} v_{N,T}(t) \tag{2.25}$$

Now, considering the quadratic mean of the truncated function in the $[-T, T]$ interval:

$$\overline{v_{N,T}(t)^2} = \frac{1}{2\cdot T}\cdot\int_{-T}^{+T} v_{N,T}(t)^2\ dt = \tag{2.26}$$

$$= \frac{1}{2\cdot T}\cdot\int_{-T}^{+T} v_{N,T}(t)\cdot v_{N,T}(t)\ dt = \tag{2.27}$$

$$= \frac{1}{2\cdot T}\cdot\int_{-T}^{+T} v_{N,T}(t)\cdot\left[\int_{-\infty}^{+\infty} V_{N,T}\left(f\right)\cdot e^{j\cdot 2\cdot\pi\cdot f\cdot t}\ df\right] dt \tag{2.28}$$

where $V_{N,T}\left(f\right)$ is the Fourier transform of $v_{N,T}$. Inverting the order of integration:

$$\overline{v_{N,T}(t)^2} = \frac{1}{2\cdot T}\cdot\int_{-\infty}^{+\infty} V_{N,T}\left(f\right)\cdot\left[\int_{-T}^{+T} v_{N,T}(t)\cdot e^{j\cdot 2\cdot\pi\cdot f\cdot t}\ dt\right] df \tag{2.29}$$

The second integral does not change if the integration extremes go to $\pm\infty$, due to the truncated function value being 0 for $|t| > T$, so

$$\overline{v_{N,T}(t)^2} = \frac{1}{2\cdot T}\cdot\int_{-\infty}^{+\infty} V_{N,T}\left(f\right)\cdot V^*_{N,T}\left(f\right)\ df \tag{2.30}$$

$$= \frac{1}{2\cdot T}\cdot\int_{-\infty}^{+\infty} \left|V_{N,T}\left(f\right)\right|^2\ df \tag{2.31}$$

where $V^*_{N,T}\left(f\right)$ is the complex conjugate of the Fourier transform of $v_{N,T}$. Now, due to (2.25):

$$\overline{v_N(t)^2} = \lim_{T\to\infty}\frac{1}{2\cdot T}\cdot\int_{-\infty}^{+\infty} \left|V_{N,T}\left(f\right)\right|^2\ df \tag{2.32}$$

$$= \int_{-\infty}^{+\infty}\lim_{T\to\infty}\frac{\left|V_{N,T}\left(f\right)\right|^2}{2\cdot T}\ df \tag{2.33}$$

assuming that the limit exists. By defining the bilateral power spectral density as

$$N_V\left(f\right) = \lim_{T\to\infty} \frac{\left|V_{N,T}\left(f\right)\right|^2}{2\cdot T} \tag{2.34}$$

the relation between the quadratic mean and the bilateral power spectral density can be written as:

$$\overline{v_N\left(t\right)^2} = \int_{-\infty}^{+\infty} N_V\left(f\right)\ df \tag{2.35}$$

Representing the noise as a power spectral density is useful to show its dependency on frequency, highlighting the contributions from different noise sources. The result shown in (2.35) describes the relationship between the quadratic mean and the bilateral power spectral density $N_V\left(f\right)$, but has little practical significance, since the integral is calculated over positive and negative frequencies. In order to have a more useful result, it can be considered that, since the bilateral power spectral density is an even function, one can define a unilateral power spectral density $S_V\left(f\right)$ as

$$S_V\left(f\right) = 2\cdot N_V\left(f\right) = \lim_{T\to\infty} \frac{\left|V_{N,T}\left(f\right)\right|^2}{T} \tag{2.36}$$

hence

$$\overline{v_N\left(t\right)^2} = \int_{0}^{\infty} S_V\left(f\right)\ df \tag{2.37}$$

expressing the quadratic mean as a function of the unilateral spectral power density, which can actually be measured by means of a spectrum analyzer and therefore will be shown further down in this section. The RMS value and, more generally, the quadratic mean of a noise signal can provide a meaningful figure of the intensity of the noise itself, but no information about is frequency components. For this purpose, the autocorrelation function can be used:

$$R_N\left(\tau\right) = \overline{v_N\left(t\right)\cdot v_N\left(t+\tau\right)} = \lim_{T\to\infty} \frac{1}{2\cdot T}\cdot \int_{-T}^{T} v_N\left(t\right)\cdot v_N\left(t+\tau\right)\ dt \tag{2.38}$$

The autocorrelation function is a convolution of the signal with itself; this function shows high values for highly correlated signals (e.g. a constant) and can show negative values if the signal changes sign. The autocorrelation function can be related to the

bilateral power spectral density following steps similar to the ones used to deduce (2.35):

$$R_N(\tau) = \lim_{T\to+\infty} \frac{1}{2\cdot T} \cdot \int_{-T}^{+T} v_{N,T}(t) \cdot \left[\int_{-\infty}^{+\infty} V_{N,T}(f) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot(t+\tau)}\, df \right] dt \tag{2.39}$$

$$= \lim_{T\to+\infty} \frac{1}{2\cdot T} \cdot \int_{-\infty}^{+\infty} V_{N,T}(f) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot\tau} \cdot \left[\int_{-T}^{+T} v_{N,T}(t) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot t}\, dt \right] df \tag{2.40}$$

$$= \lim_{T\to+\infty} \frac{1}{2\cdot T} \cdot \int_{-\infty}^{+\infty} V_{N,T}(f) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot\tau} \cdot \left[\int_{-\infty}^{+\infty} v_{N,T}(t) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot t}\, dt \right] df \tag{2.41}$$

$$= \lim_{T\to+\infty} \frac{1}{2\cdot T} \cdot \int_{-\infty}^{+\infty} V_{N,T}(f) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot\tau} \cdot V^*_{N,T}(f)\, df = \tag{2.42}$$

$$= \int_{-\infty}^{+\infty} \underbrace{\lim_{T\to+\infty} \frac{1}{2\cdot T} \cdot \left| V_{N,T}(f) \right|^2}_{N_V(f)} \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot\tau}\, df = \tag{2.43}$$

$$= \int_{-\infty}^{+\infty} N_V(f) \cdot e^{j\cdot 2\cdot\pi\cdot f\cdot\tau}\, df \tag{2.44}$$

Different types of noise sources may be present in a single device: the first taken into consideration will be the sources of thermal noise. Thermal noise is, in resistors such as wound wire resistors or thick/thin film resistors, generated by the random motion of electrons due to thermal excitation; independently of the actual current flowing into the resistor, the unilateral power spectral density is proportional to the absolute temperature T and the resistance value R:

$$\frac{d\overline{e_R^2}}{df} = 4KTR \tag{2.45}$$

where K is the Boltzmann constant. The power spectral density in (2.45) is based on a Thevenin model, but a Norton equivalent can be used as well, with a power spectral density:

$$\frac{d\overline{i_R^2}}{df} = \frac{4KT}{R} \tag{2.46}$$

In a MOSFET device, the sources of thermal noise are mainly two: the channel thermal noise and the gate series resistance contribution. Regarding the channel thermal noise, given Nyquist's [35] interpretation, it can be shown that, since thermal noise is due to temperature-induced agitation of carriers, it is present in any conductor, MOSFET

channel included. The channel thermal noise voltage expression is [36]:

$$S_w\left(f\right) = 4kT\Gamma\frac{1}{g_m} \tag{2.47}$$

Note that in the equation presented by Tsividis in the reference, $\Gamma = 2/3$, which is true for long channel devices; actually, Γ varies with the MOSFET operating region, from a minimum value of $1/2$ to a maximum value of $2/3$. An expression for Γ as a function of the inversion coefficient I_{C0}, previously discussed in section 2.1.2.1, is [37]:

$$\Gamma\left(I_{C0}\right) = \frac{1}{1+I_{C0}}\left[\frac{1}{2}+\frac{2}{3}I_{C0}\right] \tag{2.48}$$

Equation (2.47) shows that, as the subscript w suggests, the noise is white, meaning that it has a constant power spectral density, independent of frequency; in addition, due to the transconductance g_m at the denominator, the channel noise scales with the square root of the drain current, at least for devices in strong inversion. Other ohmic sections of the device can contribute in terms of thermal noise; the materials constituting the source, drain and gate themselves present a small resistance adding its thermal noise contribution. Usually the source and drain resistances are negligible for wide transistors, at least with respect to the gate one, which can be quite large due to the higher resistivity of the gate material. Figure 2.22 shows a schematic representation of the device and its distributed resistances contributing to thermal noise. As it can be seen, some layout measures, namely the use of multiple gate contacts and fingers, can drive the contribution down significantly. A different type of noise in a MOSFET device appears in the form of a flicker, or also known as $1/f$, term. The $1/f$ noise is caused by a variation in the carrier density due to fluctuations in the surface potential; carriers are trapped and later released near the oxide-channel interface [38], and this has a more prominent effect at lower frequencies. The expression of $1/f$ noise voltage is:

$$S_{1/f}(f) = \frac{K_f}{C_{ox}^2 WL}\frac{1}{f^{\alpha}} \tag{2.49}$$

where K_f is a technology dependent parameter, C_{ox} is the gate oxide capacitance per unit area and α is a coefficient regulating the frequency dependence (ideally equal to 1). The equation shows that the power spectral density is inversely proportional to the gate area, which can be intuitively explained by considering that a larger area leads to more occurrences of trapping/releasing of carriers, which tend to average out, resulting in a smaller flicker noise. A different interpretation found in literature explains $1/f$ noise with the mobility fluctuation given by phonon scattering in the bulk portion just below the

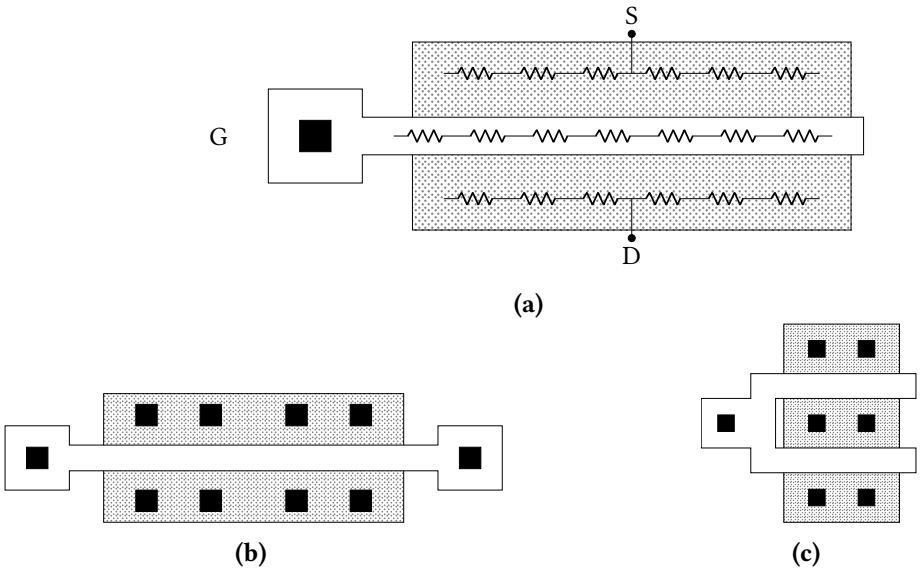


Figure 2.22 – Schematic representation of the distributed resistance in the device. Noise is generated by the charge distributing across the terminal (gate, drain or source), whose material has a given distributed resistance **(a)**. The gate poly has usually a higher resistivity and, consequently, a higher noise contribution comes from its distributed resistance. Subfigures **(b)** and **(c)** show layout designs meant to reduce the distributed resistance, among other effects, by contacting the gate from two sides **(b)** or by increasing the number of fingers **(c)**.

oxide interface. This model accounts for a dependency on the overdrive voltage [39]:

$$S_{1/f}(f) = \frac{K_f\left(V_{OV}\right)}{C_{ox}WL}\frac{1}{f^{\alpha}} \tag{2.50}$$

Since usually the power spectral density for flicker noise has been observed to be independent of the inversion coefficient in NMOS, the model presented in (2.49) is generally used for N-channel MOS devices; on the other hand, since in P-channel devices the power spectral density of the flicker noise has been observed to increase with the inversion coefficient, the model presented in (2.50) is usually associated with PMOS.

The literature also highlights the contribution of the Lorentzian noise: the traps present at the $SiO_2 - Si$ interface trap and release charges with given time constants τ_k, depending on the position of the trap and its energy level. This results in a variation of the drain current Δi whose shape is defined in signal theory as Random Telegraph Signal (RTS). The power density spectrum of said current signal is [40]:

$$\frac{d\overline{i^2}}{df} = \Delta i^2 \frac{4\tau_k}{1+\omega^2\tau_z^2} \tag{2.51}$$

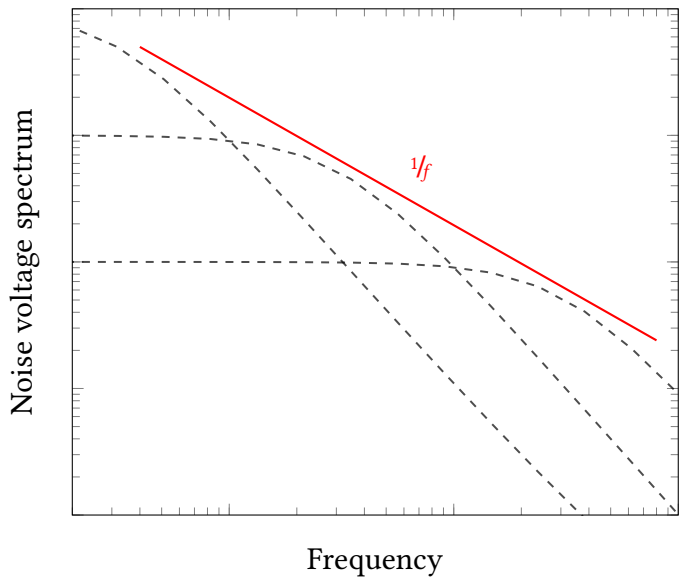


Figure 2.23 – $1/f$ origin from multiple Lorentzian components. Each trap noise power spectral density exhibits a corner frequency based on its characteristic time constant. A widespread range of traps with different time constants leads to a "smooth" overall effect whose power is inversely proportional to frequency.

According to the McWorther model already presented in (2.49), $1/f$ noise may be originated by a superposition of various Lorentzian contributions developing from a sufficiently large number of traps having different time constants, like shown in fig. 2.23. The current noise power spectral density of Lorentzian contributions is[41]

$$S_{Lorentzian}\left(f\right) = \frac{\left(\Delta I\right)^2 \tau}{1 + \left(2\pi f \tau\right)^2} \tag{2.52}$$

where τ is the characteristic time constant of the trap[II]. The Lorentzian noise is also known as Random Telegraphic Noise (RTS), for its unique behavior of jumping between two definite levels; the difference between the two levels is what is reported as ΔI in the previous equation. The power spectral density appears usually as characterized by a constant plateau, followed by a $\frac{1}{f^2}$ roll-off after the corner frequency[42]

$$f_c = \frac{1}{2\pi\tau} \tag{2.53}$$

[II]The underlying assumption here is that the noise is symmetrical, i.e. the time constants linked with the trapping and release of carriers are the same.

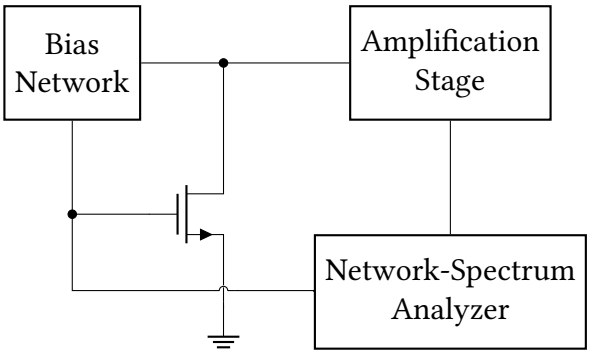


Figure 2.24 – Block diagram of the entire setup for the noise measurements. Bias network and amplification stage are hosted on the same PCB as the socket needed to connect the DUT.

2.2.1 Test setup and methodology

The scope of this set of measurement is to measure the noise referred to the input of the DUT. In order to accomplish this, a spectrum analyzer is needed; the measurements that will be reported have been performed using an HP4195A manufactured by Hewlett Packard. The instrument itself is a source of noise, most likely comparable with the noise of the DUT, hence an amplification stage is needed to ensure that the noise of the device is significantly higher than the noise coming from the analyzer. Furthermore, an adjustable bias circuit is needed to characterize the noise of the device in different bias conditions, creating a meaningful dataset useful for the design of analog circuits involving the characterized technology. A schematic view of the testing setup is presented in fig. 2.24. As it can be seen from the scheme, the amplified noise is used as input for the spectrum analyzer, but the desired result would be the input-referred noise. In order to calculate the input referred noise power spectral density, the measured (amplified) output can be divided by the squared module of the transfer function, computed $s = j\omega$, from the input, i.e. the gate of the DUT, to the output, i.e. the output of the amplifier. The HP4195A analyzer has a network analyzer functionality that can be employed to perform this kind of operation. Overall, the measurement procedure should follow these steps:

1. set the instrument to act as a network analyzer, connect its stimulus port to the gate of the DUT and its input port to the output of the amplifier;

2. bias the DUT in a bias point of choice;

3. use the network analyzer to extract the frequency response $T\left(j\omega\right)$: this is accomplished by the network analyzer by sending a sinusoidal signal into the stimulus port, and measuring the response at the input port. The frequency of the stimulus signal is swept to construct the frequency response in the desired frequency band;

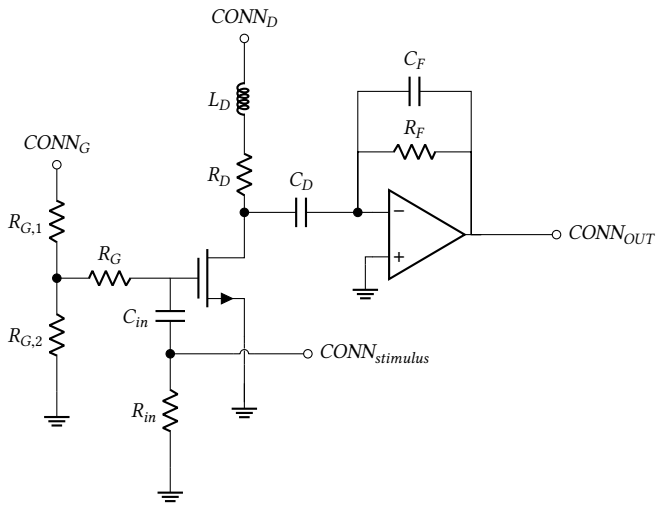


Figure 2.25 – Partial schematic of the PCB used in the noise measurements. Socket connection to the DUT and amplifier supply voltages are omitted for simplicity.

4. set the instrument to act as a spectrum analyzer and measure the output (amplified) noise power spectral density $\frac{d\overline{V_{n,out}^2}}{df}$. During this step the network analyzer stimulus port is physically disconnected from the DUT to ensure that no noise contribution is coming from that port;

5. the noise spectral density referred to the input $\frac{d\overline{V_{n,in}^2}}{df}$ can then be calculated using the frequency response module $\left|T\left(j\omega\right)\right|$ and the output noise power spectral density:

$$\frac{d\overline{V_{n,in}^2}}{df} = \frac{1}{\left|T\left(j\omega\right)\right|^2} \cdot \frac{d\overline{V_{n,out}^2}}{df} \tag{2.54}$$

Note that, in order to have a cleaner noise spectrum, steps 4 and 5 are repeated multiple times and their mean is taken as the final result of the measurement. The results shown in this work come from averaging over 20 individual measurements. In order to perform these steps, a PCB has been designed to host the adjustable DUT bias network and the amplifier; a partial schematic is presented in fig. 2.25. The schematic is constituted by two main blocks, covering the main purposes of the circuit: bias and amplification. The bias circuitry needs to be able to set the DC voltages at the gate and drain terminals in order to put the DUT in the desired working condition. The gate bias consists of by the resistive divider $R_{G,1}$ and $R_{G,2}$, used to set the gate DC voltage; resistor R_G in series with the gate and capacitor C_{in} form a low-pass filter to remove the bias resistors $R_{G,1}$, $R_{G,2}$ noise contribution in the bandwidth of interest. Capacitor C_{in} also doubles as a high-pass filter for the network analyzer stimulus, while R_{in} is a 50 Ω termination for the network analyzer stimulus transmission line. It should be noted that the resistor

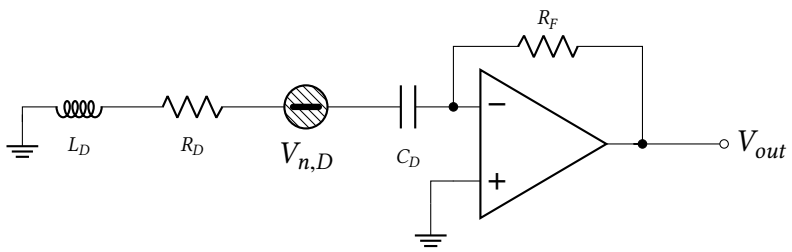


Figure 2.26 – Simplified schematic to model the noise contribution due to drain resistance R_D and its filtering.

R_{in} acts as a noise source during the spectrum analyzer steps (while the stimulus is disconnected); during the noise measurements, an additional 50 Ω termination has been connected in parallel to R_{in} by screwing it on the stimulus connector. The resulting 25 Ω resistor contribution has been compensated post-measurements by quadratically subtracting it from the measured data. A small test point has been added between R_G and the DUT gate to allow precise measurement of the gate voltage. The drain bias has been designed to be controlled by varying the voltage at the connector $CONN_D$, and a high-accuracy shunt resistor has been added in series with the drain for measuring the DUT drain current I_D. Inductor L_D and capacitor C_D have been added to filter the noise contribution of resistor R_D, the capacitor doubling as a high-pass filter for the amplifier input. This filter can be analyzed with the simplified scheme reported in fig. 2.26; the transfer function is:

$$\frac{V_{out}}{V_{n,D}} = \frac{sC_D R_F}{1 + sC_D R_D + s^2 C_D L_D} \tag{2.55}$$

which shows a low-pass response filtering the contribution of the noise in the bandwidth of interest. The poles lie at a frequency of:

$$f_t = \frac{1}{2\pi\sqrt{C_D L_D}} \tag{2.56}$$

In these measurements, the effect of the bulk bias voltage on the noise performance has not been studied. The bulk was therefore kept directly to ground. The circuit is meant to be able to perform measurements on both N-channel and P-channel devices: the source of the device is always connected to ground, and the voltages at $CONN_G$ and $CONN_D$ are driven positive for NMOS and negative for PMOS; the bulk, being connected to ground, is always shorted to the source. The amplifier has been chosen alongside commercially available ones: regarding the characteristics needed by this integrated circuit, the scheme in fig. 2.27 should be considered. In this, the amplifier has been

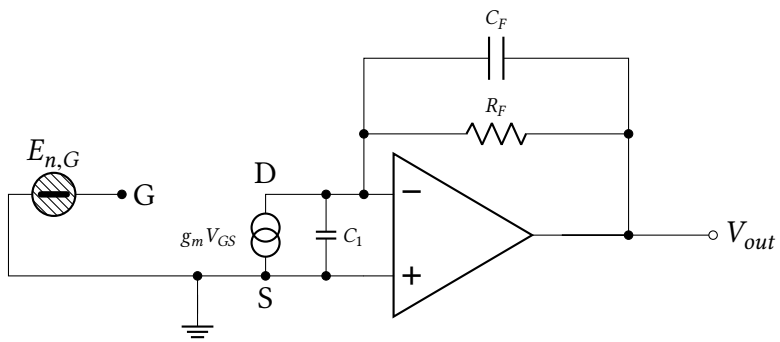


Figure 2.27 – Simplified scheme of the amplifier and small-signal model of the DUT used to analyze the transfer function from the equivalent series noise generator $E_{n,g}$ to the TIA output V_{out}.

simplified to show a single-pole response with the following frequency response:

$$A\left(j\omega\right) = \frac{A_0}{1 + j\omega\tau_0} \tag{2.57}$$

The C_F represented in the amplifier feedback can be as low as the parasitic capacitance between input and output, but a footprint has been foreseen on the PCB to grant the possibility of changing its value. C_1 is the sum of various contributions, namely the drain-source capacitance of the DUT, the input capacitance of the amplifier and the parasitic capacitance shorting the DUT drain, which originates from the package, the metal lines and the PCB itself. The parameter under measurement is the input-referred noise, modeled by the noise generator $E_{n,g}$. The transfer function from the input-referred noise to the output, with the previously mentioned assumption of the amplifier being single-pole, will be of the form

$$T_N\left(j\omega\right) = \frac{V_{out}}{E_{n,g}} = \frac{A_f}{s^2 + s\frac{\omega_0}{Q} + \omega_0^2} \tag{2.58}$$

where A_f is

$$A_f = \frac{g_m A_0}{\tau_0\left(C_1 + C_F\right)} \tag{2.59}$$

which, by assuming $C_1 \gg C_F$, becomes

$$A_f \simeq \frac{A_0}{\tau_0}\frac{g_m}{C_1} = \frac{g_m 2\pi GBP}{C_1} \tag{2.60}$$

while ω_0^2 and $\frac{\omega_0}{Q}$ are

$$\omega_0^2 = \frac{1 + A_0}{\tau_0 R_F \left(C_1 + C_F\right)} \simeq \frac{A_0}{\tau_0} \frac{1}{R_F C_1} = \frac{2\pi GBP}{R_F C_1} \tag{2.61}$$

$$\frac{\omega_0}{Q} = \frac{\tau_0 + R_F \left[C_1 + C_F \left(1 + A_0\right)\right]}{\tau_0 R_F \left(C_1 + C_F\right)} \simeq \frac{A_0 C_F}{\tau_0 R_F C_1} = 2\pi GBP \frac{C_F}{C_1 R_F} \tag{2.62}$$

where $GBP = \frac{A_0}{2\pi\tau_0}$ is the gain-bandwidth product of the amplifier. The DC gain is:

$$T(0) = \frac{A_f}{\omega_0^2} \simeq \frac{\frac{g_m 2\pi GBP}{C_1}}{\frac{2\pi GBP}{R_F C_1}} = g_m R_F \tag{2.63}$$

The complex-conjugated poles can lead to frequency response with a pronounced resonance peak; in order to have the maximum flatness in the pass-band the quality factor should be equal to $1/\sqrt{2}$, which poses a constraint on the values of the feedback network components:

$$Q = \sqrt{\frac{\omega_0^2}{\left(\frac{\omega_0}{Q}\right)^2}} \simeq \frac{1}{C_F} \sqrt{\frac{C_1}{R_F 2\pi GBP}} \tag{2.64}$$

The commercial amplifier chosen is the LMH6624[III], a wideband, ultra-low noise amplifier developed by Texas Instruments. The electrical characteristics of interest are the gain-bandwidth product of 1.5 GHz, the input-referred voltage noise of $0.92\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ and the input-referred current noise of $2.3\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$. Even with these expedients, the instrumentation, the amplification stage and the PCB itself contribute to the noise in the measurement in the bandwidth of interest; in order to compensate these additional contributions, the noise floor has been measured, after removing the DUT from the circuit, for each bias condition of each device and has been subsequently quadratically subtracted from the actual noise measurement. This is accounted for by adding a couple more steps to the measurement procedure:

1. the input-referred noise is measured as described previously;

2. the device is disconnected, all the biases in the circuit are kept the same and a new noise measurement is performed to get the input-referred noise floor (the

[III] Datasheet link

noise measured in these conditions is divided by the transfer function module to have it referred to the input);

3. the noise floor and the 25 Ω input resistance contributions are quadratically subtracted to obtain the final noise measurement.

Similarly to the previously described steps, 2 is repeated multiple times and the mean spectrum is used in step 3; as before, the mean spectrum is the result of 20 measurements. All measurements shown later in this chapter have been compensated for the noise floor and the contribution of the 25 Ω resistor[IV] at the DUT gate and have been performed over a frequency range from 3 kHz to 70 MHz at drain currents of 25 µA, 50 µA, 100 µA, 200 µA and 500 µA and at drain-source voltages $\left|V_{DS}\right|$ of 0.8 V. All measurements have been acquired by controlling the HP4195A via GPIB port with a simple python script; all further analysis has been done in MatLab.

2.2.2 Noise measurement data

This section will present the results of noise measurements performed using the setup and following the procedure described in the previous section. The power spectral density of the noise referred to the gate of the DUT can be written as:

$$S_e^2\left(f\right) = S_{WN}^2 + S_{1/f}^2\left(f\right) \tag{2.65}$$

The model is composed of a S_{WN}^2 term, independent of frequency, and a $1/f$ term inversely proportional to the frequency. The contribution of the source and drain parasitic resistances are generally negligible in devices operated in low current density conditions [43], which is the case for most devices in low-power electronics for radiation detectors. Hence, the following analysis will consider only the channel thermal noise contribution in S_{WN}^2. Given this, referencing (2.47) and (2.49), it is clear that at lower frequencies, where the flicker component is dominant, noise should appear inversely proportional to the device area; the higher frequencies, where white noise is dominant, should exhibit an inverse dependence to the device transconductance g_m and consequently to the square root of the drain current, at least for devices biased in strong inversion. This behavior can be observed in the noise measurements reported in figs. 2.28 to 2.31 reporting the square root of the noise power spectral density. Some of the devices, e.g. the W=600 µm L=120 nm NMOS device shown in fig. 2.28 appear to present a higher noise contribution at lower frequencies, most likely explained by Lorentzian components. As can be seen, the Lorentzian components seem to change alongside the device bias, as previously

[IV] All measurements have been performed at room temperature.

shown in the literature [44], albeit the dependence seems to change with the device channel length. Looking at the measurement results, a qualitative consideration can be drawn: in PMOS devices, $\frac{1}{f}$ noise (ideally a line in the shown log-scale plots) seems to shift depending on the drain current of the device, behavior that is on par with what is expected as per the McWorther model reported in (2.50), in which the K_f parameter, dependent on the overdrive voltage, acts exactly as a shift in the $\frac{1}{f}$ noise line. This behavior can be seen in most PMOS devices reported, but seems to be especially present in those with a smaller width, e.g. the W =100 µm, L =720 nm PMOS device reported in fig. 2.31.

Regarding the white noise, there seems to be good accordance between the white noise as calculated according to (2.47) and the measurements, as shown by the aggregate plots in fig. 2.33. The white noise has been extracted from the measurements by fitting the data with the function in (2.65). The method tends to fall apart when faced with the combined effects of small devices and the bandwidth limits of the measurement setup: in these cases, the $\frac{1}{f}$ noise overpowers the white by moving the noise corner at higher frequencies, reducing the amount of available white noise data. The left plots in fig. 2.33 show a visual comparison of the measured versus theoretical white noise, the first quadrant bisector being the ideal result, for all devices at a given gate length. Additionally, the errors with respect to the theoretical result can be calculated as the absolute value of the ratio between the measured and theoretical white noise, as follows:

$$\varepsilon_{WN} = \left| \frac{WN_{measured}}{WN_{theoretical}} - 1 \right| \tag{2.66}$$

The so calculated errors are presented in the same figure, in the plots on the right. The model fitting the dataset seems to fail for short channel lengths, highlighted by both the theoretical vs measurement results and the error plots; this is likely due to the fact that the data under analysis is mainly composed of $\frac{1}{f}$ noise, making the fit quite difficult. Attempts were made at weighting the data more where the measurement showed clear $\frac{1}{f}$ or white noise behavior, but results were discarded due to the weights being deemed quite arbitrary and due to the rising worry of overfitting the data, defeating the approach of estimating the white noise using the model itself.

Moving onto the $1/f$ noise, the K_f parameter shown in Hooge's model in (2.50) can be extracted. Parameter α_f is added to account for the deviation from the ideal slope of -1; note that the α_f parameter changes the unit of measure of K_f. In order to provide a meaningful comparison of the extracted K_f parameters, data analysis has been carried out in two steps: the first uses a model where both K_f and α_f are free parameters, allowing to extract α_f for each device. The second step performs another fit of the noise data, using a model where K_f is a parameter and α_f is set to the mean value found in the

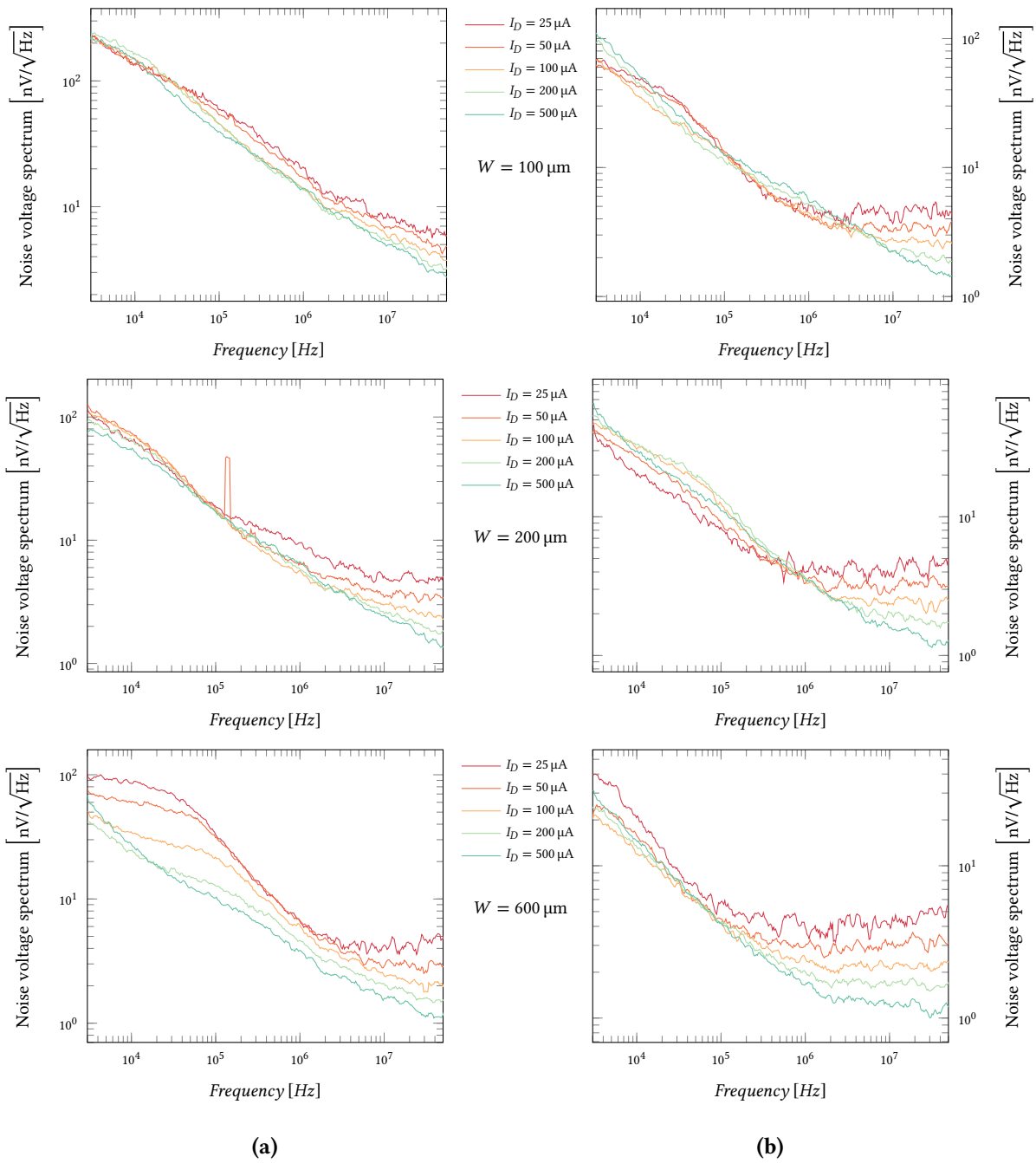
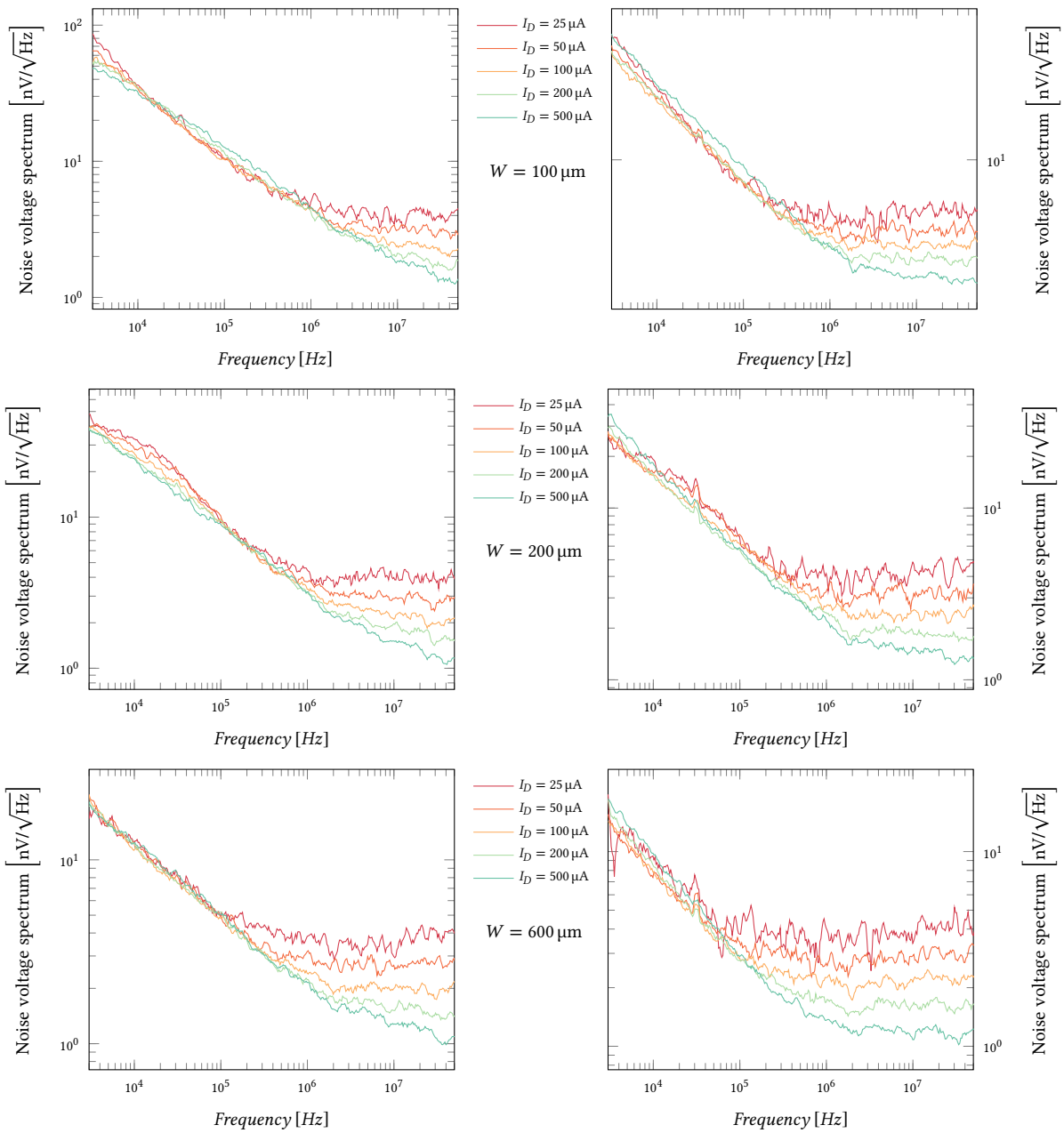


Figure 2.28 – Noise measuremens curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =120 nm, different drain currents and channel widths (from the top down) of 100 µm, 200 µm and 600 µm. Data shown is not raw, but a mean value of 20 successive measurements.



(a) **(b)**

Figure 2.29 – Noise measurement curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =240 nm, different drain currents and channel widths (from the top down) of 100 µm, 200 µm and 600 µm. Data shown is not raw, but a mean value of 20 successive measurements.

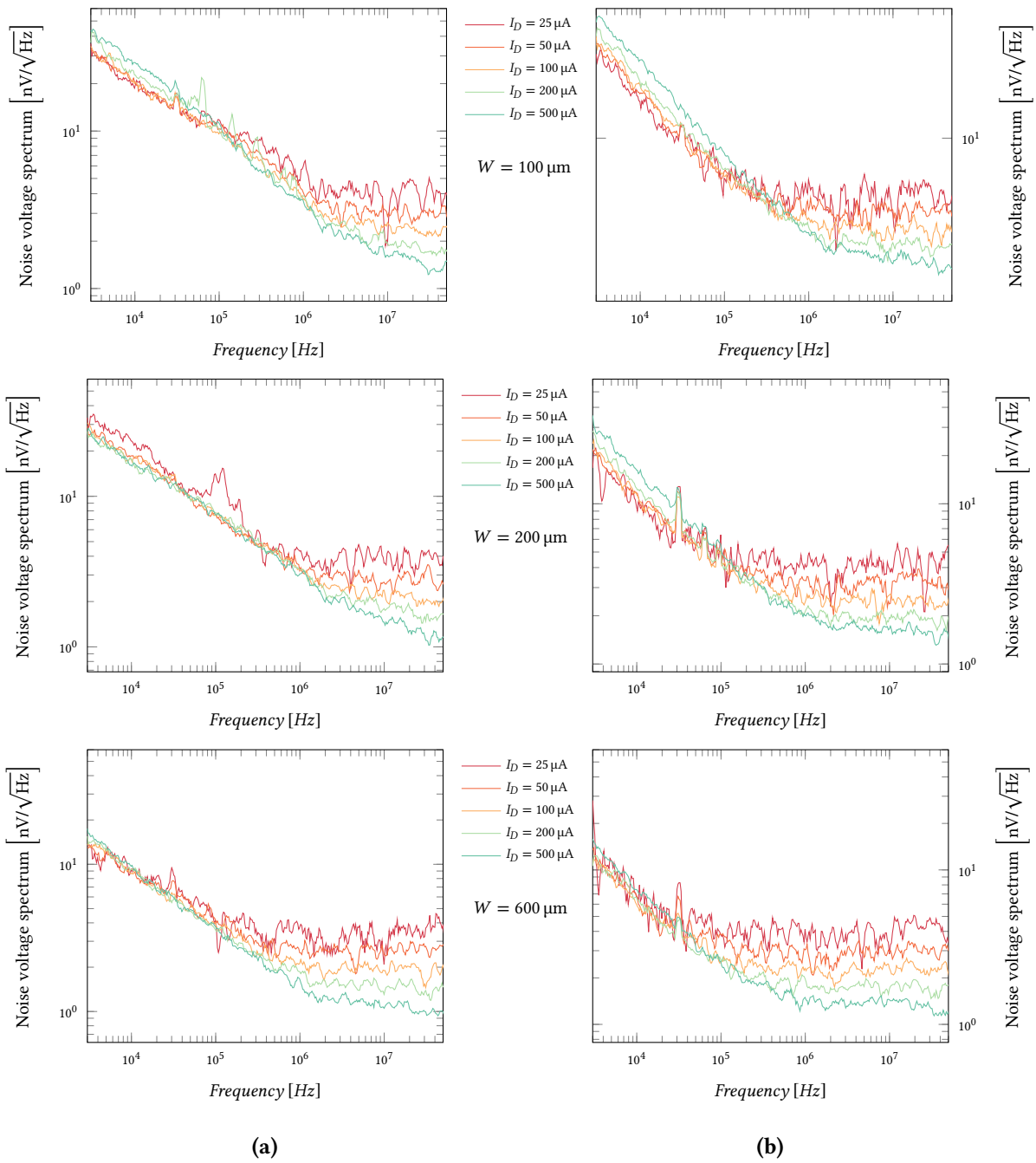


Figure 2.30 – Noise measurement curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =360 nm, different drain currents and channel widths (from the top down) of 100 µm, 200 µm and 600 µm. Data shown is not raw, but a mean value of 20 successive measurements.

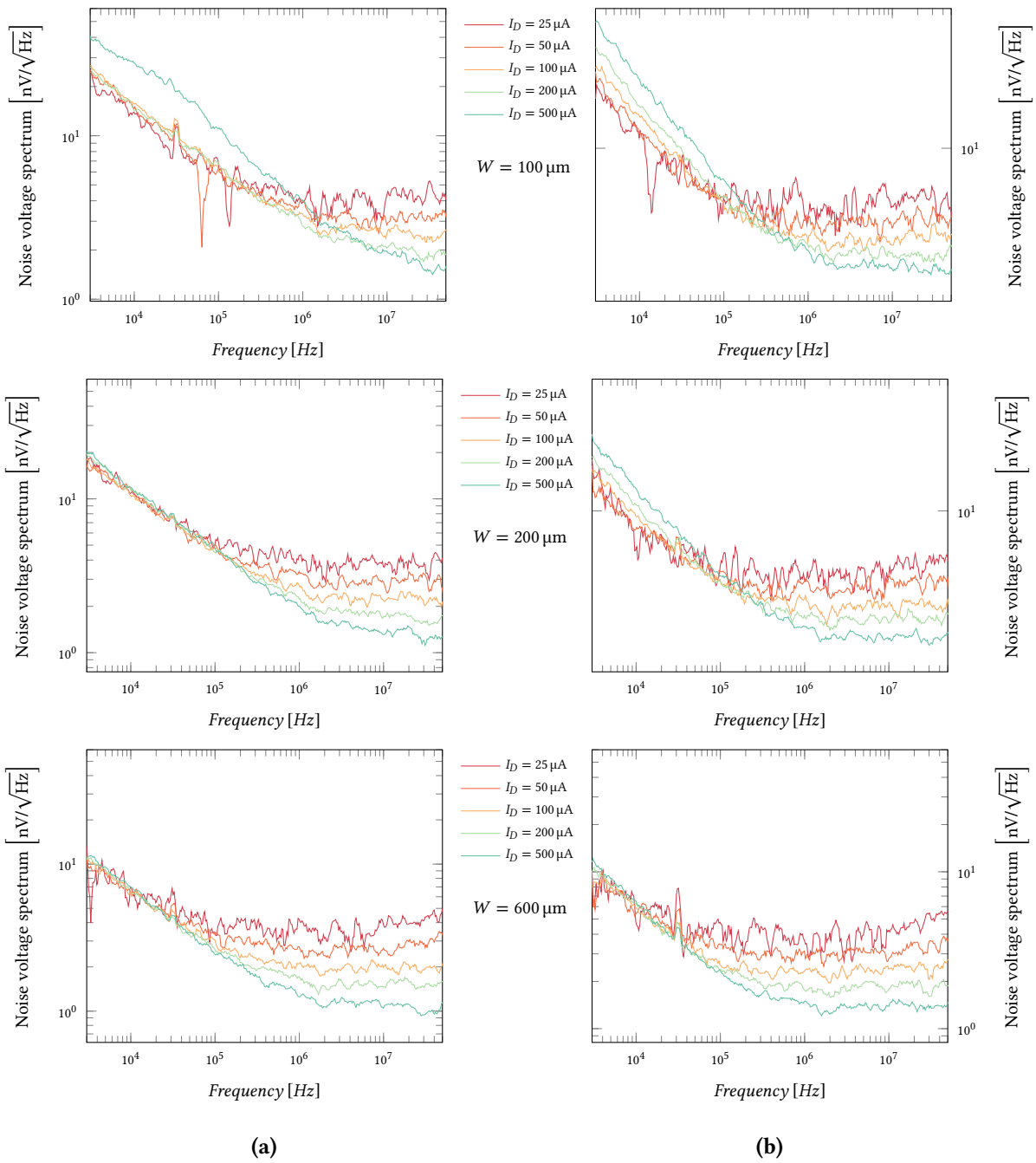


Figure 2.31 – Noise measurement curves for N channel devices **(a)** and P channel devices **(b)** with channel length of L =720 nm, different drain currents and channel widths (from the top down) of 100 µm, 200 µm and 600 µm. Data shown is not raw, but a mean value of 20 successive measurements.

previous analysis. The described procedure has been carried out separately for NMOS and PMOS devices, considering all currents (25 µm, 50 µm, 100 µm, 200 µm and 500 µm) and all devices apart from the thick oxide ones, and is reported in the form of a box plot in fig. 2.32.

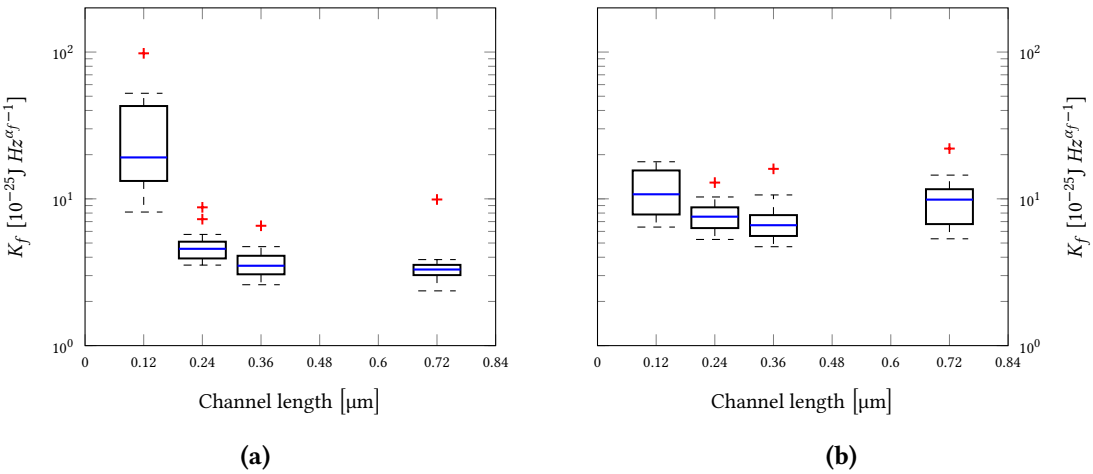
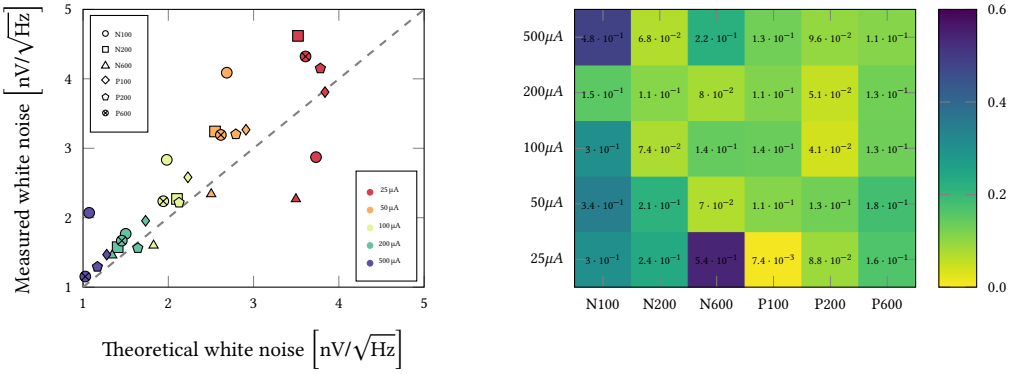
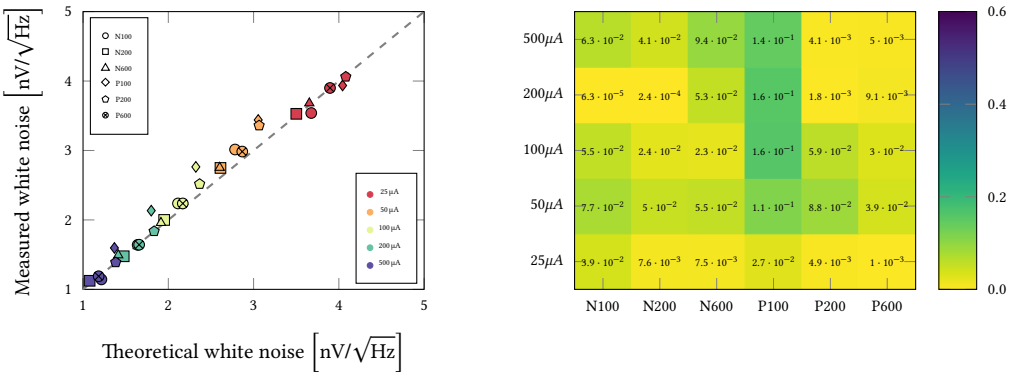


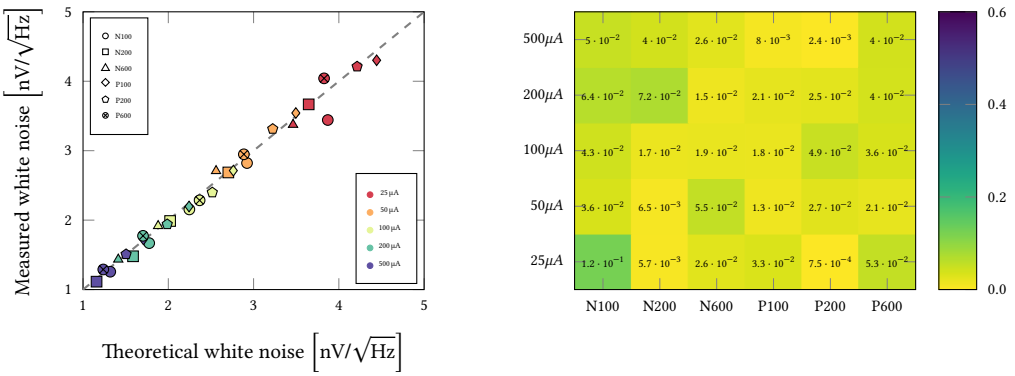
Figure 2.32 – Boxplot showing the extracted K_f values for NMOS **(a)** and PMOS **(b)** devices versus the channel length. The extracted α_f has a value of 0.8012 for NMOS devices and a value of 0.9858 for PMOS devices.



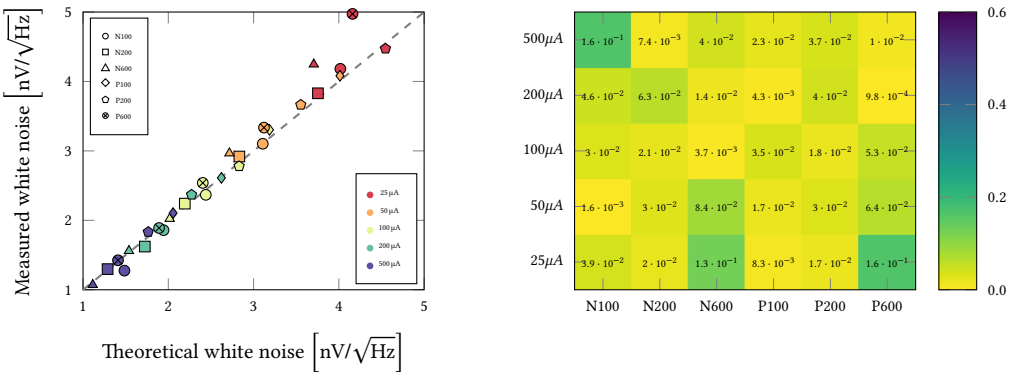
(a)



(b)



(c)



(d)

Figure 2.33 – Aggreated data of white noise measurements for channel lengths of 120 nm **(a)**, 240 nm **(a)**, 360 nm **(c)** and 720 nm **(d)**. Left plots show a comparison between theoretical and measured white noise; right plots show the error ratio.

3

Bandgap

A bandgap voltage reference, or bandgap reference (BGR) for short, is a crucial block in both analog and mixed signal circuits: it is used to provide a reference voltage constantand independent, as much as possible, of temperature, supply voltage and process variations. As mentioned, it finds application in analog circuits, for example in the generation of current and voltage references [45], to bias different blocks; or in power electronics, such as in LDOs [46], where the error amplifier relies on a stable voltage reference in order to close the loop and regulate the output. But bandgaps are widely used in mixed signal circuits as well: the presence of a stable reference allows for meaningful comparisons, which are crucial in conversion circuits like ADCs or DACs [47] and consequently in threshold comparison circuits. This chapter will cover the basic principle behind a bandgap reference and the design of a current mode implementation in the same 110 nm technology used in the rest of this thesis work.

3.1 Bandgap reference working principle

The basic idea behind a bandgap reference is quite simple: combine an element that shows characteristics proportional to the absolute temperature (PTAT) with another one that features complementary to the absolute temperature (CTAT) characteristics using a circuit that is insensitive to variations on the supply voltage or due to process parameters fluctuations. Given this premise, the working principle of the bandgap can be better understood by introducing its components first.

3.1.1 Complementary to Absolute Temperature component

Once the current is set, a PN junction forward voltage shows a dependency complementary to temperature; this is interesting in many fields, such as the design of LEDs, where a shift in temperature due to the heat dissipated in the junction causes a shift in the

forward voltage of the junction itself, changing the dominant frequency of the emitted radiation [48]. To no surprise, given the topic of this chapter, the same principle is widely used in bandgap references as a mean to obtain an element with a negative temperature coefficient; this behavior is usually exploited by using one of the PN junctions in a bipolar device, thus the following analysis will focus on this instance, although the same can be applied to other structures including a PN junction. In a bipolar device, the collector current can be expressed as

$$I_C = I_{C0} \cdot e^{\frac{V_{BE}}{V_T}} \tag{3.1}$$

where $V_T = {}^{kT}/_{q}$ is the thermal voltage, and I_{C0} is the saturation current. The latter has the following relation with temperature

$$I_{C0} \propto \mu k T N^2 \tag{3.2}$$

where N is the doping concentration. The temperature dependence of the components of this equation can be expressed as [49]

$$\mu \propto \mu_0 T^{-3/2} \tag{3.3}$$

$$N \propto T^3 \cdot e^{-\frac{E_g}{kT}} \tag{3.4}$$

where $E_g \simeq 1.12eV$ is the bandgap energy of silicon. From these dependencies, an overall expression of the saturation current I_{C0} containing the temperature can be extracted as

$$I_{C0} = T^{5/2} e^{-\frac{E_g}{kT}} \tag{3.5}$$

The base-emitter voltage is the forward voltage of a PN junction used in the bandgap reference operation; knowing that $V_{BE} = V_T \cdot ln\left(\frac{I_C}{I_{C0}}\right)$, by replacing I_{C0} with its expression in (3.5) and calculating the derivative with respect to the temperature, the temperature coefficient can be extracted.

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \cdot ln\left(\frac{I_C}{I_{C0}}\right) + V_T \cdot \left[\frac{I_{C0}}{I_C} \cdot \left(-\frac{I_C}{I_{C0}^2}\right)\right] \cdot \frac{\partial I_{C0}}{\partial T} \tag{3.6}$$

$$= \frac{\partial V_T}{\partial T} \cdot ln\left(\frac{I_C}{I_{C0}}\right) - \frac{V_T}{I_{C0}} \cdot \frac{\partial I_{C0}}{\partial T} \tag{3.7}$$

$$= \frac{k}{q} \cdot ln\left(\frac{I_C}{I_{C0}}\right) - \frac{V_T}{I_{C0}} \cdot \frac{\partial I_{C0}}{\partial T} \tag{3.8}$$

$$\frac{\partial I_{C0}}{\partial T} = \frac{5}{2} T^{3/2} e^{-\frac{E_g}{kT}} + T^{5/2} \left(-\frac{-E_g}{kT^2} \right) e^{-\frac{E_g}{kT}} \tag{3.9}$$

$$= \frac{5}{2} T^{3/2} e^{-\frac{E_g}{kT}} + T^{5/2} \frac{E_g}{kT^2} e^{-\frac{E_g}{kT}} \tag{3.10}$$

$$= \frac{5}{2} \frac{I_{C0}}{T} + I_{C0} \frac{E_g}{kT^2} \tag{3.11}$$

$$= I_{C0} \cdot \left(\frac{5}{2T} + \frac{E_g}{kT^2} \right) \tag{3.12}$$

Thus,

$$\frac{\partial V_{BE}}{\partial T} = \frac{k}{q} \cdot ln \left(\frac{I_C}{I_{C0}} \right) - \frac{V_T}{I_{C0}} \cdot I_{C0} \cdot \left(\frac{5}{2T} + \frac{E_g}{kT^2} \right) \tag{3.13}$$

$$= \frac{V_T}{T} \cdot ln \left(\frac{I_C}{I_{C0}} \right) - \frac{5}{2} \cdot \frac{V_T}{T} - V_T \frac{E_g}{kT^2} \tag{3.14}$$

$$= \frac{V_{BE} - \frac{5}{2} V_T - \frac{E_g}{q}}{T} \tag{3.15}$$

From (3.15) it can be seen that for small enough V_{BE}, the temperature coefficient is negative; furthermore, it can be deduced that the relation between V_{BE} and temperature is not linear, due to the appearance of the temperature itself in $\frac{\partial V_{BE}}{\partial T}$. Despite this, the simplified analysis further described in this chapter will assume a linear dependence, i.e. a constant temperature coefficient, with the justification that in the range of interest, $-40\,°\mathrm{C}$ to $70\,°\mathrm{C}$, the variation is roughly $13\,\mu\mathrm{V}\,°\mathrm{C}^{-2}$. Given this behavior, a CTAT voltage can be generated by using a PN junction in the circuit; a CTAT current can then be generated starting from the CTAT voltage. The scheme presented in fig. 3.1 is an example of a circuit exploiting the described behavior to generate a current with a negative temperature coefficient, i.e. a CTAT current: the amplifier inputs are kept equal by the feedback, assuming a high enough gain and not considering the offset effects. These nodes are at a voltage equal to the V_{BE} of BJT Q_3 which, as already discussed, exhibits a negative dependency on the temperature. The current generated using resistor R and assuming again no further variation due to the temperature, will also be CTAT. The one presented in fig. 3.1 is a simple circuit, but effective in highlighting some of the obstacles encountered when designing a bandgap circuit: the assumptions made in this analysis offer an idea of the possible problems. Namely, the gain and offset of the amplifier are crucial in the effectiveness of the bandgap reference. If they are not taken into account, components in the circuit can introduce further temperature dependencies, impacting the overall circuit performance.

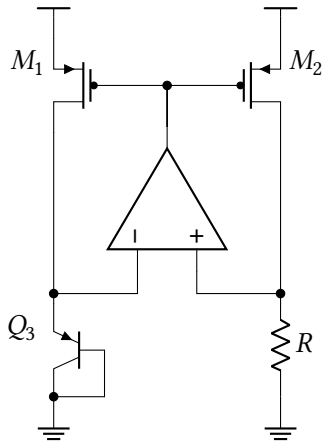


Figure 3.1 – Example schematic of a circuit capable of generating a CTAT current.

3.1.2 Proportional to Absolute Temperature component

As already mentioned, the bandgap reference working principle is based on the combination of a CTAT component and a PTAT one. Regarding the PTAT component, i.e. having a positive temperature coefficient, it can be proved that the difference between the V_{BE} of two transistors working at two different current densities $I_\rho = \frac{I_C}{I_{C0}}$ is in fact proportional to the absolute temperature. Considering that:

$$V_{BE} = V_T \log\left(\frac{I_C}{I_{C0}}\right) \tag{3.16}$$

then the difference between the base-emitter voltages of two matched[I] BJTs, one having a current density N times the current density of the other, is:

$$\Delta V_{BE} = V_T \ln\left(N\frac{I_C}{I_{C0}}\right) - V_T \ln\left(\frac{I_C}{I_{C0}}\right) \tag{3.17}$$

$$= V_T \ln(N) \tag{3.18}$$

This means that ΔV_{BE} has a direct proportionality to the temperature thanks to the presence of the thermal voltage V_T:

$$\frac{\partial \Delta V_{BE}}{\delta T} = \frac{k}{q}\ln(N) \tag{3.19}$$

Equation (3.19) also shows that the temperature dependency can be controlled by changing the ratio of the current densities N, providing a useful degree of freedom when designing the bandgap structure. The conceptual scheme reported in fig. 3.2a

[I] Having the same I_{C0}

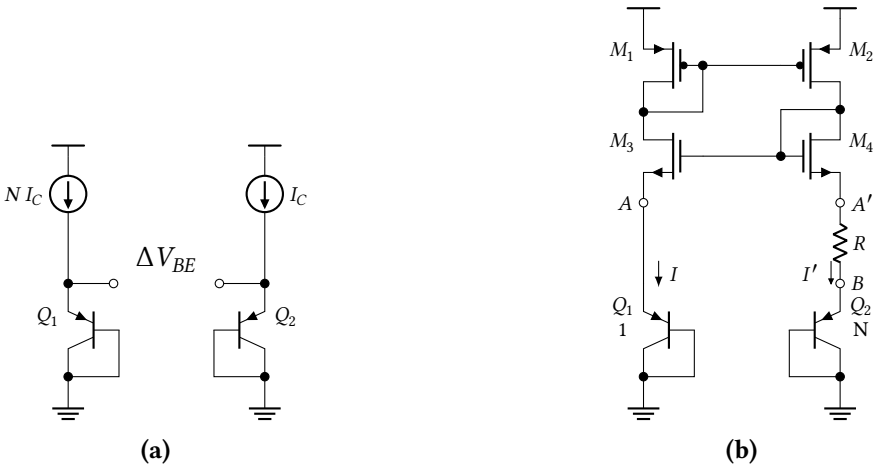


Figure 3.2 – Example schematics of circuits capable of generating a PTAT voltage **(a)** and a PTAT current **(b)**.

shows a possible implementation of a circuit generating a PTAT voltage: by biasing the nominally identical Q_1 and Q_2 transistors with different currents, the difference between their V_{BE} will have a positive temperature coefficient, although the behavior is usually exploited by having differently sized transistors, as opposed to the case of identical transistors biased with the same current. The same concept can be applied to the scheme presented in fig. 3.2b to generate a PTAT current: here the two transistors Q_1 and Q_2 are biased with the same current, and the difference in current density is obtained by making the junction area of Q_2 N times bigger than Q_1 junction area. The double current mirror formed by MOSFETs M_1, M_2, M_3 and M_4 is used as a simple mean of keeping nodes A and A' equal, in addition to keeping $I = I'$. Knowing this, the voltage at nodes A and A' are equal to the V_{BE} of Q_1, and voltage at node B is the V_{BE} of Q_2; hence, the voltage drop across resistor R is the V_{BE} difference between two transistors with different current densities[II] and thus is PTAT. Assuming no further contribution due to the temperature coefficient of the resistor, the current I' (and consequently, I) is PTAT as well.

3.1.3 Bandgap reference: voltage mode

As mentioned at the beginning of this section, the operating principle of the bandgap reference circuit is based on the combination of a PTAT and a CTAT component. The classic design makes use of an amplifier to implement the schematic function represented in the block diagram of fig. 3.3. The mostly known bandgap voltage reference circuit has been proposed by Widlar[50] and its first implementation, which is reported in fig. 3.4, by Kuijk[51]. In the circuit, the amplifier, assuming the gain is relatively high,

[II]Considering the difference in junction area.

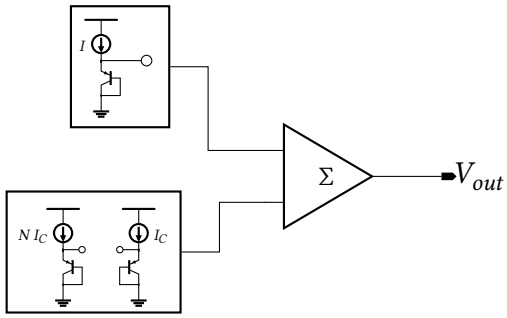


Figure 3.3 – Block diagram of a voltage-mode bandgap reference. The reference voltage is generated as a weighted sum of two components with opposite temperature dependencies.

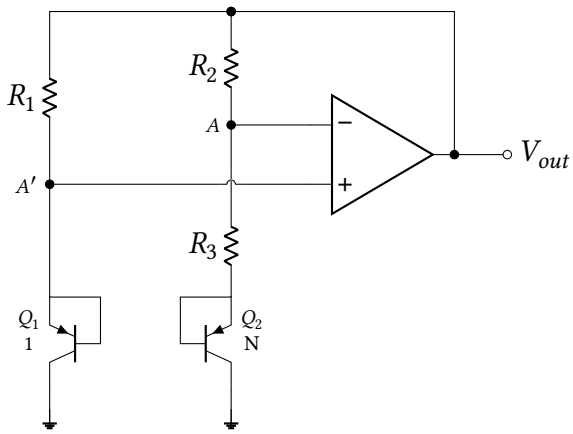


Figure 3.4 – Voltage mode bandgap reference circuit as initially proposed by Kuijk.

will keep node A and A' at roughly the same voltage; the voltage at node B is just equal to the V_{BE} of transistor Q_2. From the results shown in (3.18) it is clear that the voltage between node A' and B, i.e. the voltage across the resistor, is $V_T \ln(N)$. From this, the current in the resistor branch must be $I_{R_3} = \frac{V_T \ln(N)}{R_3}$, making the output voltage equal to:

$$V_{out} = \overbrace{V_{BE,Q2}}^{CTAT} + \overbrace{V_T \ln(N)}^{PTAT} + \overbrace{\frac{V_T \ln(N)}{R_3} \cdot R_2}^{PTAT} \tag{3.20}$$

$$= \underbrace{V_{BE,Q2}}_{CTAT} + \underbrace{V_T \ln(N) \left(1 + \frac{R_2}{R_3}\right)}_{PTAT} \tag{3.21}$$

which is independent of the temperature coefficient of resistors R_3 and R_2, assuming it equal for both (a fair assumption if the resistors are made from the same material). The output voltage of a bandgap reference constructed as described is usually around 1.2 V, which is usually indicated as the source of the bandgap name, given its similarity to the

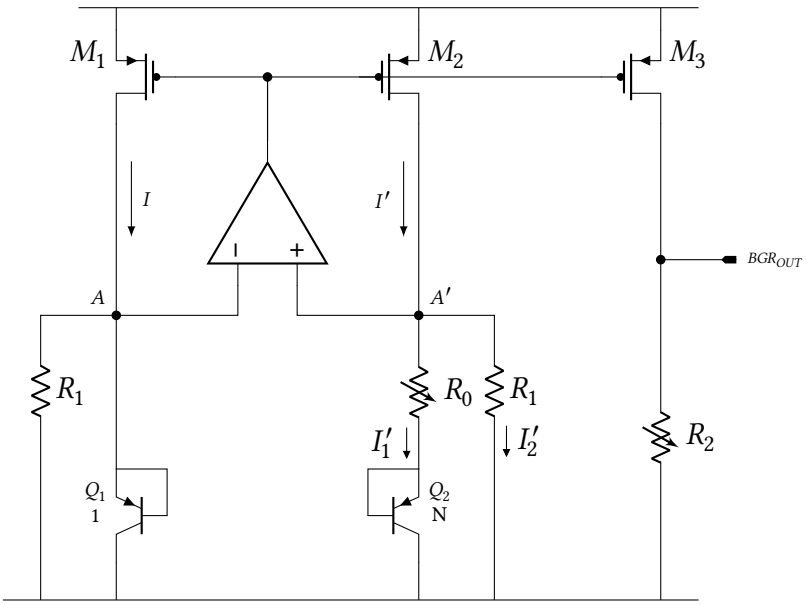


Figure 3.5 – Current mode bandgap reference circuit.

1.21 eV energy bandgap of the silicon. Although this is an interesting piece of trivia, the 1.2 V output actually constitutes a limit for the currently widespread technologies, which employ a core supply voltage most likely comparable to, if not lower than, the 1.2 V output of a voltage mode bandgap reference, making this design approach not feasible. This is true for the 110 nm technology used in this thesis work, which has a core supply voltage of exactly 1.2 V.

3.1.4 Bandgap reference: current mode

In order to circumvent the problem posed by the low supply voltage, the circuit must be slightly modified in order to be able to output a smaller than 1.2 V reference voltage; the basic principle remaining the same, this second approach will make use of currents as the PTAT and CTAT elements[52]. The schematic of the so-called current-mode bandgap reference is reported in fig. 3.5. The circuit can be seen as a combination of the schematics reported in figs. 3.1 and 3.2b. As before, the amplifier allows node A and A' to stay at roughly the same voltage; now, looking at current I', it is clear that it can be traced back to the sum of two components, I_1' and I_2'. The I_1' component is quite familiar, since it appeared in the previous analysis of the voltage mode bandgap: said current can be seen as the voltage drop across resistor R_0 divided by its resistance, and since the voltage drop is ΔV_{BE}, it is PTAT, as previously shown in (3.19). The second component, I_2', is just $\frac{V_{A'}}{R_1}$, and thanks to the amplifier, $V_{A'} = V_A = V_{BE,Q1}$, thus CTAT,

as shown in (3.15). Thus, the current I' is

$$I' = \frac{V_{BE,Q1}}{R_1} + \frac{V_T \ln(N)}{R_0} \tag{3.22}$$

R, R_1 and N can be used as design parameters in order to minimize the temperature dependence of the current. Additional comments should be done on the remainder of the circuit: MOSFETs M_1 and M_2 are nominally equal and are biased in the same conditions, thus $I = I'$. Given this, resistor R_1, connected to the A node, can be used in order to have the same current flowing into transistors Q_1 and Q_2, thus having the control onto the current density just by changing the design parameter N. Lastly, the mirror formed by MOSFETs M_2 and M_3 is then used to output the I current and obtain a reference voltage by adding a simple resistor R_2. Thus the output voltage of the current-mode bandgap becomes

$$V_{out} = \left[\frac{V_{BE,Q1}}{R_1} + \frac{V_T \ln(N)}{R}\right] R_2 \tag{3.23}$$

It is worth noting that, again assuming it to be equal for all resistors, the output voltage is not influenced by the temperature coefficient of the resistors.

3.2 Bandgap reference design in a 110 nm CMOS technology

This section will be focused on the design of the current-mode bandgap reference that will be used in the ARCADIA project. The adopted scheme is an improved version of a current-mode bandgap; the main differences in the design are the adoption of diode-connected MOSFETs instead of BJTs and the usage of a second amplifier to improve the current mirror used by the output branch current.

3.2.1 Radiation damage mechanisms in semiconductor devices

The choice of substituting BJTs with MOSFETs warrants a brief introduction to the radiation damage effects in silicon devices which will be briefly covered in this section. Generally speaking, radiation damage mechanisms can be traced back to two sources:

Ionization damage Energy deposited in a semiconductor or in insulating layers (e.g. silicon dioxide, SiO_2) will release electron-hole pairs, which could get trapped in imperfections of the lattice, mostly present at the interface region between two different materials. Some may recombine, their effects lasting only a brief while as a passing transient, but others may lead to unwanted, relatively large, dose

and/or dose rate dependent concentration of charge in some regions of a device, leading thus to equally unwanted electric fields. These effects are mainly caused by the interaction with X-rays, γ-rays and charged particles with silicon dioxide and mainly affect devices based on surface conduction, such as MOSFETs.

Displacement damage Heavy particles can hit atoms present in the semiconductor, physically dislodging them from their site and damaging the lattice structure, creating defects and therefore altering the electric properties of the crystal. These effects are mainly caused by high energy neutrons and, in some cases, by charged particles, such as protons or ions, but also by high energy photons (e.g. γ-rays from ^{60}Co) through secondary electrons recoiled due to Compton scattering[53]. Devices based on bulk conduction, such as BJTs, are mostly affected by displacement damage.

The high fluences expected in some of the target applications of ARCADIA, namely the future colliders, are a non-negligible point against the use of bipolar structures, due to their sensitivity to bulk damage. In the past, radiation-hard ASICs were fabricated using silicon-on-insulator (SOI) technology, for their resilience to single event upsets[III] and single event latchups[IV] [54]. Even though SOIs show resilience to SEUs and SELs, the presence of an insulating layer, more often than not implemented using SiO_2, will make this technology sensitive to ionization damage: positive charge will accumulate in the insulation layer, generating a parasitic field, leading to performance degradation in analogue blocks. A solution proposed in the literature based on a biasing layer buried between two insulation layers, can be used to counteract the onset of the parasitic field, but leads to an increase in the manufacturing cost. As an upside, technology scaling has played a major role in the increase of the radiation hardness of the technology nodes: while the decreasing thickness of the gate oxides and the rise of tunneling phenomena can lead to undesired effects, such as an increase in the gate current, the same can be useful in helping get rid of the accumulated charge generated in the oxide after an ionizing particle interaction. For this reason, rad-hard bandgap circuits usually employ MOSFETs instead of BJTs in their design[54, 55, 56], and some employ some additional layout technique, such as the use of enclosed layout devices in order to minimize the radiation damage effects. The reason why MOSFET devices are used when faced with radiation hardness requirements should be clear, but replacing BJTs with MOSFETs is not as straightforward as it may have sounded so far, and this warrants a brief explanation.

[III] Briefly, a single event upset (SEU) is a radiation-induced bitflip in a memory array.

[IV] Equally briefly, a single event latchup (SEL) is a radiation-induced activation of a parasitic bipolar structure, leading to the creation of a low-impedance path between the circuit rails, often ending in a permanent thermal failure.

3.2.2 MOSFET behaviour in the sub-threshold region

A MOSFET which is biased in the weak inversion region has an $I_D - V_{GS}$ showing an exponential characteristic with a dependency on temperature[57], similar to the one of a BJT. The drain current follows the equation[58]:

$$I_D\left(V_{GS}, T\right) = \mu\left(T\right) C_{OX} \frac{W}{L} V_T^2 e^{\frac{V_{GS}-V_{th}(T)}{n_{sub}V_T}} \tag{3.24}$$

where μ is the carrier mobility, C_{OX} the gate oxide capacitance, V_T the thermal voltage and n_{sub} the subthreshold slope factor, already discussed in section 2.1.2.1. In the presented model, many of the parameters show some dependency on the temperature: the thermal voltage $\frac{kT}{q}$ has a direct dependency, while the mobility and threshold show more complex behaviors. In particular, the mobility can be expressed as [59] [V]

$$\mu\left(T\right) = \mu\left(T_0\right)\left(\frac{T}{T_0}\right)^{-m} \tag{3.25}$$

where m is a technology parameter with values ranging from 1.2 to 2; the threshold voltage temperature dependence can be expressed as well with a simplified equation

$$V_{th}\left(T\right) = V_{th}\left(T_0\right) - \gamma\left(T - T_0\right) \tag{3.26}$$

showing that the threshold lowers at higher temperatures [61]. Substituting then (3.25) and (3.26) in (3.24) as an expression of V_{GS}

$$V_{GS} = V_{th} + n_{sub}V_T \ln\left[\frac{I_D}{\mu\left(T_0\right)\left(\frac{T}{T_0}\right)^{-m} C_{OX}\frac{W}{L}V_T^2}\right] \tag{3.27}$$

Assuming $m = 2$ and no temperature dependence in the drain current, the logarithm argument loses its dependence from T; using these assumptions and deriving the previous equation, the temperature dependence of the V_{GS} of a MOSFET biased in weak inversion is

$$\frac{\partial V_{GS}}{\partial T} = \frac{\partial V_{th}}{\partial T} + n_{sub}\frac{k}{q}\ln\left[\frac{I_D L q^2}{\mu\left(T_0\right) T_0^2 C_{OX} W k^2}\right] \tag{3.28}$$

Even by a rough estimation, the logarithm argument appears to be less than 1, thus making the second term of the sum CTAT; (3.26) shows that the threshold voltage is also CTAT, thus the temperature coefficient of the V_{GS} of a MOSFET biased in weak

[V] A less simplified approach would be to consider different contributions to the mobility such as phonon scattering, surface roughness scattering and Coulomb scattering and their temperature dependency [60]. The presented analysis uses a simplified equation presented by Tsividis.

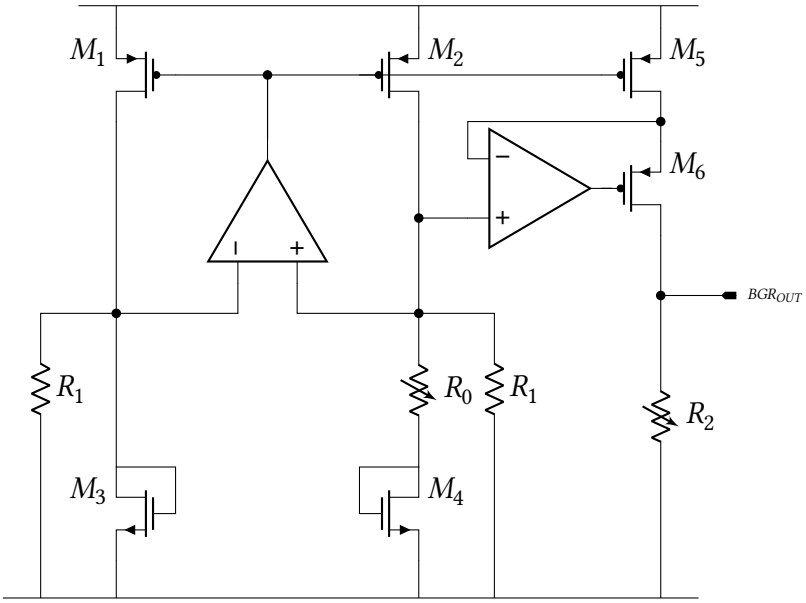


Figure 3.6 – Architecture of the current mode, rad-hard bandgap voltage reference developed in a 110 nm technology for the ARCADIA project.

inversion is negative, just as the BJT case seen previously. A proof very similar to the one presented in section 3.1.2 can show that the difference between two V_{GS} is PTAT, making the replacement of BJT with a MOSFET possible.

3.2.3 Implemented bandgap reference scheme

The complete schematic of the implemented current mode bandgap based on MOSFETs biased in the sub-threshold region and designed in the 110 nm technology already discussed in previous chapters is shown in fig. 3.6 and targets an output voltage of 600 mV. With respect to the current-mode implementation shown in fig. 3.5, the scheme used for ARCADIA employs diode-connected MOSFETs M_3 and M_4 instead of the BJTs for the reasons discussed in the previous section. It also includes an additional amplifier, maintaining the same V_{DS} for M_2 and M_5, therefore improving the performance of the M_2-M_5 current mirror. Furthermore, as can be seen from the reported schematic, an additional difference can be found in the use of programmable resistors R_0 and R_2; by looking at (3.23) it can be inferred that resistor R_0 can control the CTAT part of the equation, providing a well-defined trimming point for user compensation of the reference voltage vs temperature slope, $\partial V_{out}/\partial T$. A very similar point can be made for resistor R_2, which can be used to trim the absolute value of the reference voltage. As can be seen from (3.24) and (3.28), a threshold shift will influence both the current in the MOSFET and the temperature coefficient used in the bandgap operation: from what was discussed in the very brief introduction to radiation damage, ionizing particles

will induce electron-hole pairs, with holes typically trapped in insulating layers such as the gate oxide. When this happens, the trapped charge will induce a parasitic field, which will affect the threshold voltage of the MOSFET itself[62]. The trapped charge in the oxide is usually positive, leading to different effects based on the device type. For PMOS transistors, the trapped charge works against the creation of the channel, leading to an increase, in absolute value, of the device threshold; the effect is opposite in NMOS devices, in which the positive charge helps creating the channel, often leading to a decrease in the device threshold[VI]. This effect, when related to the total dose that ARCADIA will need to withstand in some applications[VII], and additional variations due to process or mismatch contributions are, in the author's opinion, a fair reason to advocate for the presence of trimming capabilities. The practical implementation of the programmable resistors will be discussed shortly.

3.2.3.1 Startup circuit

A common problem of a bandgap circuit can be found in its stable operating points; from a rough analysis of the circuit presented in fig. 3.6, it can be seen that a condition in which no currents are flowing in the circuit is definitely possible: both input nodes of the amplifier will be at ground, the amplifier output at VDD and BGR_{OUT} at ground as well; the circuit has in fact two stable operating points, one being the just described one. Such a state of the circuit is unarguably not very useful, thus the bandgap reference needs an auxiliary circuit to move into a functioning operating point. The rough idea of such a circuit, aptly named startup, is to lower the voltage at the output of the amplifier, forcing currents in MOSFETs M_1 and M_2 and thus forcing the operating point near the second stable one; the circuit should then disconnect once its task has been accomplished. The approach used is the one of a dynamic startup circuit, shown in fig. 3.7, where the timing of the disconnection of the startup from the bandgap core is given by the time constant of a charging capacitor. The behavior of the circuit can be seen by first considering capacitor C discharged: during startup, V_{DD} will rise to its nominal value, and a current will start flowing in the capacitor branch. M_{11} will have no impact on this operation, given that its gate is tied to V_{DD}. The capacitor charging current is mirrored by M_9-M_{10} onto the left branch of the circuit; given that the bandgap will probably be in the zero-current state described earlier, the gate of M_7 will be at ground, thus the drain current of M_9 will cause the gate of M_8 to rise, which in turn

[VI]This is true for relatively low doses, where the threshold shift is mainly affected by the charge trapped in the gate oxide, but the trend changes at higher doses, where the threshold begins to increase due to the increase in charge trapping at the interface between the oxide and the substrate[63]. In both cases, a case can be made for the addition of trimming capabilities in the bandgap circuit.

[VII]In the applications targeted at future colliders, the total dose can be up to ∼1 Grad[64].

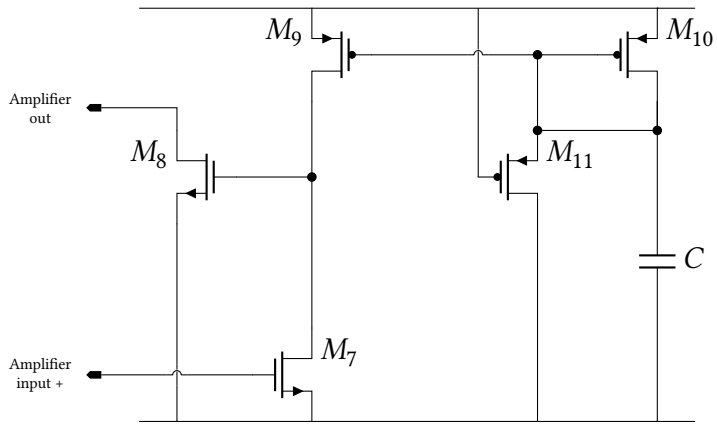


Figure 3.7 – Schematic of the implemented dynamic startup circuit for ARCADIA's bandgap voltage reference.

will lower the amplifier output node and move the bandgap core into its functional operating point. The rise of currents in the bandgap core will mean a rise in the voltage at the amplifier inputs and thus at the gate of M_7, bringing the gate of M_8 towards ground. Once C has fully charged, no current will flow in the mirror and the power consumption of the circuit will theoretically be zero. MOSFET M_{11} will come into play when V_{DD} is removed, discharging the capacitor and resetting the circuit to its initial state, ready for another startup.

3.2.3.2 Operational amplifier

Both operational amplifiers seen in fig. 3.6 have been designed starting from a mirrored cascode architecture, then ending the design phase with the scheme reported in fig. 3.8. The classic mirrored cascode amplifier is based onto a differential input pair whose generated differential current is mirrored by a cascode current mirror[65]. The mirror also deals with the differential to single ended conversion. Due to very little headroom simulated with the four transistor stack of the classic scheme, the cascoded current mirror has been replaced by a simple current mirror, lowering the gain of the amplifier. The input pair has been designed with a large area and interdigitated layout in order to minimize mismatch effects leading to an increased offset. The tail current generator is biased by a Widlar current generator, not shown in the reported scheme. During the design, the bandwidth of the amplifier has been considered to be less important than gain and offset, due to the fact that the temperature effects will act in the very low frequency range and thus the amplifier has been loaded with a capacitor at the output to ensure stability. The input pair has been chosen to be PMOSs mainly due to the need of a relatively low common mode, around 0.3 V, but has the added benefit of having a lower $1/f$ noise[66], as can be seen from the measurements reported in section 2.2.2. In

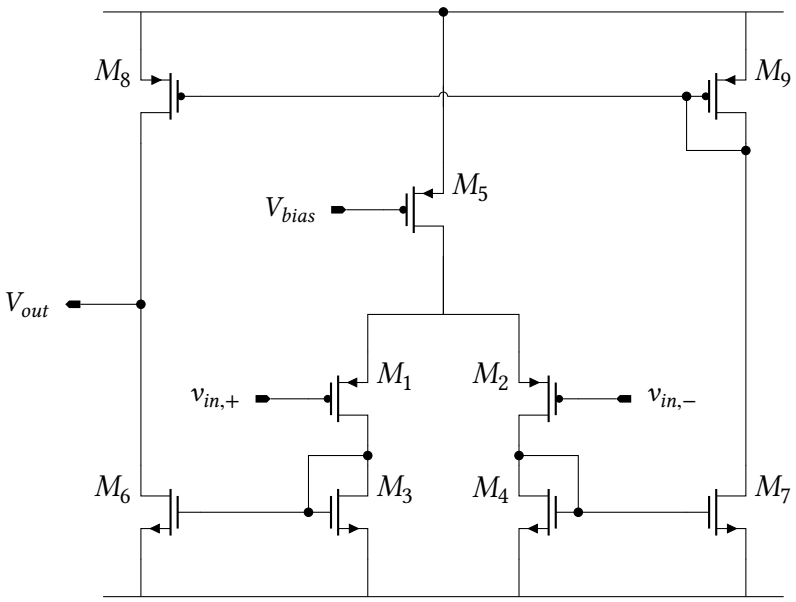


Figure 3.8 – Simplified mirrored cascode amplifier used in ARCADIA's bandgap.

order to minimize the impact onto the bandgap power consumption, the amplifier has been biased with a relatively low current resulting in a power consumption of 3.89 µW. The corner simulation results for gain and phase are reported in fig. 3.9. The amplifier has an open-loop gain ranging from 62 dB to 66 dB (66 dB typical) and a 61 deg to 69 deg (64 deg typical) phase margin.

3.2.3.3 Programmable resistors

Both programmable resistors R_0 and R_2 have been designed as a series of binary weighted resistors, each of which has a pass-gate in parallel to short the resistor terminals; the schematic is reported in fig. 3.10.

3.2.4 Bandgap simulations

A custom test-bench has been created in order to evaluate the performance of the bandgap reference. The chosen parameters of interest exposed by the test-bench are introduced fig. 3.11, showing a visual representation of most of the parameters used to evaluate the bandgap performance. As it can be seen, due to second order effects not considered in the analysis shown previously, a bandgap output without second order compensation shows a parabolic dependency to the temperature. The parameters of interest, given this shape, are the mean output, expected to lie near the 600 mV target, and the variation across the temperature range, which by specification is −40 °C to 70 °C. In order to evaluate the latter, three distinct parameters have been considered,

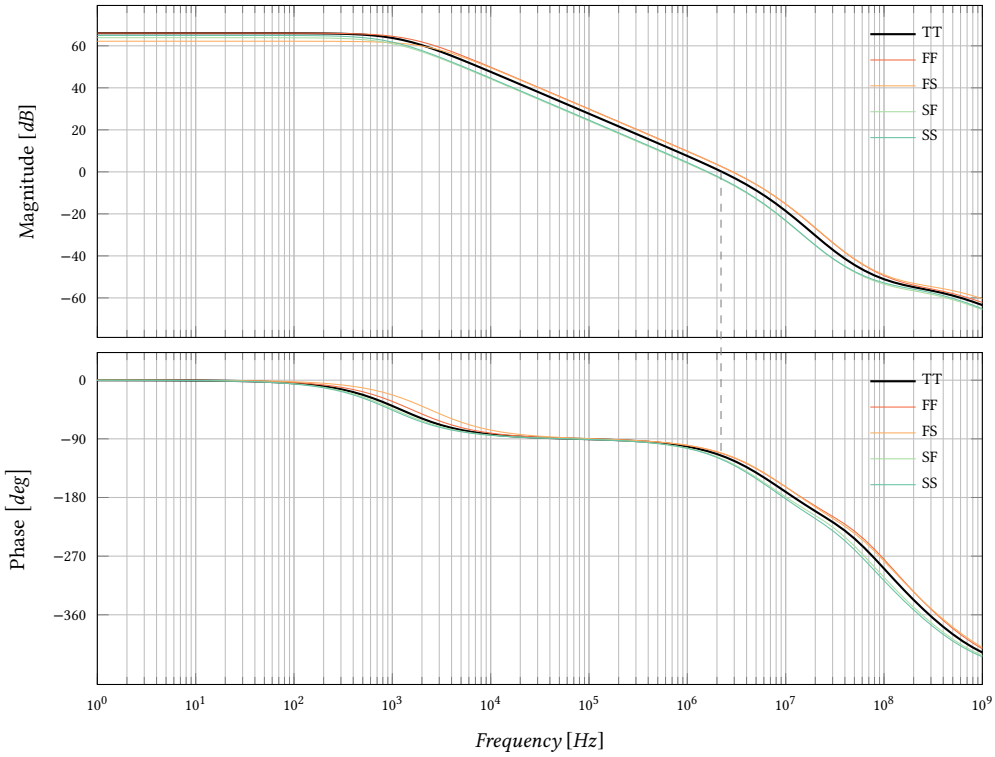


Figure 3.9 – Simulated open-loop magnitude and phase response of the amplifier used in ARCADIA's bandgap for the different MOSFETs corners.

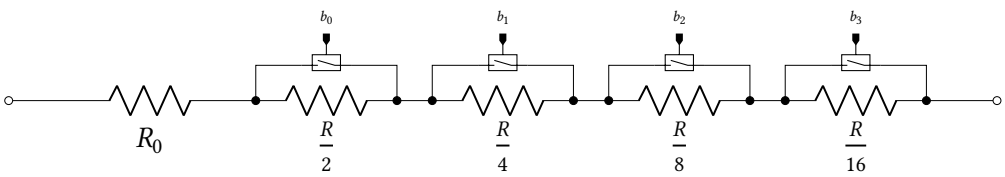


Figure 3.10 – Schematic of the implementation of the bandgap programmable 4-bits resistors. The represented switches have been implemented as pass-gates.

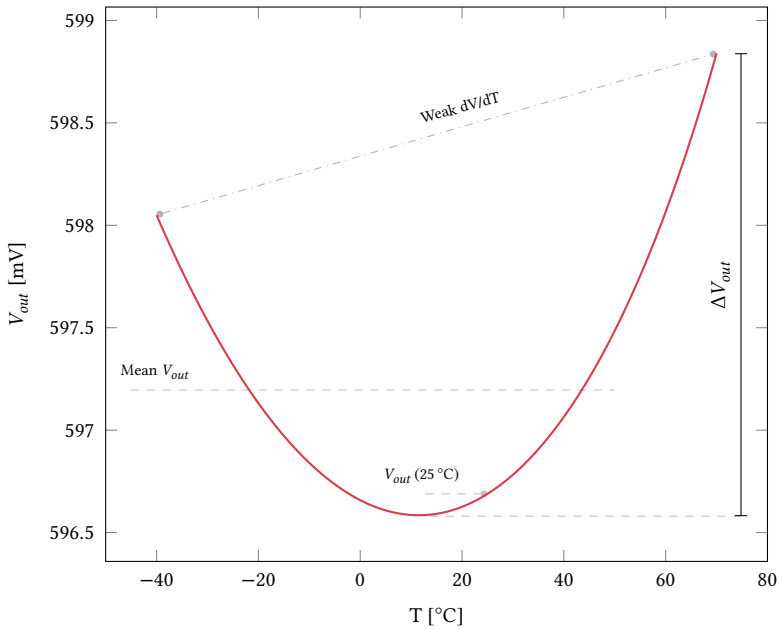


Figure 3.11 – Simulated bandgap reference output showing most of the parameters used in the test-bench to evaluate the circuit performance.

two of which are conceptually the same: in order to evaluate the "droop" of the parabola, the voltage variation across the entire temperature range has been evaluated as

$$\Delta V_{BGR} = MAX\left(V_{BGR}\right) - MIN\left(V_{BGR}\right) \tag{3.29}$$

and, in a very similar fashion, the same can be evaluated as a part-per-million variation by referring it at the mean output or, in this case, the output measured at room temperature and to the entire temperature range, as

$$\Delta_{PPM} = \frac{\Delta V_{BGR}}{\frac{V_{BGR}(25\,°\mathrm{C})}{110\,°\mathrm{C}}} \tag{3.30}$$

The third parameter considered in the test-bench has been the so-called "weak $^{\partial V}/_{\partial T}$", defined as the slope of the line connecting the two points at the extremes of the temperature range. This has been chosen as an indicator of performance given the fact that, in an "ideal" case, the bandgap output would show a parabola whose ends (−40 °C and 70 °C) share the same output voltage, since this is the circumstance in which the parabola features the most contained variation; the absolute value of the parameter can be considered since the target is 0, thus making the minimization of this parameter one of the design goals. A more rigorous definition of the weak $^{\partial V}/_{\partial T}$ can be described as

$$Weak\ \frac{\partial V}{\partial T} = \left|\frac{V_{BGR}\left(-40\,°\mathrm{C}\right) - V_{BGR}\left(70\,°\mathrm{C}\right)}{110\,°\mathrm{C}}\right| \tag{3.31}$$

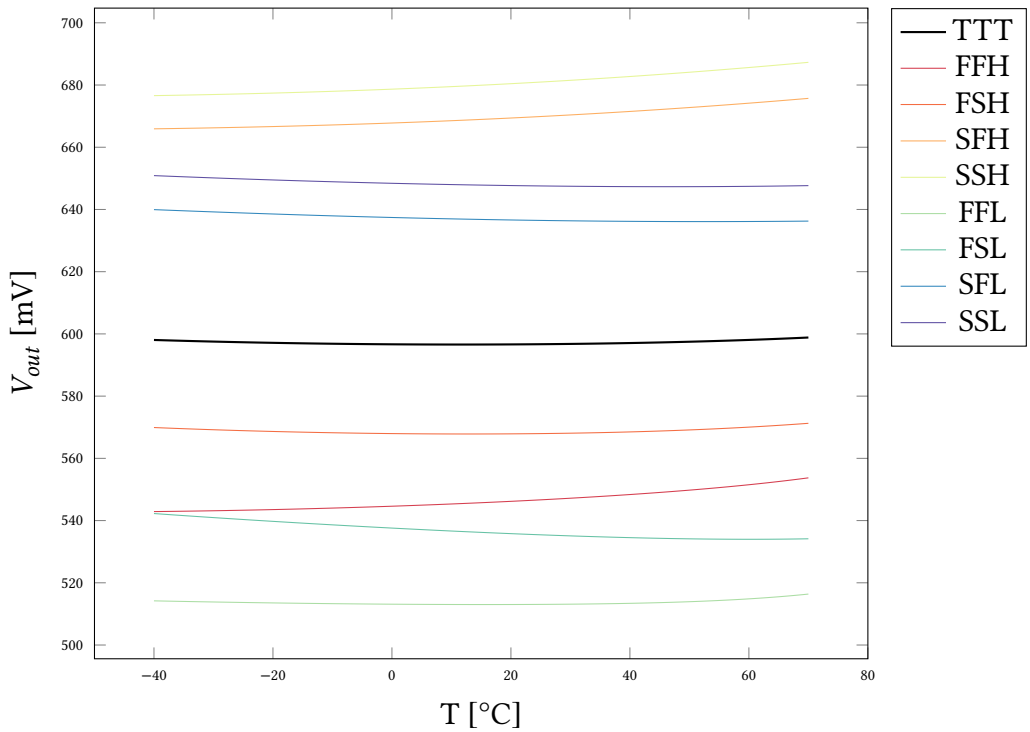


Figure 3.12 – Simulated bandgap output for all corners. The first two letters in the corner name indicate the MOSFET corner, the last one indicates the polyresistor corner, where H stands for high (higher value resistors) and L stands for low (lower value resistors).

Moving onto line regulation, the bandgap reference should be able to reject variations onto the supply voltage, usually considering a variation of ±10 %; the line regulation is defined as

$$LR = \frac{\delta V_{BGR}}{V_{BGR}\,(25\,°\mathrm{C})} \cdot \frac{100}{\Delta V_{DD}} \tag{3.32}$$

expressed in $\%\,\mathrm{V}^{-1}$. Furthermore, the circuit should be able to reject supply noise, behavior that is often summarized with the power supply rejection ratio, or PSRR

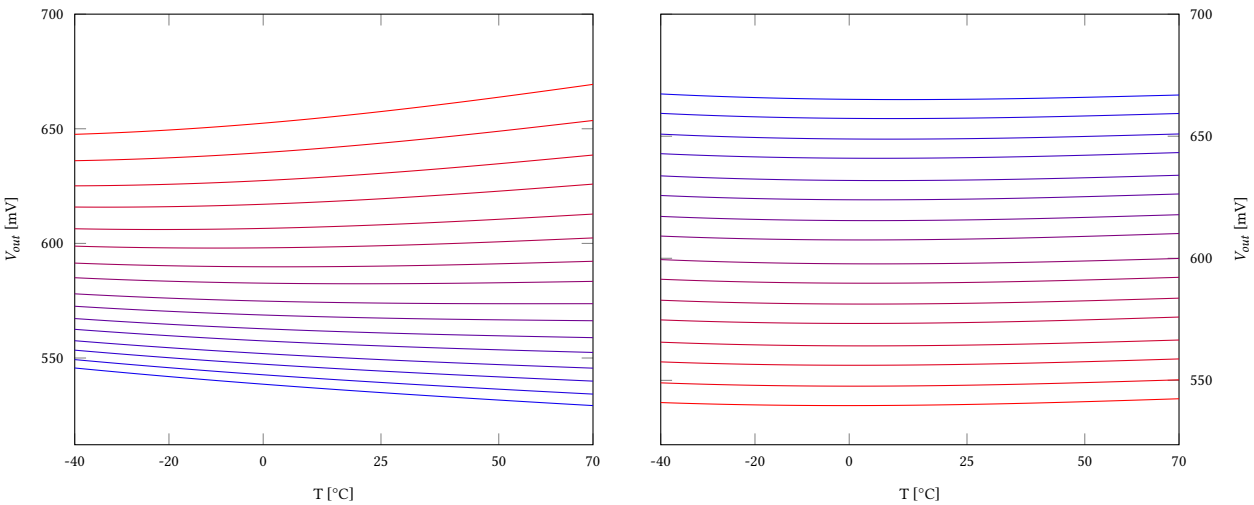
$$PSRR = 20 \log \left(\frac{V_{BGR}}{V_{DD}} \right) \tag{3.33}$$

obtained by doing a frequency sweep onto a superimposed signal on V_{DD} and measuring the response onto V_{BGR}. The temperature-related parameters have been simulated for all MOSFETs corners plus two additional ones to account for polysilicon resistors variation, resulting in a total of 9 simulated cases, shown in fig. 3.12 and summarized in table 3.1.

Figures 3.13a and 3.13b show the results obtained for different configuration bits of R_0 and R_2, respectively. As is to be expected, acting on R_0 changes $\partial V_{BGR}/\partial T$, but also slightly modifies the mean, due to the current flowing through R_2 changing, meaning that R_0 causes a change in both the slope and mean of the output curve; this behavior,

	TTT	FFH	FSH	SFH	SSH	FFL	FSL	SFL	SSL
ΔV_{out} [mV]	2.255	10.840	3.342	9.791	10.700	3.367	8.275	3.845	3.559
Mean V_{out} [mV]	597.2	546.5	568.7	669.6	680.6	513.7	536.8	637.2	648.3
PPM [ppm/°C]	34.36	180.3	54.93	132.9	142.9	59.67	140.5	54.92	49.96
Weak dV/dT[µV °C^{-1}]	-7.18	-98.54	-12.35	-89.01	-97.27	-19.81	73.75	33.54	29.25
Line reg. [% V^{-1}]	2.13	1.82	5.18	8.76	9.31	1.37	4.74	4.76	5.06
Power [µW]	74.4	92.5	94.9	99.9	100.6	60.7	61.8	62.8	63.0

Table 3.1 – Summarized data for the bandgap corner simulations. The first two letters in the corner name indicate the MOSFET corner, the last one indicates the polyresistor corner, where H stands for high (higher value resistors) and L stands for low (lower value resistors).

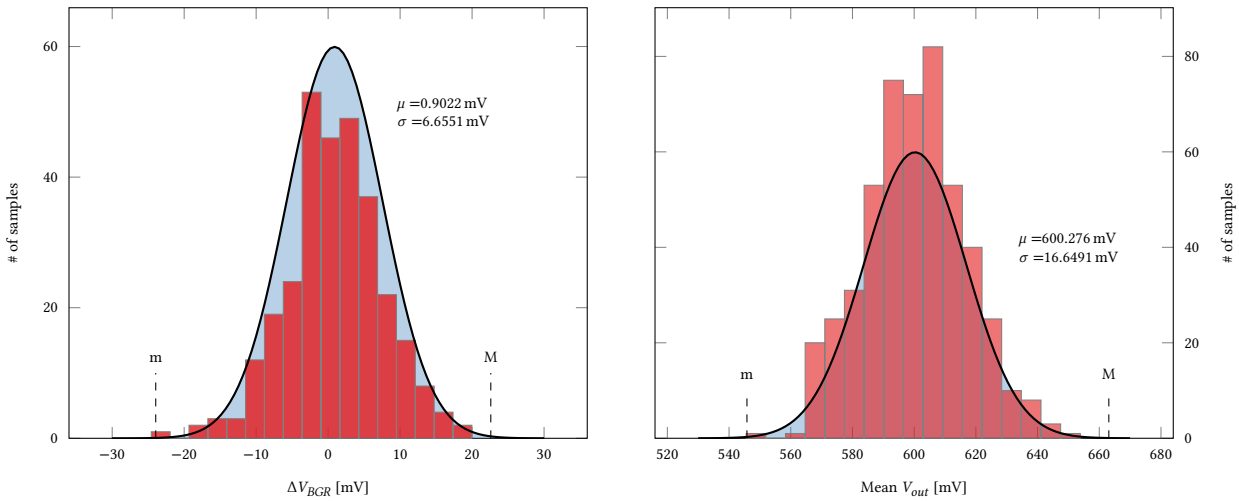


(a) Simulation of the effects of different R_0 values on the output voltage slope, obtained by changing the 4-bit configuration word.

(b) Simulation of the effects of different R_2 values on the output voltage mean, obtained by changing the 4-bit configuration word.

Figure 3.13 – Bandgap voltage reference programmable resistor simulations in the typical corner. Results show the bandgap output for the considered temperature range −40 °C to 70 °C.

even tough undesirable, can be easily fixed by counteracting the output mean change by acting on R_2 configuration bits. The compensation range for resistors R_0 and R_2 has been set by evaluating the process and mismatch variations seen in Monte Carlo simulations: figs. 3.14a and 3.14b show the span of the compensation range referred to Monte Carlo simulation data for R_0, the temperature-slope controlling resistor, and R_2, the mean controlling resistor, respectively. The compensation has been further tested by implementing a simple script using Cadence OCEAN scripting language; given the seed, the script performs a Monte Carlo simulation twice, using the set number of points which, for the data reported here, was 500. The first Monte Carlo run is performed using the standard configuration for both resistors R_0 and R_2, and is used to evaluate the variation in slope and mean; the script then finds the best configuration for both



(a) Monte Carlo process and mismatch variations on the ΔV_{BGR}.

(b) Monte Carlo process and mismatch variations on the bandgap output mean.

Figure 3.14 – Bandgap voltage reference programmable resistor Monte Carlo simulations in the typical corner. Thresholds indicated as m and M indicate the limits of the compensation network.

	This work	Gromov[54]	Cao[55]	Bofuss[67]	Vergine[56]	Traversi[68]
Supply voltage [V]	1.2	1.2	1.2	2.5	1.2	1.2
Operating voltage range [V]	1.08 - 1.32	0.85 - 1.4	0.85 - 1.5	2.5	1.08 - 1.32	1.08 - 1.32
Design reference voltage [mV]	600	405	600	1500	330	400
Temperature coefficient [ppm/°C]	34.4	30.5	15	470	130	16
Temperature range [°C]	−40 °C - 70 °C	0 °C - 80 °C	−40 °C - 125 °C	−40 °C - 200 °C	−40 °C - 80 °C	−40 °C - 100 °C
Line regulation [% V^{-1}]	2.13	2.72	-	-	0.25	4
Power consumption[μW]	70	-	60	50	240	165
Layout area [mm^2]	0.006	0.064	0.056	0.09	0.018	0.028

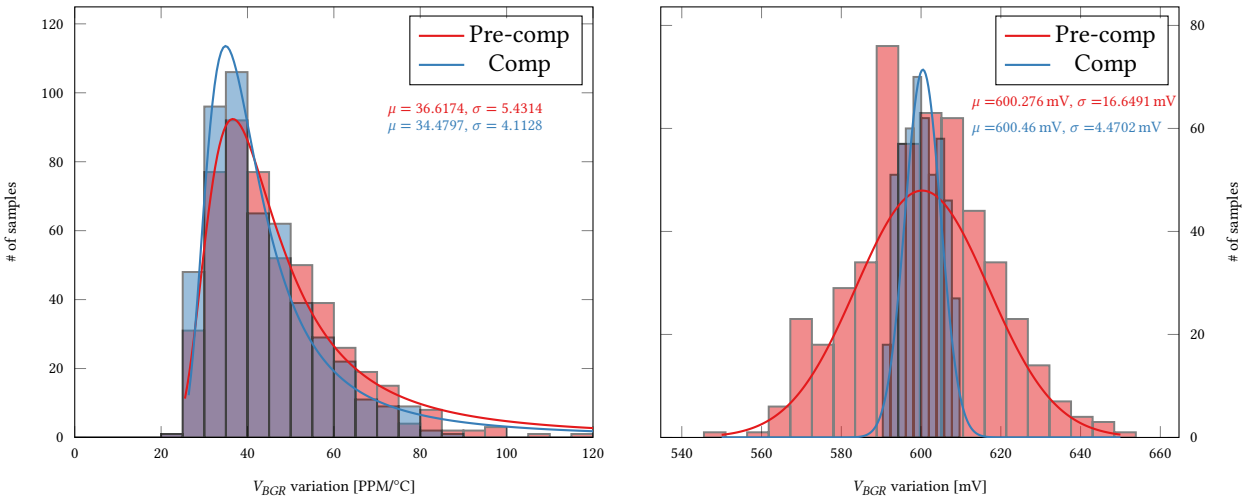
Table 3.2 – Comparison between the presented bandgap and other works in the literature developing references for similar applications.

resistors that compensates for the parameter variation most effectively and, after setting the newly found resistor configuration bits, re-launches the same Monte Carlo run keeping the same seed, obtaining a simulated post-compensation result. A comparison between the pre- and post-compensation simulated results can be seen in figs. 3.15a and 3.15b for the output voltage PPM value, as defined in (3.30), and the mean output voltage, respectively.

Lastly, the simulated PSRR is reported in fig. 3.16, showing a value of −38 dB up until 1 kHz. A comparison with other bandgap voltage references for similar applications found in the literature is reported in table 3.2.

3.2.5 Reference current from bandgap core

Additionally, ARCADIA's main demonstrator has the need of a block generating a reference current for the front-end bias networks: after some considerations, the most



(a) Compensation results for bandgap circuit, temperature related variations in the range −40 °C to 70 °C. Red distribution shows uncompensated results, blue distribution shows compensated ones.

(b) Compensation results for bandgap circuit, output mean value. Red distribution shows uncompensated results, blue distribution shows compensated ones.

Figure 3.15 – Simulated Monte Carlo results showing the trimming resistors action for compensating the temperature slope and the mean value errors of the bandgap output.

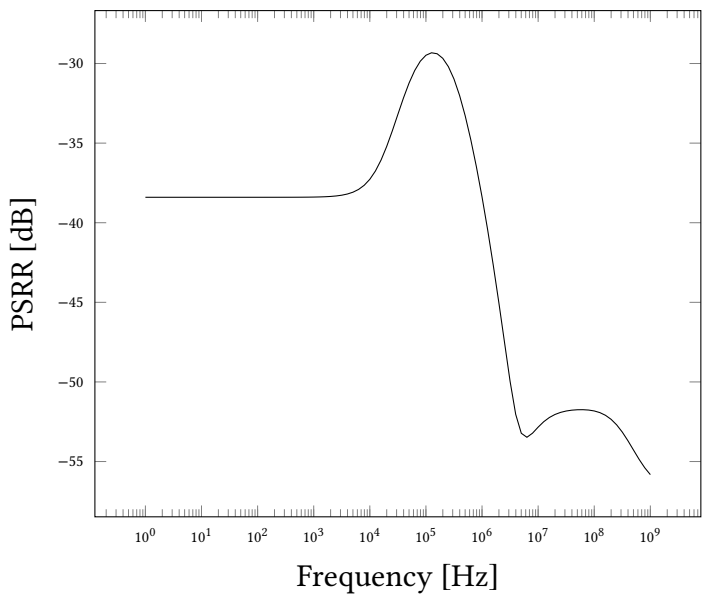


Figure 3.16 – Simulated bandgap reference PSRR in the typical corner.

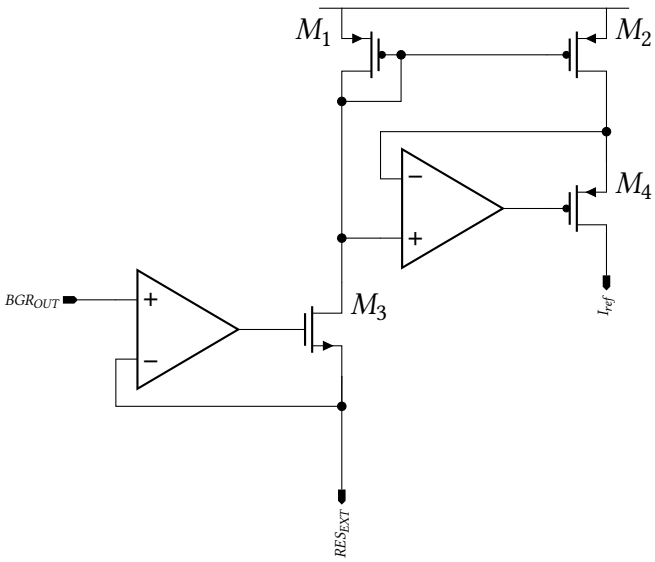
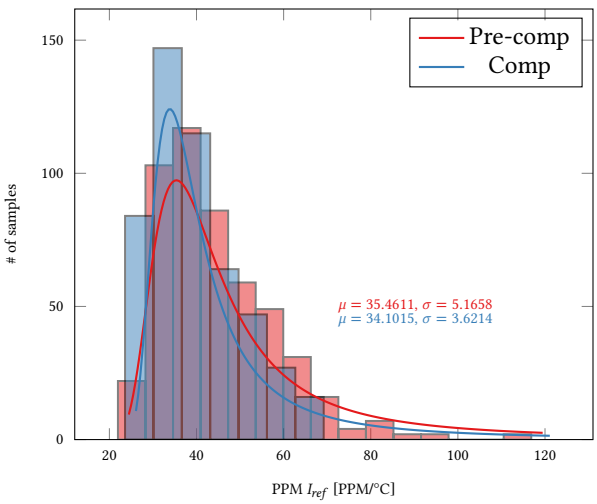


Figure 3.17 – Scheme of the implemented block used to generate a 2.4 µA reference current starting from the bandgap output. Pin RES_{EXT} is routed to the padrig and will be connected to a low-temperature coefficient resistor placed on the board hosting the chip.

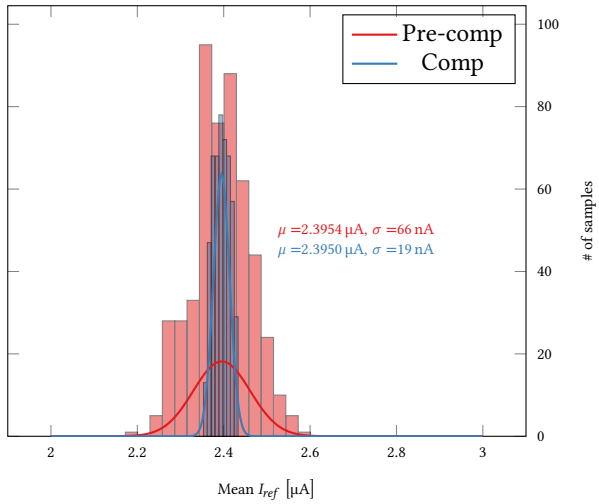
straightforward way to develop this circuit has been deemed the addition of a block cascaded to the discussed bandgap reference, in order to use it to generate the needed reference current. The specifications require a current of 2.4 µA, with a maximum variation of ±10 %. In the first implementation, the current is generated from the buffered bandgap output with the use of an external resistor; the complete scheme is reported in fig. 3.17. The overall concept is quite basic: the bandgap output is buffered onto an external resistor, generating the current and then mirroring it onto an internal pin used by the bias generation block to mirror the current to other parts of the chip. The chosen external resistor[VIII] is a 249 kΩ (E96 series) with a 5 ppm/°C temperature coefficient and a 0.05 % tolerance to minimize both the absolute and the temperature-related errors.

The errors in the generated current strongly depend on the accuracy of the bandgap reference voltage, thus the same script-based compensated simulation can be run simulating both blocks to evaluate the effect of process and mismatch variations onto the generated current and the effectiveness of the compensation. The simulation results are shown in fig. 3.18; the simulation does not take into account the variation introduced by the external resistor, but taking into account a $\pm 5\sigma$ variation from the compensated results and adding the contribution from the ±0.05 % resistor tolerance, the error is ±3.7 %.

[VIII] https://docs.rs-online.com/01b4/0900766b814887ca.pdf



(a) Compensation results for bandgap current reference circuit, temperature-related variations in the range −40 °C to 70 °C. Red distribution shows uncompensated results, blue distribution shows compensated ones.



(b) Compensation results for bandgap current reference circuit, mean value. Red distribution shows uncompensated results, blue distribution shows compensated ones.

Figure 3.18 – Simulated Monte Carlo results showing the trimming resistors action for compensating the temperature slope and the mean value errors of the reference current.

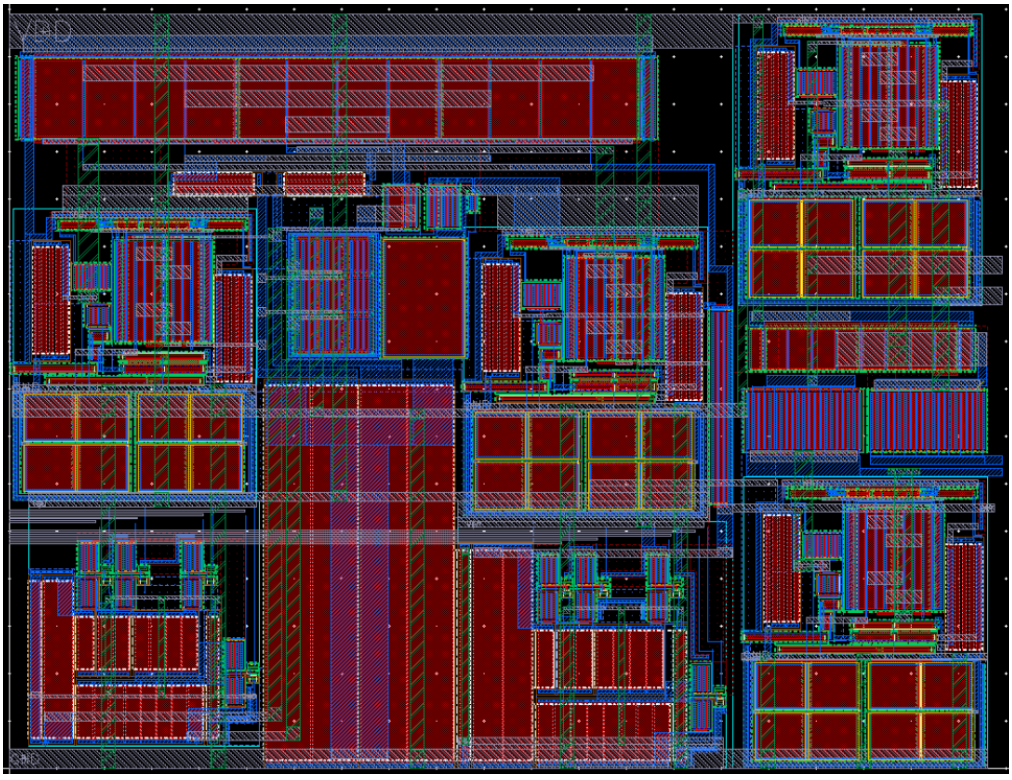


Figure 3.19 – Layout of ARCADIA bandgap voltage reference block, developed in a 110 nm technology. The layout, comprised of a bandgap circuit and a reference current generation circuit, has an area of 100 × 80 µm^2.

3.2.6 Non-linear compensation approach

The following section describes a different approach at the bandgap trimming, which was later set aside and not implemented in the shown bandgap circuit. The section will merely present a theoretical approach at a different compensation scheme and a possible, albeit quite basic, implementation; although it was initially envisioned for the bandgap compensation circuit, the following could be used in other applications where trimming is involved.

3.2.6.1 Linear (classic) compensation approach

Generalizing, the role of a compensation network is to reduce the dispersion of the parameter of interest, where this cannot be further reduced by design or layout choices. To abstract the problem, this analysis will consider the final goal of a compensation network to be the minimization of the compensated distribution standard deviation. As an underlying hypothesis, all distributions taken into account are considered to be normal distributions. It is the author's opinion that considering only normal distributions offers a good starting point for the analysis and that the hypothesis does not narrow down the window of applications too much, given the central limit theorem, while keeping the calculations simple. Given the starting hypothesis, the problem can be formulated as having the to-be-compensated parameter normally distributed as $\mathcal{N}_0\left(\mu_0, \sigma_0\right)$ and trying to obtain a post-compensation parameter distributed as $\mathcal{N}_C\left(\mu_C, \sigma_C\right)$, keeping $\mu_C = \mu_0$ while lowering σ_C as much as possible or, equivalently, minimizing $\frac{\sigma_C}{\sigma_0}$. In order to actually perform the trimming operation, a certain number of points 2^n is chosen, each one indexed using an n bit word; these compensation points offer, when chosen by programming the desired word, the possibility of applying a fixed offset to the parameter of interest, effectively shifting the distribution mean. The first case covered will be that of equally spaced compensation points, used in other works [69] and in the already discussed bandgap reference circuit, meaning that:

$$P_C(k) = (k - 2^{n-1}) \cdot \Delta C \ \mid \ k \in \{0, 1, \dots, 2^n - 1\} \tag{3.34}$$

Such points are chosen so that the middle point $P_C(2^{n-1})$ lies exactly on the original distribution mean, making $P_C(2^{n-1})$ the standard configuration. Taking the point definition one step further, the distance between two compensation points ΔC can also be related to the original distribution standard deviation σ_0 via a multiplicative factor ϑ so that $\Delta C = \frac{\vartheta \cdot \sigma_0}{2^n}$; the parameter ϑ influences the compensation dynamic of the system, or how much of the original distribution is covered by the compensation points. Referring back to the trimmable BGR architecture, such a compensation can be imple-

mented as already shown in fig. 3.10, but the same principle is achieved, with a different implementation, in most trimming DACs. Considering the distribution depicted in fig. 3.20 and assuming an error-less compensation[IX], choosing a certain compensation point $P_C(k)$ is equivalent to shifting the mean value of the entire distribution by a value equal to $\left(2^{n-1} - k\right) \cdot \Delta C$ so that $P_C(k)$ referred to the compensated (shifted) distribution falls exactly onto the original distribution mean μ_0, as demonstrated by the dashed distribution in fig. 3.20. The choice to use the $k - th$ compensation point is taken when the particular sample from the original distribution that needs to be trimmed falls into the interval $\left\{P_k - \frac{\Delta C}{2}, P_k + \frac{\Delta C}{2}\right\}$. This leads to a specific sample of the distribution being compensated perfectly (falling exactly on the mean of the distribution) if said sample is exactly on a compensation point; otherwise, the error after compensation will be[X]

$$\varepsilon_C \in \left\{ -\frac{\Delta C}{2} \leq \varepsilon_C \leq +\frac{\Delta C}{2} \right\} \tag{3.35}$$

The compensated standard deviation σ_C can be written as [70]:

$$\begin{aligned} \sigma_C = \Bigg\{ \int_{-\infty}^{+\infty} \Bigg[\sum_{i=1}^{2^n-2} & \mathscr{H}\left(\frac{\Delta C}{2} - |x|\right) \cdot \mathscr{N}\left[\mu_0 + (i - 2^{n-1}) \cdot \Delta C, \sigma_0\right] \\ & + \mathscr{H}\left(\frac{\Delta C}{2} - x\right) \cdot \mathscr{N}\left(\mu_0 + (-2^{n-1} \cdot \Delta C), \sigma_0\right) \\ & + \mathscr{H}\left(x + \frac{\Delta C}{2}\right) \cdot \mathscr{N}\left(\mu_0 + (2^{n-1} - 1 \cdot \Delta C), \sigma_0\right) \Bigg] \cdot x^2 \, dx \Bigg\}^{\frac{1}{2}} \end{aligned} \tag{3.36}$$

where $\mathscr{H}\left(f(x)\right)$ is the Heaviside step function. It can also be demonstrated [70] that the ratio $\frac{\sigma_C}{\sigma_0}$ is independent of σ_0; thus, this analysis will be conducted assuming $\sigma_0 = 1$, so any σ_C reported can be considered as the $\frac{\sigma_C}{\sigma_0}$ ratio. Considering now that

$$\begin{cases} \int_{-\infty}^{+\infty} \mathscr{H}\left(\gamma - |x|\right) \cdot f(x)\, dx = \int_{-\gamma}^{+\gamma} f(x)\, dx \\ \int_{-\infty}^{+\infty} \mathscr{H}\left(\gamma - x\right) \cdot f(x)\, dx = \int_{-\infty}^{+\gamma} f(x)\, dx \\ \int_{-\infty}^{+\infty} \mathscr{H}\left(\gamma + x\right) \cdot f(x)\, dx = \int_{-\gamma}^{+\infty} f(x)\, dx \end{cases}$$

and defining then the list of points $m_C(k)$ as the mid-points between two compensation points P_C which, as already said, delineate the limit of the boundaries according to

[IX]Meaning that using any compensation point influences the parameter distribution mean μ_C only, and does not effect other distribution parameters, such as its standard deviation or its shape.

[X]Note that this is a simplification, since it does not take into account the tails of the distribution.

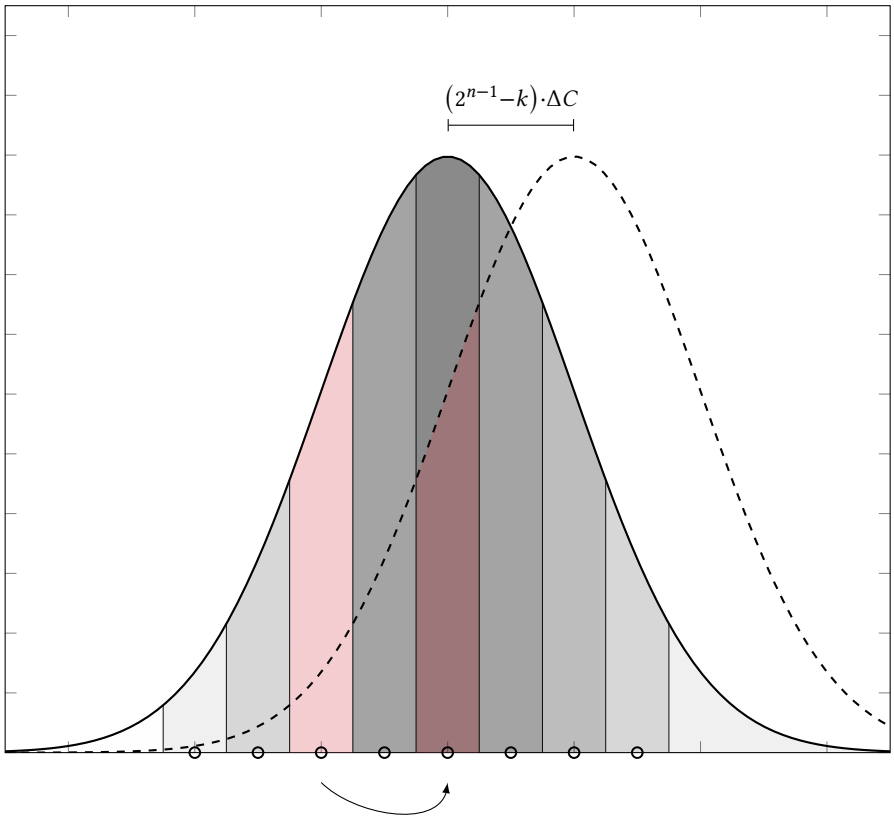


Figure 3.20 – Compensation effect and original distribution $\mathcal{N}(\mu_0, \sigma_0)$. Round markers indicate the position of the compensation points P_C and black lines mark the edges of the compensation intervals, referred to in this text as mid-points m_C. If a particular case falls into a certain colored area, then the compensation point related to that particular area should be used. The dashed distribution shows the ideal mean-shift that the compensation accomplishes: choosing a compensation point and performing the compensation procedure corresponds to re-distributing the points that previously were located in the distribution tail around the desired value.

which a compensation point is chosen over the next

$$m_C(k) = \frac{P_C(k) + P_C(k+1)}{2} \quad | \quad k \in \{0, 1, \ldots, 2^n - 2\} \tag{3.37}$$

and finally taking into account that, using simple substitution rules,

$$\int_{-\gamma}^{+\gamma} \mathcal{N}(\mu_0 + \varphi, \sigma_0) \cdot x^2 \, dx = \int_{\varphi-\gamma}^{\varphi+\gamma} \mathcal{N}(\mu_0, \sigma_0) \cdot (w - \varphi)^2 \, dw \tag{3.38}$$

then (3.36) can be re-written as

$$\sigma_C = \left\{ \sum_{i=1}^{2^n-2} \left[\int_{m_C(i-1)}^{m_C(i)} \mathcal{N}\left[\mu_0, \sigma_0\right] \cdot (x - P_C(i))^2 \, dx \right] + \int_{-\infty}^{m_C(0)} \mathcal{N}\left[\mu_0, \sigma_0\right] \cdot (x - P_C(0))^2 \, dx + \int_{m_C(2^n-2)}^{+\infty} \mathcal{N}\left[\mu_0, \sigma_0\right] \cdot (x - P_C(2^n - 1))^2 \, dx \right\}^{\frac{1}{2}} \tag{3.39}$$

Computing (3.39) instead of (3.36) using numerical methods should be faster due to the absence of the Heaviside step functions. The value of σ_C has then been numerically computed using Matlab, for different number of bits, influencing the number of compensation points, and for different values of ϑ, influencing the compensation dynamic. The results are showed as solid lines in fig. 3.21. The results show behaviors that can be intuitively explained: first, as it is to be expected, a higher number of bits corresponds to a better compensation, due to the increased resolution in the trimming or, more rigorously, due to a decrease of ΔC and its impact on the error shown in (3.35). Secondly, and possibly more interestingly, the function exhibits a minimum as a function of ϑ, that can be intuitively explained by considering that a low value of ϑ corresponds to the compensation circuit not being able to cover the entire distribution, and a higher value of theta corresponds to a decrease in the compensation resolution (higher ΔC).

3.2.6.2 Non-linear compensation approach

Having evaluated the performance, in terms of σ_C, that can be obtained in the case of equally spaced compensation points, this section will focus on the comparison of said performance against the ones achievable with non-equally spaced compensation points. The idea sparks from the intuitive concept that, being the original distribution a gaussian, more refined compensation points should be needed near the mean value of the distribution, where most of the cases will fall; this can be paid with a coarser compensation (points more distant one from another) along the tails of the distribution. For this reason, the function used to sample the position of the non-equally spaced compensation points has been arbitrarily chosen as a cubic function. In order to obtain the expression of the compensated standard deviation $\sigma_{C,ne}$, the position of the compensation points must be redefined. When defining the equally spaced compensation points, the compensation dynamic $\vartheta \cdot \sigma_0$ was defined so that $P_C(2^n - 1) + \frac{\Delta C}{2} - \left(P_C(0) - \frac{\Delta C}{2}\right) = \vartheta\sigma_0$, which is, with

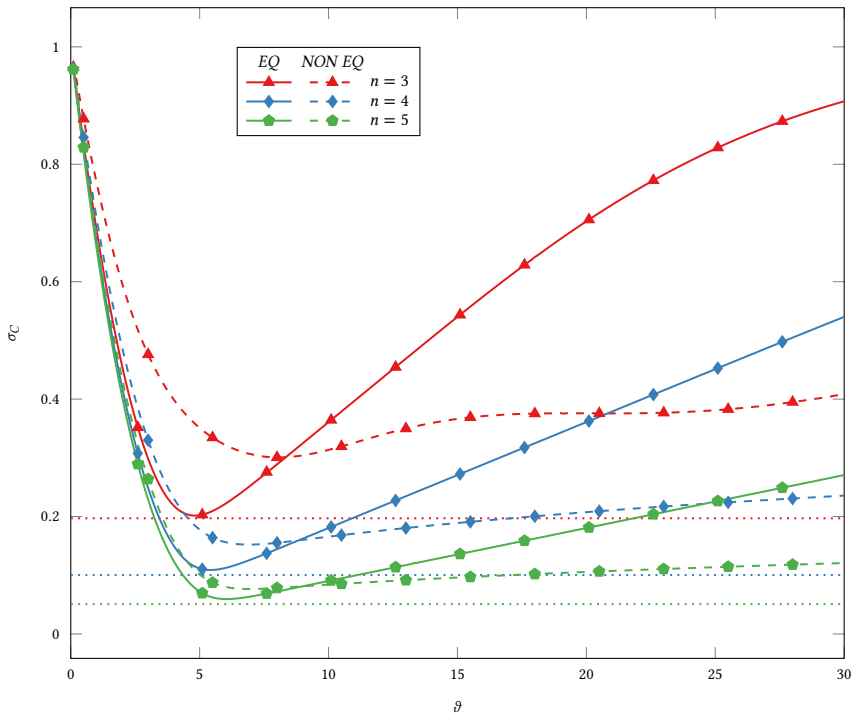


Figure 3.21 – Normalized compensated distribution standard deviation versus compensation dynamic, for 3, 4 and 5 bits schemes and for equally spaced and non equally spaced (cubic) point distribution. The dotted lines indicate the minimum σ_C obtained as the result of a minimization problem.

a little abuse of notation, the same as the distance between $m_C(-1)$[XI] and $m_C(2^n-1)$[XII]. In order to keep consistency in the comparison, the compensation dynamic for the cubic-sampled points has been defined in the same way. As such, imposing that:

$$\begin{aligned} m_{C,ne}(2^n-1) - m_{C,ne}(-1) &= \vartheta\sigma_0 \\ \frac{P_{C,ne}(2^n)+P_{C,ne}(2^n-1)}{2} - \frac{P_{C,ne}(0)+P_{C,ne}(-1)}{2} &= \vartheta\sigma_0 \end{aligned} \tag{3.40}$$

the list in the case of cubic-sampled (non-equally spaced) compensation points is:

$$P_{C,ne}(k) = \frac{\vartheta\sigma_0 \cdot 2^{2-n}}{2^{2n}+6} \cdot (k-2^{n-1})^3 \;\;|\;\; k \in \{0,1,\dots,2^n-1\} \tag{3.41}$$

Having defined the compensation points and using the introduced notation for the midpoints, where $m_C(k)$ is the midpoint between $P_C(k), P_C(k+1)$, the compensated

[XI] $m_C(-1)$ is an abuse of notation of (3.37), intended as the middle point between $\{P_C(-1), P_C(0)\}$.
[XII] Following the streak of notation abuse, $m_C(2^n-1)$ is intended as the middle point between $\{P_C(2^n-1), P_C(2^n)\}$.

standard deviation $\sigma_{C,ne}$ equals

$$\sigma_{C,ne} = \Bigg\{ \int_{-\infty}^{m_{C,ne}(0)} \mathcal{N}\left[\mu_0, \sigma_0\right] \cdot (x - P_{C,ne}(0))^2 \, dx + \sum_{i=1}^{2^n-2} \left[\int_{m_{C,ne}(i-1)}^{m_{C,ne}(i)} \mathcal{N}\left[\mu_0, \sigma_0\right] \cdot (x - P_{C,ne}(i))^2 \, dx \right] + \int_{m_{C,ne}(2^n-2)}^{+\infty} \mathcal{N}\left[\mu_0, \sigma_0\right] \cdot (x - P_{C,ne}(2^n - 1))^2 \, dx \Bigg\}^{\frac{1}{2}} \tag{3.42}$$

Again, the compensated standard deviation has been numerically calculated for different lengths of the compensation word and for different values of ϑ; results are shown in fig. 3.21. The calculations shown in this section have been performed starting from a distribution $\mathcal{N}_0(0, 1)$; the equation showing that the ratio $\frac{\sigma_C}{\sigma_0}$ is independent of σ_0 presented in [70] can be generalized so that the same can be applied to the results presented here. Considering a generalization of the σ_C expression in (3.36):

$$\sigma_C = \Bigg\{ \int_{-\infty}^{+\infty} \frac{1}{\sigma_0\sqrt{2\pi}} \cdot \left[\mathscr{H}\left(f_{h,left}(x)\right) \cdot \exp\left[-\frac{(x + f_m(0))^2}{2\sigma_0^2}\right] + \sum_{i=1}^{2^n-2} \mathscr{H}\left(f_{h,central}(x, i)\right) \cdot \exp\left[-\frac{(x + f_m(i))^2}{2\sigma_0^2}\right] + \mathscr{H}\left(f_{h,right}(x)\right) \cdot \exp\left[-\frac{(x + f_m(2^n - 1))^2}{2\sigma_0^2}\right]\right] \cdot x^2 \, dx \Bigg\}^{\frac{1}{2}} \tag{3.43}$$

it can be shown that, following the same procedure highlighted in [70], $\frac{\sigma_C}{\sigma_0}$ is independent of σ_0 as long as a factor σ_0 can be extracted for the $f_{h,central}$,$f_{h,left}$,$f_{h,right}$ and f_m functions. In the case of the non-equally spaced compensation points distributed along a cubic

function, as described in this chapter, it holds true that:

$$
\begin{aligned}
f_{h,central}(x,i) =& \frac{P_{C,ne}(i+1) - P_{C,ne}(i-1)}{4} \\
& - \left| x - \frac{P_{C,ne}(i+1) - 2P_{C,ne}(i) + P_{C,ne}(i-1)}{4} \right| \\
f_{h,left}(x,i) =& \frac{P_{C,ne}(1) - P_{C,ne}(0)}{2} - x \\
f_{h,right}(x,i) =& \frac{P_{C,ne}(2^n-1) - P_{C,ne}(2^n-2)}{2} + x \\
f_m(i) =& P_{C,ne}(i)
\end{aligned}
\tag{3.44}
$$

3.2.6.3 Compensation points position as a minimization problem

Lastly, this problem may be approached as an optimization problem with 2^n parameters, those being the position of the compensation points; the cost function is the expression of σ_C. The problem can be defined as:

$$
\min_{P_{C,opt} \in \mathbb{R}^{2^n}} \sigma_{C,opt}\left(P_{C,opt}\right) \tag{3.45}
$$

bound to

$$
\begin{cases}
P_{C,opt}(2^{n-1}) = \mu_0 \\
P_{C,opt}(2^{n-1}-k) = P_{C,opt}(2^{n-1}+k), \forall k \in \mathbb{N} \mid 0 \leq k < 2^{n-1} \\
P_{C,opt}(k) < P_{C,opt}(k+1), \forall k \in \mathbb{N} \mid 0 \leq k < 2^n - 1
\end{cases}
$$

so that the middle compensation point lies exactly on the original distribution mean, the points are mirrored with respect to the middle one and are monotonically increasing. The so defined problem has been solved using the `fmincon` MATLAB built-in function, which implements an interior-point algorithm meant to solve convex optimization problems. The compensated standard deviation $\sigma_{C,opt}$ obtained with this method has been reported in fig. 3.21 as a dotted line indicating the minimum $\sigma_{C,opt}$ obtained; this representation of $\sigma_{C,opt}$ has dropped the ϑ dependence because of the ambiguity of its definition, that requires knowing the position of $P_{C,opt}(-1)$ and $P_{C,opt}(2^n)$ in order to obtain $m_{C,opt}(-1)$, $m_{C,opt}(2^n-1)$ and consequently, following the same concept highlighted in (3.40), define the ϑ obtained using this method. A somewhat loose definition of ϑ can be obtained by fitting the PC, opt points in order to extract the values of $P_{C,opt}(-1)$ and $P_{C,opt}(2^n)$. The point distribution obtained using this minimization method somewhat follows the initial idea that the compensation should be precise

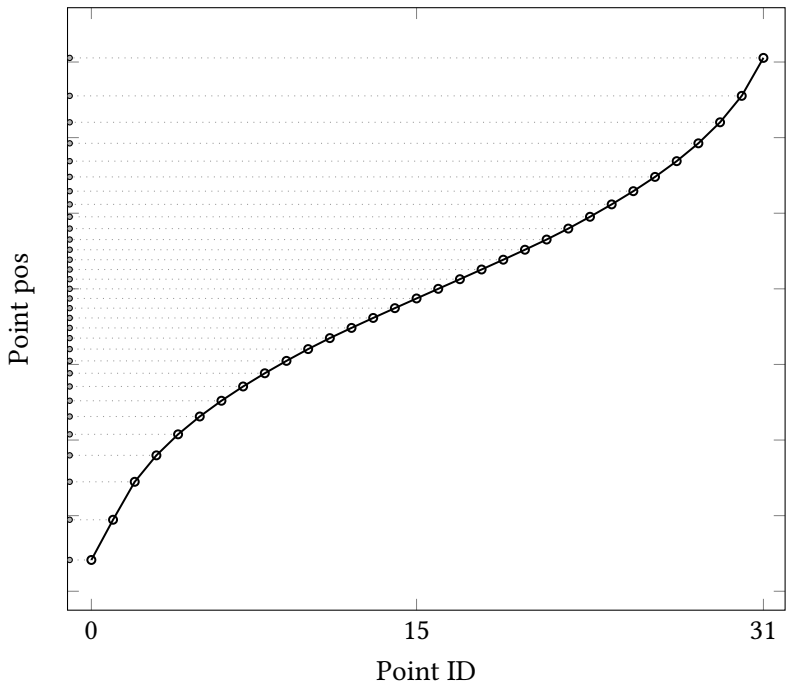


Figure 3.22 – Distribution of points that satisfy the numerical minimization problem described in section 3.2.6.3.

towards the mean value of the gaussian, having the compensation points more bunched up, and coarser onto the distribution tails, where the points are more spaced apart, as can be seen from the plot in fig. 3.22. Trying to fit the curve does not yield definitive results: a third-degree polynomial fit results in an adjusted R^2 value of 0.9982 and an RMSE of 6.9%; an hyperbolic sine curve, even though seemingly a good possible candidate, has similar results, with an adjusted R^2 value of 0.9988 and a RMSE of 5.8%. Using the third degree polynomial fit function, ϑ can be re-introduced, by using the relative position of the compensation points and scaling them accordingly in order to obtain $\vartheta = P_{C,opt}(2^n) - P_{C,opt}(-1)$. A comparison between the previous results and the third degree polynomial fit is reported in fig. 3.23; the fit function obtains slightly better results, the main being a slower increase of the compensated σ_C for $\vartheta > \vartheta_{optimal}$.

3.2.6.4 Basic implementation of the non-linear compensation

A very crude implementation of such a compensation can be obtained by recreating the non-linear increments using a series of resistors and controlling them using an appropriate digital circuit, such as a binary to thermometric converter; a 3-bits example is reported in fig. 3.24. The obvious drawback to such an implementation is the area: the addition, with respect to the implementation presented in fig. 3.10, of more resistors (which scales exponentially with the number of bits, given the $2^n - 1$ resistors needed by this implementation) should not lead to a dramatic increase of the area, under

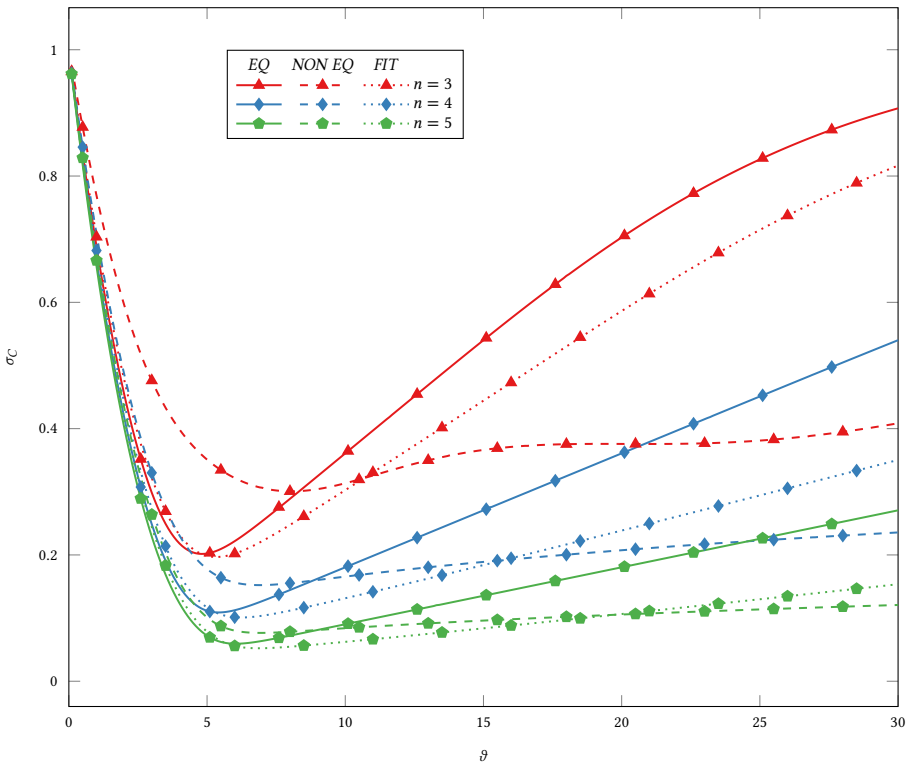


Figure 3.23 – Compensated distribution standard deviation versus compensation dynamic, for 3, 4 and 5 bits schemes and for equally spaced and non equally spaced (cubic) point distribution. The dotted lines indicate the σ_C obtained using the points result of the minimization problem.

the assumption that, given that the sum of resistors in the implementation shown in fig. 3.10 and fig. 3.24 differ by $\frac{R}{16}$. The addition of the binary to thermometric converter, whose complexity scales non-linearly with the number of bits, is a different matter; the complexity in the digital circuitry for the non-linear implementation is considerably higher than the linear one, making this the main cause of the area increase.

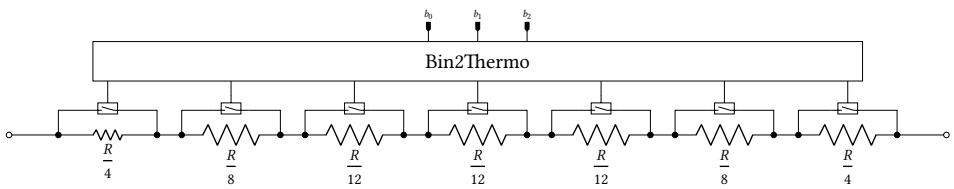


Figure 3.24 – Possible implementation of a resistor network implementing third-degree polynomial-like increments. The bin2thermo block is a binary to thermometric converter that allows the progressive selection of the resistors.

4

Custom LVDS Transmitter and Receiver

Nowadays, it is hard to find circuits that accomplish a complex task in a purely analog fashion; said circuits are accompanied, often controlled by digital blocks: for example, considering outbound information, data is very often transmitted after an analog-to-digital conversion or, considering inbound information, a chip internal blocks are often configurable and programmed by sending binary data. The presence of data leads to the need for transmitting and receiving it, thus the need of specific blocks meant for this exact purpose. Two blocks consitute the bread and butter of this topic: a receiver and a transmitter. A receiver is, roughly speaking, an amplifier meant to regenerate the incoming signal, and a transmitter is a circuit capable of driving the transmission line load, based on the input signal[I]. Many transmission standards have been developed, such as RS-485, RS-422 or SCSI, all with their strong points and drawbacks. The standard that is of interest for this chapter is the Low Voltage Differential Signaling (LVDS) one, defined by the Telecommunication Industry Association (TIA) in the ANSI/TIA EIA-644[71] standard[II] as a general purpose, high speed (recommended frequency ranges given by the original standard were $655\,\mathrm{Mbit\,s^{-1}}$ to $1.923\,\mathrm{Gbit\,s^{-1}}$) and low power standard. The LVDS standard defines a common mode of 1.2 V and a voltage swing of $\pm 350\,\mathrm{mV}$, making this standard unfeasible for technologies with V_{DD} lower or equal to 1.2 V. This chapter will present the design, simulation and measurements performed on a test vehicle including a custom sub-LVDS transmitter and receiver pair developed in a 110 nm technology.

[I] Although here is assumed that the transmitter and receiver are meant for digital signals, the concept of line driver and line receiver can be found in the analog world as well.

[II] LVDS is also standardized by the IEEE, in their SCI (1596-1992) standard [72].

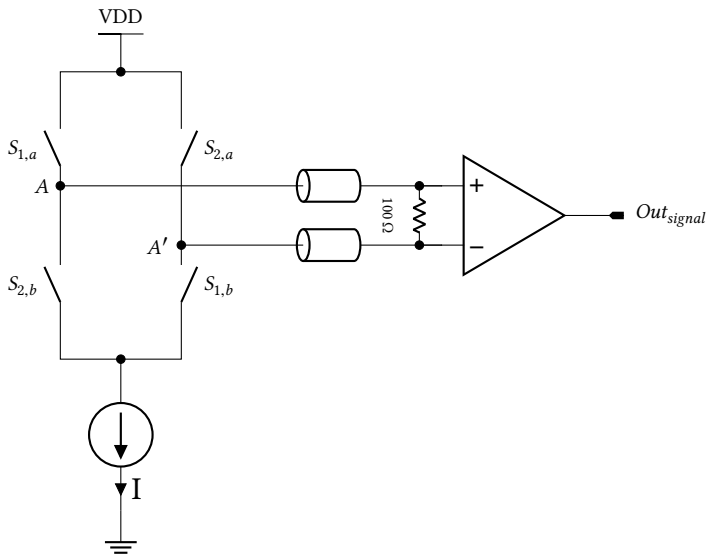


Figure 4.1 – Simple schematic of an LVDS driver/receiver pair. The current I is usually 3.5 mA to obtain a $\pm$350 mV voltage swing on the 100 Ω termination resistor.

4.1 Problems and performance evaluation of a line driver/receiver

As it may have been guessed from the introduction, LVDS is not a standard defining a single ended transmission, but a differential one. A crude implementation of a transmitter/receiver pair is reported in fig. 4.1. The line driver can be seen as an alternating current generator driving a current onto the termination resistor; in the reported schematic, the input signal would control switch pairs alternately: closing switches $S_{1,a}$ and $S_{1,b}$ forces a current flowing from node A to node A', through the 100 Ω termination resistor, generating a 350 mV signal, in the case of the standard LVDS's 3.5 mA current, at the amplifier inputs. In the opposite case, i.e when the input signal is the opposite of the previous case, switches $S_{2,a}$ and $S_{2,b}$ are closed, making the current I flow in the opposite direction, thus generating the same signal with opposite polarity at the amplifier inputs. These signals are modulated on top of a DC component, known as common mode; in the LVDS standard, the common mode value is 1.2 volt, leading to signals similar to the one reported in fig. 4.2. The receiver deals with the conversion from differential to a single-ended signal. The advantages of such a transmission are found mainly in the ability of rejecting common-mode interference: a spurious variation that affects both signals in the same way is compensated for when looking at the differential signal; obviously this comes at the disadvantage of having to use two transmission lines instead of one.

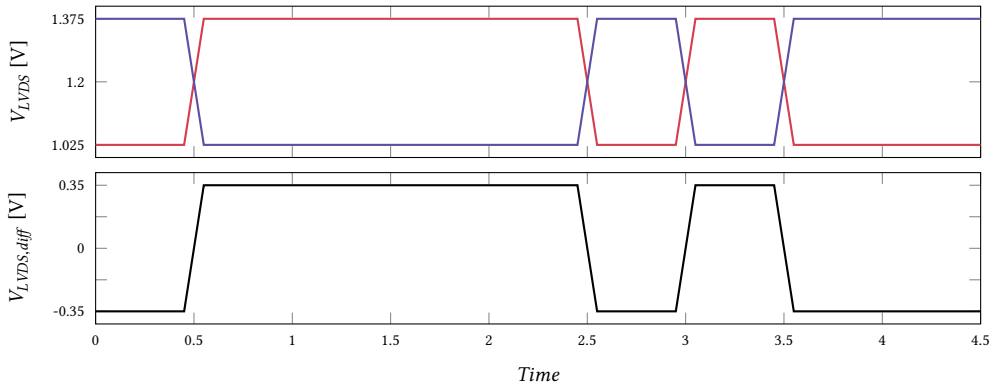


Figure 4.2 – Example of a LVDS signal. Plotted are the single ended (i.e. referred to ground) voltages seen at the two receiver inputs, followed by the differential signal.

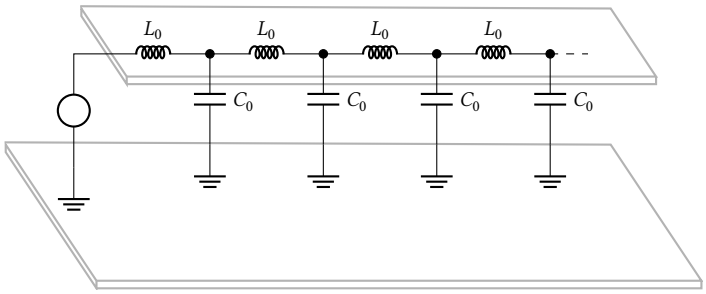


Figure 4.3 – Transmission line model as a distributed LC ladder.

4.1.1 Transmission line and termination mismatch effects

A good deal of the problems linked to the design of a transmitter and a receiver pair lie in the domain of the connection between the former and the latter; unfortunately, treating the transmission medium using a lumped parameter approach and analyzing it using Kirchhoff's laws can hold up to certain frequencies, but not quite so for higher ones. By definition, a wavelength λ is the distance traveled by the wave to accumulate a 360 deg phase shift; a rule of thumb considers a phase shift of 36 deg to be negligible, and thus a distance smaller than $\frac{\lambda}{10}$ is usually considered electrically short [73]. An electrically short distance can be treated using Kirchhoff's laws and lumped-circuit analysis; for higher frequencies, or higher distances, the connection should be considered as a transmission line. Transmission lines are analyzed as distributed contributions, usually as a distributed LC ladder, as depicted in fig. 4.3, where L_0 and C_0 are the inductance and capacitance per unit length, respectively. Note that this model does not take into account any line resistance, there are models based off π-networks that take into account series resistance and dielectric leakage[74]. A fundamental property of the transmission line can be defined by envisioning an input step with amplitude V_0 at one end, resulting in a current step with amplitude I_0, from which the characteristic impedance of the line

can be defined as $Z_0 = \frac{V_0}{I_0}$ and it can be shown[75] to be equal to (for a lossless line)

$$Z_0 = \sqrt{\frac{L}{C}} \tag{4.1}$$

which is a real value without any reactance, due to the line being lossless, and can be seen as the instantaneous input impedance of the transmission line. Now, assuming no discontinuity in the connection, the characteristic impedance will stay constant, thus the voltage step V_0 envisioned before will travel up to the load, inducing a I_0 current in the line. At the load, different situations can occur: in the case of a perfectly matched load, i.e. $R_L = Z_0$, the relationship $\frac{V_0}{I_0} = R_L$ is valid and the transient happens on the load as expected. A different case can be experienced in the case of a mismatched load impedance: in the case of an open circuit at the load, $R_L = \infty$, thus the current I_L should be zero; once the I_0 "forward" step reaches the load, an opposite step must propagate towards the initial point so that $\frac{V_0}{R_L} = I_L = 0$ and no current flows on the "load". The current traveling back generates a voltage step of amplitude V_0, since the characteristic impedance of the line is still Z_0, raising the voltage on the line at $2V_0$. Same goes for the completely opposite case, when the load is a short circuit, $R_L = 0$, thus the voltage should be zero; once again, when the voltage step reaches the load, a reflection of amplitude $-V_0$ must propagate towards the beginning of the line, lowering the voltage along the line to 0. These examples are obviously extremes, but the concept remains the same: a mismatch on the load (i.e., the so-called termination) causes a reflection, which is seen as an additional variation on top of the voltage read by the receiver, and can lead to bit errors. By the same phenomena, the reflection itself can cause further reflections, called secondary reflections, if the transmitter output impedance is not matched. The reflected wave can be defined with respect to the forward traveling one in terms of a reflection coefficient Γ[76]:

$$\Gamma_{RX} = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{4.2}$$

so, given the initial forward voltage step V_0, the reflected wave will be of amplitude $\Gamma_{RX} V_0$; if the transmitter impedance is not matched, this will cause a secondary reflection with a reflection coefficient of

$$\Gamma_{TX} = \frac{Z_S - Z_0}{Z_S + Z_0} \tag{4.3}$$

leading to a behavior depicted in the scheme reported in fig. 4.4 which will appear as noise on the line itself.

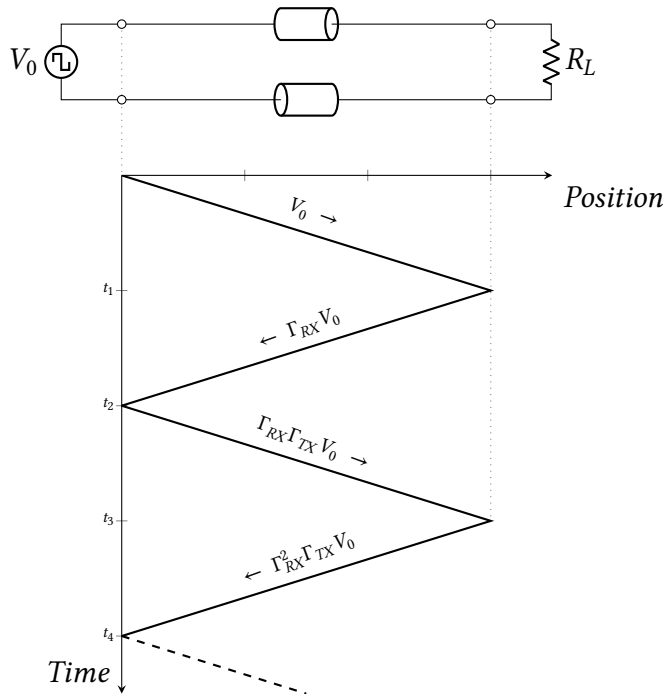


Figure 4.4 – Schematic representation of the timings involved in primary and secondary reflections when dealing with unmatched terminations on a lossless line.

4.1.2 Bandwidth related effects

When analyzing lossy lines, the presence of a distributed series resistance combined with the distributed capacitance leads to a low-pass filtering effect, lowering the bandwidth of the system. This filtering leads to slower edges, due to the longer fall and rise times linked to the capacitance charge and discharge cycles; an example for a periodic signal is shown in fig. 4.5a. Obviously the effect is undesirable, but may be easily circumvented by adjusting the threshold which separates the logical 1 and logical 0 values. When dealing with random data instead, the behavior is a bit different: long sequences of a single bit allow the line capacitance to fully charge, while it may occur that single bits do not provide enough time to completely charge. This behavior leads to a moving signal mean, making the previous solution of adjusting the discriminator level between a 1 and a 0 unfeasible or less effective. In any case, the threshold will be set in such a way that some cases will be barely over threshold, leading to a signal very sensitive to noise. This effect is called Inter-Symbol Interference (ISI) due to the fact that a previous symbol, or even better, the previous recent history of symbols, interferes with the current one. An example is reported in fig. 4.5b. The effects of ISI can be somewhat mitigated with the use of DC-balanced encoding schemes[77], such as 8b/10b. In most cases, 8b/10b encoding is used to avoid long ”runs”, i.e. long repetitions of the same symbol, to facilitate the clock recovery and to avoid localized (time-wise) DC components. This

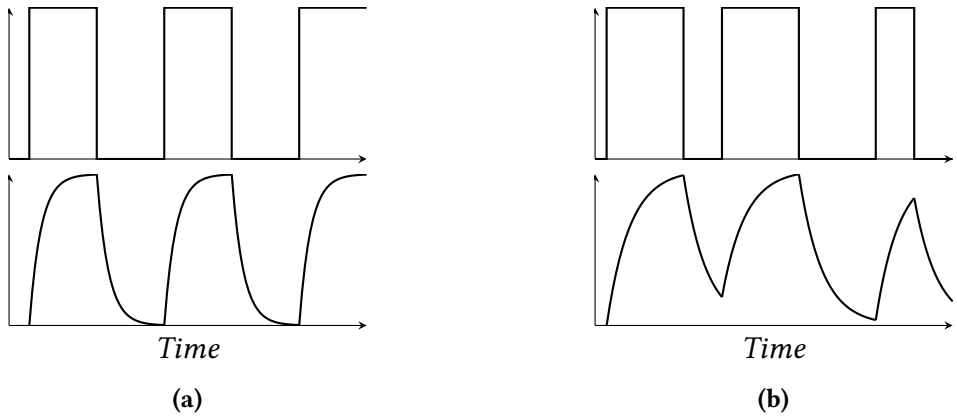


Figure 4.5 – Effects of bandwidth limitation in transmission lines for periodic data **(a)** and random data **(b)**. The "movements" in the signal mean due to the bandwidth limitations shown in **(b)** lead to Inter-Symbol Interference (ISI).

is accomplished by forcing a loosely DC-balanced code, adding data that carries no information but aims at keeping the mean value of the signal stable. In the case of 8b/10b encoding for example, the 8 bit input is subdivided into two blocks constituted by 5 and 3 bits each, which are encoded using 6 and 4 bits, respectively. The encoded 6-bit and 4-bit blocks will either have a difference between the number of ones and the number of zeros of ± 2, or 0 (meaning in the block there is an equal number of ones and zeros). The transmission is characterized by a state, also called running disparity, due to the fact that it indicates the current surplus of ones in the transmission; this disparity is initialized at a value of -1. Depending on the value of the running disparity, a block is chosen so that the difference of ones and zeroes is opposite in sign: for example, the first block will see a running disparity of -1, since that is the initialization value, thus a block with a surplus of ones (+2) will be chosen. In this case the running disparity is flipped, so that the next block chosen will have more zeros than ones (-2). As mentioned before, some blocks are encoded with an even number of ones and zeros: if such a block is chosen in the encoding, the running disparity stays the same. Following this encoding, the running disparity will alternate between -1 and +1, leading to a DC balance on the line. ARCADIA's main demonstrator will implement an 8b/10b encoding. Another interesting consideration can be found in the analysis of a NonReturn-to-Zero (NRZ) random binary transmission. NRZ encoding, as the name suggests, transmits data by encoding a binary zero with one line level and a binary one with another line level, as opposed to Return-to-Zero (RZ) encoding, where a symbol is always followed by a zero, making the line "return" to that level after every symbol[III]; examples of NRZ and RZ encoding are shown in fig. 4.6. The case of NRZ encoded data is relevant to this

[III]The usage of RZ encoding allows for easier clock recovery due to the fact that the signal exhibits a frequency equal to the bit rate.

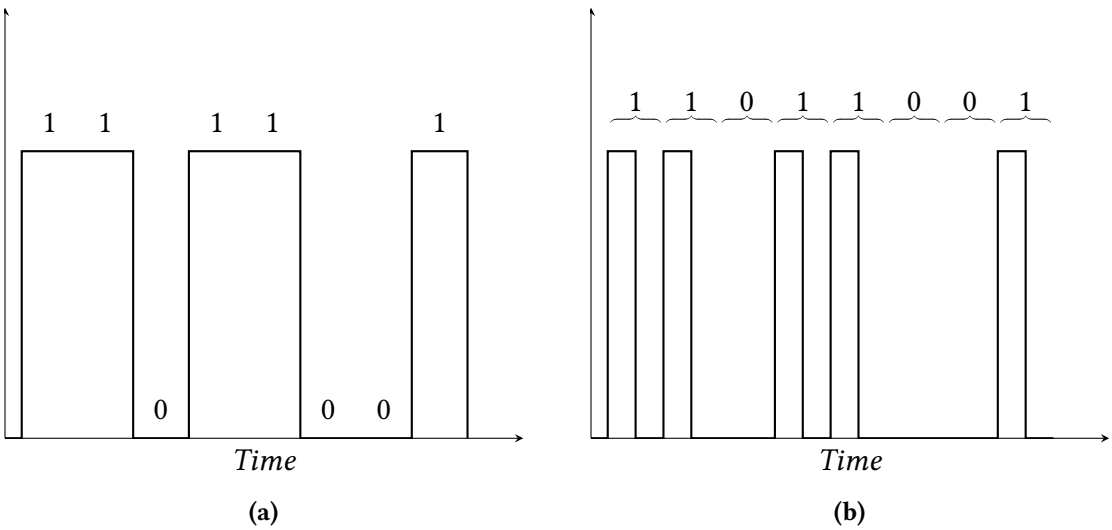


Figure 4.6 – Example of binary data stream encoded as NonReturn-to-Zero **(a)** and Return-to-Zero **(b)** showing byte 0xD9.

chapter since it is usually the choice for LVDS systems. The starting assumption for this analysis is that the data stream ones and zeros have the same probability and there is no correlation between the symbol transmitted at time T_k and the one at T_{k+m} with $m \in \{\mathbb{Z}^*\}$. Starting from this assumption and from what has been already discussed in section 2.2 (see (2.38)), the autocorrelation function can be written as:

$$R_{XX} = \lim_{T\to\infty} \frac{1}{2 \cdot T} \cdot \int_{-T}^{T} rect\left(\frac{t}{T_b}\right) \cdot rect\left(\frac{t}{T_b} + \tau\right) dt \tag{4.4}$$

where T_b is the bit time; since the convolution of a rectangle function with itself is a triangular function, the autocorrelation will be

$$R_{XX} = \begin{cases} A_b^2\left(1 - \frac{|t|}{T_b}\right), & |t| < T_b. \\ 0, & |t| \geq T_b. \end{cases} \tag{4.5}$$

where A_b is the amplitude of the rectangle function. From what was showed in (2.44), the power spectral density can be obtained by the inverse Fourier transform of the autocorrelation function: for a triangular function, the inverse Fourier transform is a cardinal sine function, thus the power spectral density of a random NRZ signal is

$$S_{NRZ}\left(f\right) = A_b^2 T_b \operatorname{sinc}^2\left(\pi f T_b\right) \tag{4.6}$$

the plot of which, with frequencies normalized to $\frac{1}{T_b}$, is shown in fig. 4.7. If an essential bandwidth is defined for the signal[78] as the one containing 95 % of the signal power, it

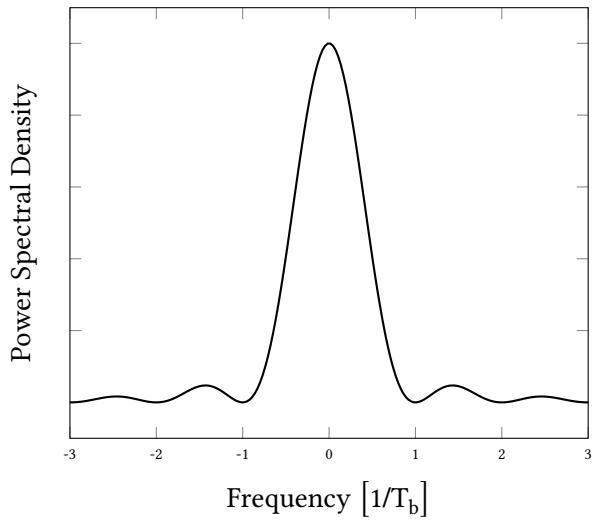


Figure 4.7 – Plot of the bilateral power spectral density for a NRZ signal, showing frequencies normalized to $\frac{1}{T_b}$.

can be concluded that the essential bandwidth for a NRZ random data stream is $\sim \frac{0.75}{T_b}$.

4.1.3 Noise and jitter effects

A close-to-ideal digital transmission signal has a trapezoidal shape, where the bit is sampled in the middle of the flats and the rise and fall times are finite and equal to each other. In reality, this shape is degraded by the presence of noise, which can come from different sources, such as receiver or transmitter noise, external interferences or line parasitics. In the scope of this analysis, the time instant of particular interest is the one at which the bit value is sampled, i.e. in the middle of the flat part of the ideal trapezoidal shape and, since the noise is a stochastic process, the analysis should be carried out using its probability density function. Given these premises, the probability density function of the ideal signal is constituted by only two values, each having a probability of ½[IV]; regarding the probability density function of the noise, it should be safe to assume that it is normally distributed. The aim of this analysis is to consider the probability density function of the sum of the two presented variables; given that they are independent, the probability density function of their sum is the convolution of their pdfs, resulting in the distribution represented in fig. 4.8. A first analysis can be carried out by looking at the separation between the two distributions, if hopefully present: in that separation range a threshold can be chosen to effectively distinguish between a logical zero and a logical one, without having the noise flip a bit. If the separation is not present, the two distribution tails overlap, leading to ones falsely interpreted as zeros and vice-versa. Later in this chapter the analysis tool of the eye diagram will

[IV] This assumes that the ones and zeros have equal probability of appearing in the data stream.

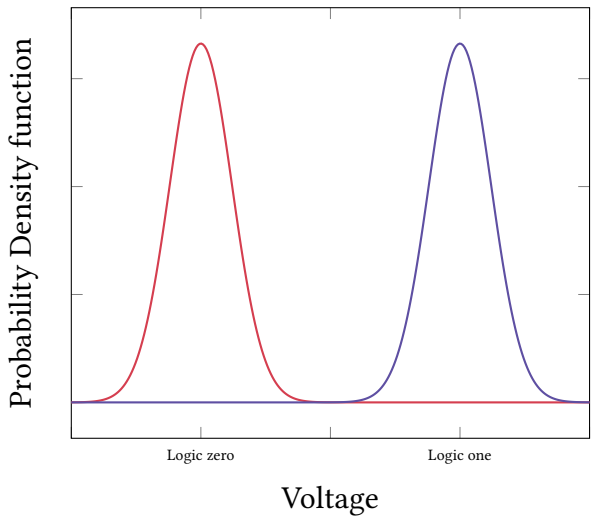


Figure 4.8 – Plot of the probability density function for a noisy digital signal, obtained as the convolution between the pdf of the ideal signal and the pdf of the noise, assumed normally distributed.

be presented, as an effective mean of evaluating the two discussed distributions; some oscilloscopes also provide a way to extract them from the eye diagram itself. So far the analysis has been focused on the amplitude noise, which is the effect of noise onto the line voltages, and that could lead to errors in the bit evaluation, but noise can have an effect on the timing of the signal, leading to the so-called jitter. The jitter can be defined as the short-term variation of the significant instants of a timing signal from their ideal position in time[79], but can also be defined as the deviation of an oscillator period[80]; this chapter is concerned with the first definition. These phenomena come primarily from noise but other sources, such as system or data-related phenomena, can result in jitter. System phenomena are related to the analog nature of the transmitter and receiver circuits: some examples can be cross-talk and impedance mismatch induced reflections; data-related phenomena are, as the name suggests, born out of patterns in the data which affect the jitter, such as the previously described Inter-Symbol Interference or pseudo-random bit sequences periodicity. More generally, the "total" jitter is often modeled as born of various components[81], also represented in the block diagram reported in fig. 4.9:

Deterministic Jitter (DJ). Jitter correlated to some other variable in the system and thus repeatable and predictable; its amplitude is bounded.

Data-Dependent Jitter (DDJ). Sub-category of the Deterministic Jitter correlated to data patterns, such as the jitter caused by Inter-Symbol Interference. The fact that it is caused by the data itself makes it predictable and thus falls into the deterministic jitter category.

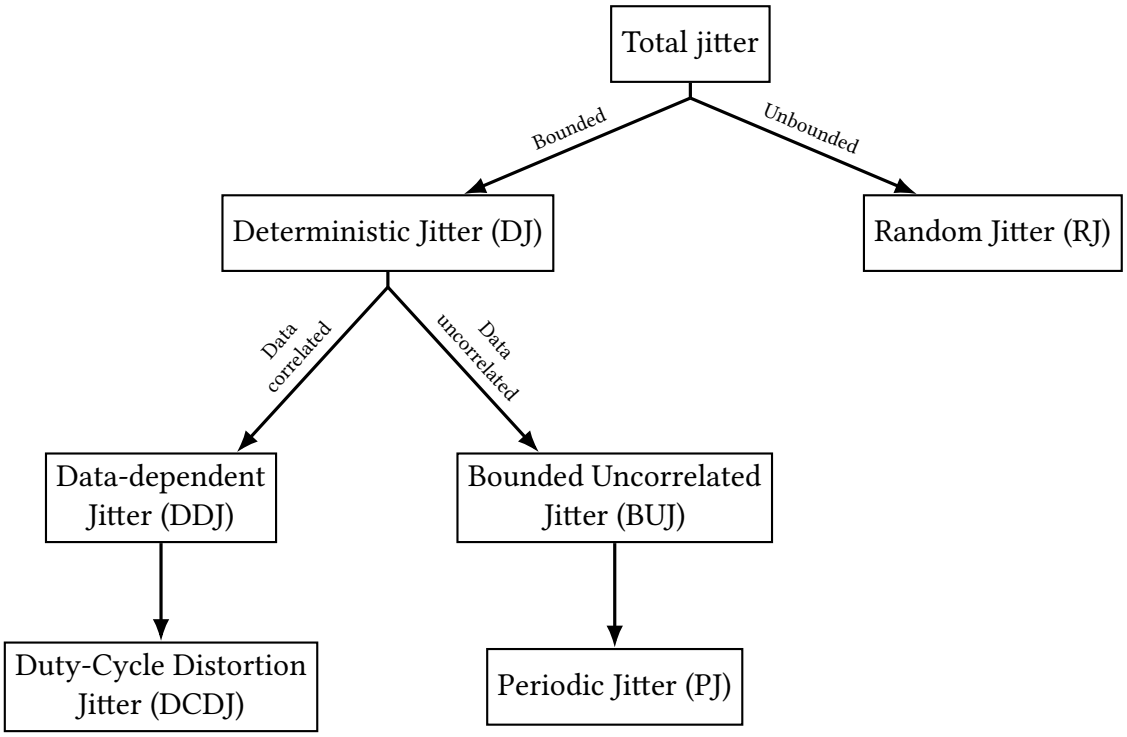


Figure 4.9 – Block diagram of the decomposition of Total Jitter into its basic components.

Duty-Cycle Distortion Jitter (DCDJ). The most straightforward case of DDJ, where the jitter is introduced only by the bit preceding the signal edge that shows jitter; can be caused for example by asymmetries in rise/fall times of the signal.

Bounded Uncorrelated Jitter (BUJ). The counterpart of the DDJ, encompasses the jitter caused by phenomena unrelated to the data patterns, such as crosstalk from other lines or signal reflections due to impedance mismatches on the line.

Periodic Jitter(PJ). A sub-category of the BUJ that shows a periodic component; one example could be the jitter induced by a switching power supply used to power the circuit under analysis: the supply lines will have a periodic ripple due to the switching action of the supply, leading to a periodic change in the jitter.

Random Jitter (RJ). Jitter caused by stochastic processes such as noise.

The two concepts introduced in this section can be related one to the other, in the sense that since noise can cause jitter, an expression of the jitter as a function of the noise amplitude can be provided[82]; considering the noise as the sum of two components as described before, one trapezoidal ideal shape and the random noise as an additive component,

$$V_{signal}(t) = V_{ideal}(t) + V_N(t) \tag{4.7}$$

considering V_{thr} as the voltage crossing threshold at the receiver between two bit values and $V_{ideal}\left(t_{ideal}\right) = V_{thr}$, i.e. defining t_{ideal} as the time at which the ideal shape crosses the threshold, the jitter is

$$\mathcal{J}itter = t_{ideal} - \frac{V_N\left(t_{ideal}\right)}{\frac{\partial V_{ideal}}{\partial t}} = t_{ideal} - \frac{V_N\left(t_{ideal}\right)}{SR} \tag{4.8}$$

where SR is the slew rate. This analysis assumes that the noise keeps the value $V_N\left(t_{ideal}\right)$ in the vicinity of the t_{ideal} time instant. This result entails that an higher slew rate (or lower noise) will lead to a lower random jitter.

4.1.4 Eye diagrams

The chapter has so far discussed some indicators of performance in a digital transmission, but these are somewhat difficult to see in a standard plot or timing diagram; evaluation of characteristics such as jitter require the acquisition and comparison of features repeated in the signal at different times. The most used tool to analyze these features is the eye diagram. The basic idea of an eye diagram is that, exploiting the periodicity of the signal itself, more than one "time-window" of the signal can be shown in the same plot: by sub-dividing the signal into intervals, usually 2-bits wide, then superimposing all the obtained intervals into a single plot, all the characteristics such as ISI, jitter, noise, crossing points can be extracted from this single diagram. An example is presented in fig. 4.10: the figure shows a random, quite clean string of data, which is sub-divided into 3 segments, as reported below. The three segments are then plotted on the same graph, aligning all the left (and right, since the intervals are all the same width) edges, resulting in the eye diagram seen on the right. Usually, instruments capable of plotting eye diagrams use a colormap to indicate how many lines are superimposed on top of each other. The jitter and its distribution can be inferred from the spread and color of the lines around, and the same goes for the distribution of noise around the two logic levels; some instruments are able to compute and show these distributions automatically. A simple eye diagram highlighting some of the useful parameters that can be extracted from it is reported in fig. 4.11.

4.2 Custom LVDS transmitter design and simulations

The scheme of the implemented transmitter is shown in fig. 4.12. As previously shown in this chapter, the transmitter is composed of a H-bridge structure formed by MOSFETs M_1,M_2, M_3 and M_4 meant to drive a current with two different polarities into the

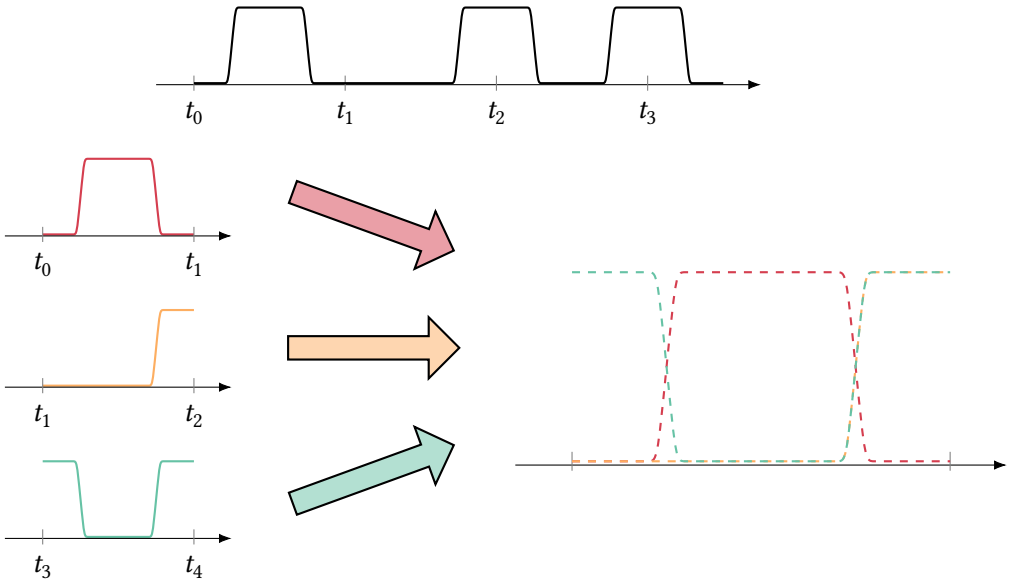


Figure 4.10 – Example of the principle used to construct an eye diagram. Folded signals in the eye diagram are plotted dashed for the sole purpose of showing superimposed signals.

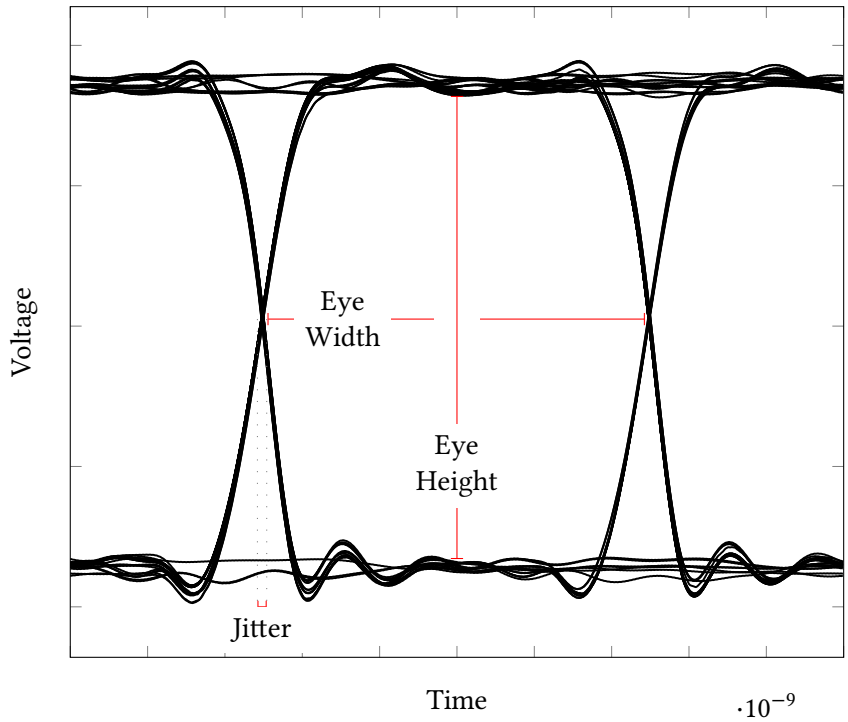


Figure 4.11 – Schematic representation of the parameters that can be extracted from an eye diagram.

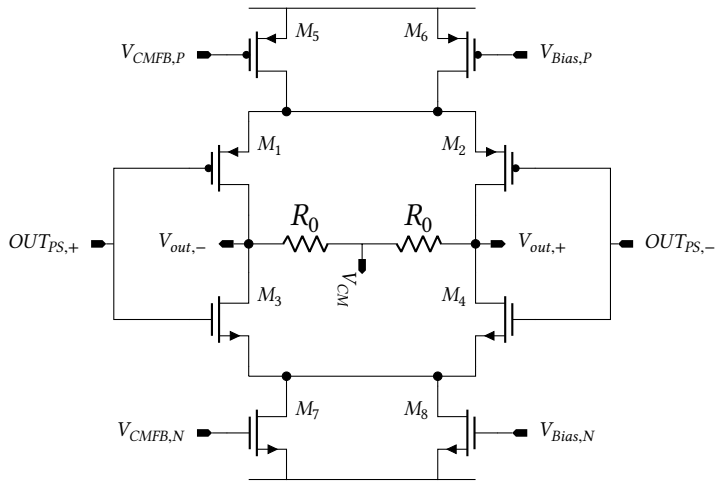


Figure 4.12 – Schematic of the implemented custom LVDS transmitter used in the ARCADIA main demonstrator.

termination resistor, by controlling pairs M_1/M_3 and M_2/M_4 with two counter-phase signals; the two signals are generated starting from a CMOS input data stream using the phase splitter that will be described later in this section. These MOSFETs have been chosen to have a large area to ensure a low $R_{ds,ON}$. The driver current is controlled by MOSFETs M_5/M_7 and M_6/M_8: in order to satisfy different constraints of power consumption and speed from the different application targeted by ARCADIA, the transmitter power consumption has been designed to be programmable; this has led to the addition of an enable signal, and to the possibility of having a selectable driver current. More will be discussed together with the bias blocks, but for what concerns the bridge structure itself, the pair M_5/M_7 is used to generate a dynamic current controlled by the Common Mode FeedBack amplifier (CMFB), and M_6/M_8 are used to generate a static current ranging from 250 µA to 2 mA, thus obtaining a voltage signal onto the 100 Ω termination resistor that ranges from ±25 mV to ±200 mV. The resistors R_0, with a value of 50 kΩ, have been added as a mean of sampling the common mode, and have been implemented using high resistivity polysilicon.

4.2.1 Common Mode Feedback amplifier

The differential nature of the transmitter requires an additional circuit to ensure that the DC common mode of its output does not drift during transmission or due to PVT variations. To achieve this goal, the approach used in this work has been to adjust the gate voltages of the NMOS and PMOS controlling the driver current[83], to ultimately adjust the common mode error against a reference. The common mode feedback amplifier scheme implemented is reported in fig. 4.13. The CMFB used is a very simple current differential amplifier comparing the sampled common mode voltage V_{CM} with the ref-

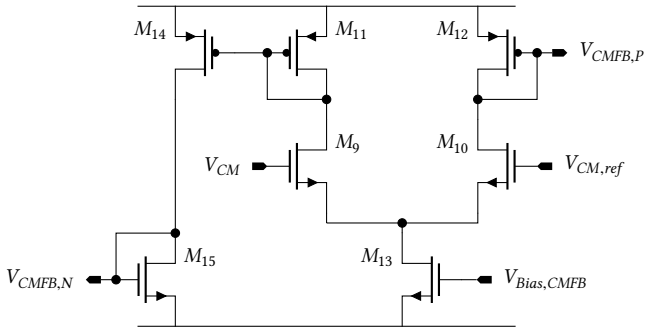


Figure 4.13 – Schematic of the Common-Mode Feedback amplifier used in ARCADIA's custom LVDS transmitter.

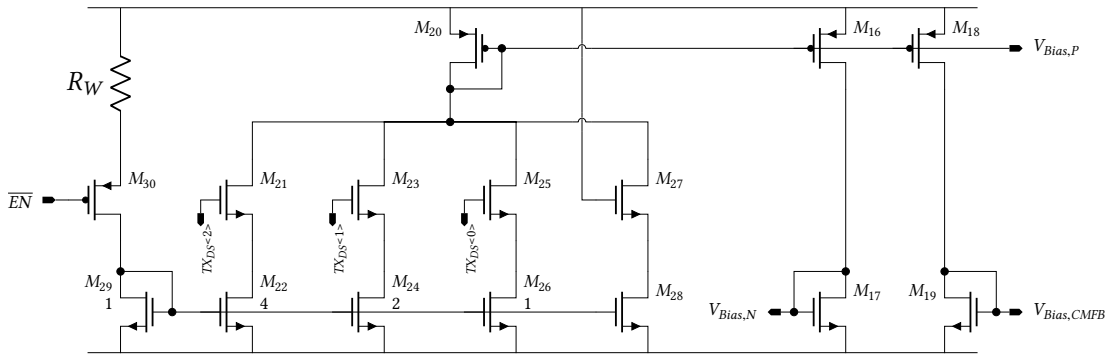


Figure 4.14 – Schematic of the implemented bias and 3-bit control blocks for ARCADIA custom LVDS transmitter.

erence $V_{CM,ref}$ =450 mV generated by a resistive divider. When the two are the same, the currents in the differential pair M_9/M_{10} are balanced and, thanks to the M_{11}/M_{14} plus M_{15}/M_7 mirrors for the NMOS and M_{12}/M_5 mirror for the PMOS, the current just adds to the driver current, leaving the common mode unchanged. In the case of an unbalance, e.g. $V_{CM} > V_{CM,ref}$, the currents in the differential pair will unbalance as well, causing ultimately, in this example, a rise in M_7 current and a drop in M_5 current, both found in the bridge structure reported in fig. 4.12, thus lowering the value of V_{CM}; the same applies to the opposite case. The mirror ratios from M_{11} to M_7 and from M_{12} to M_5 have been kept equal to have a null systematic offset, and have also been chosen quite high (1:100) to minimize the power consumption of the amplifier.

4.2.2 3-bit control of the driver current

The programmability needed to ensure different power consumption settings has been implemented by using a 3-bit current-switching DAC to control the system bias currents, as depicted in fig. 4.14. The DAC current is selected by the bits TX_{DS}<0:2>, plus one additional bit for the circuit enable, $\overline{EN}$, that cuts the current completely. One additional dummy MOS has been added in the always-active branch used to generate the minimum

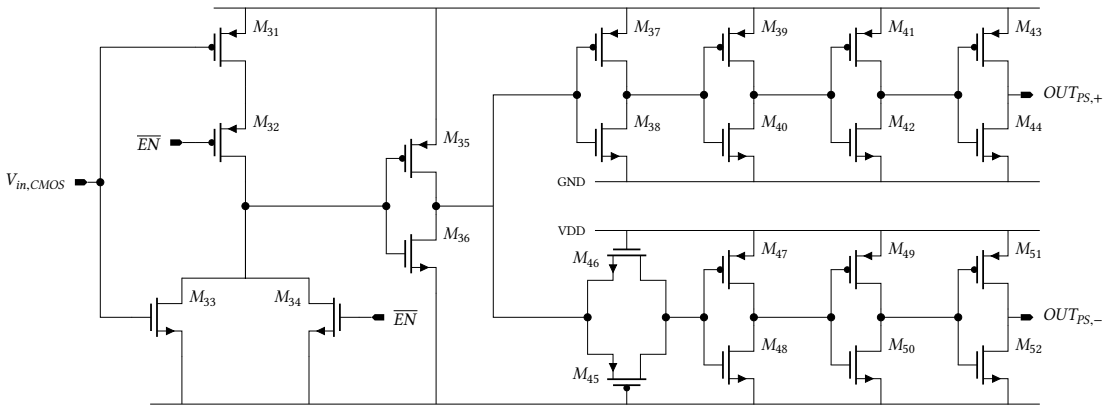


Figure 4.15 – Schematic of the implemented phase-splitter for ARCADIA custom LVDS transmitter.

current configuration in order to maintain symmetry. The resulting current is then mirrored in various parts of the circuit as needed.

4.2.3 Phase splitter

Lastly, as mentioned at the beginning of this section, the driver bridge structure is controlled by a phase splitter. Its purpose is twofold: to convert a CMOS signal into two counter-phase ones to control the output signal and to actually drive the MOSFETs constituting the bridge. Although having very large MOSFETs in the bridge structure helps in having a small $R_{ds,ON}$, it comes at the price of a high gate capacitance, due to its direct dependence on the device area ($C_G \simeq C_{OX}\,WL$). Both functions can be obtained by using two inverter chains, one with an even number of inverters and the other with an odd number of inverters, of increasing W/L ratio to be able to ensure that the inverter can drive the load imposed by the following circuit (the last one will drive the large bridge MOSFETs). Since this design would entail a delay between the two phases due to the different number of inverters in each chain, an always-on switch has been added to mitigate this effect. The $\overline{EN}$ signal that was mentioned in the control and bias section is also present here, as an input of a NOR gate; the input signal is let through only when $\overline{EN}$ is low, otherwise the input is tied to ground by MOSFET M_{34}. The schematic is reported in fig. 4.15.

4.2.4 Transmitter simulation

This section will present the results of the transmitter simulations; the test bench used is reported in fig. 4.16. The test bench simulates the transmitter using a 2 Gbit s^{-1} Pseudo-Random Bit Sequence (PRBS) CMOS signal as input, fed through a technology-library

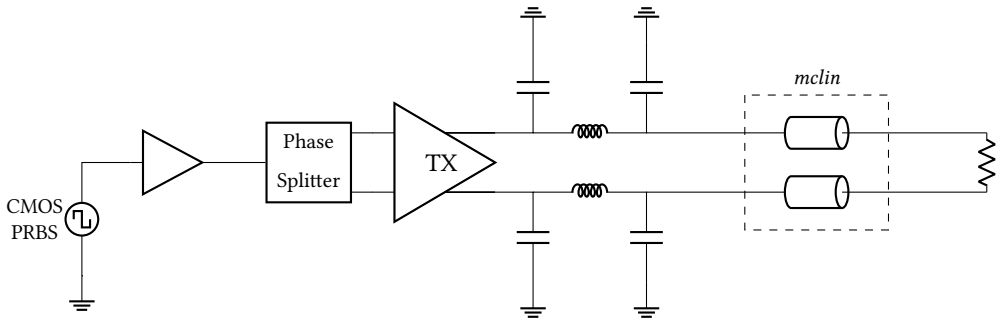
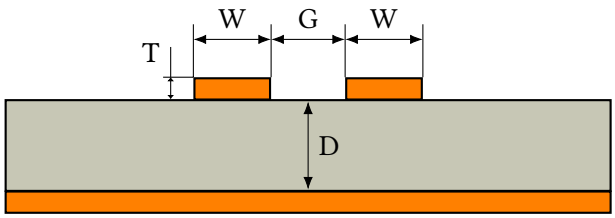


Figure 4.16 – Schematic of the test bench used for the ARCADIA transmitter simulations.

buffer[V] which drives the transmitter input. The transmission line is constituted by a couple of 500 fF parasitic capacitances aimed at providing a model of the pad and chip line capacitances and a series inductance to model the bonding. These are followed by a stripline model (*mclin*) provided by the rfTlinelib Cadence library; this model covers a 15 cm stripline with a differential characteristic impedance of 100 Ω over a standard FR4 dielectric. Of course, this model cannot cover all the cases for PCB implementations that have a 100 Ω differential stripline, but the choice of the width of the traces (127 µm, or 5 mil), which is based on the conservative assumption that 127 µm is a small trace width for a PCB[VI] and thus will provide a ceiling for the line resistivity. The other parameters used in the model have been reported in table 4.1. The two conductors have been assumed to be routed on top of a continuous reference plane and the dielectric has been assumed to be FR4 due to its widespread use; the dielectric constant D_k value for FR4 dielectric around the frequency of operation of 2 GHz is 4.2[85], value that has been used in the simulations that will be presented in this chapter, although FR4 dielectric constant can vary quite a lot[86]. In such cases, the differential characteristic impedance can be adjusted by widening the conductors, if the dielectric constant is lower or by increasing the distance between them, if the dielectric constant is higher; the only concern could be the increase in line capacitance if the dielectric constant is large, but it should be contained enough to be negligible. Altium Designer solver for flat conductors models estimates the capacitance of the line at $54\,\mathrm{pF\,m^{-1}}$ in the simulated case, $60\,\mathrm{pF\,m^{-1}}$ in the case of $D_k = 5.2$ and $48\,\mathrm{pF\,m^{-1}}$ in the case of $D_k = 3.2$. Simulations for the different driving strengths, selected by bits TX_DS<2:0> are shown

[V]This buffer is not actually part of the transmitter, but its implementation in the ARCADIA main demonstrator includes this buffer and for this reason it was added to the simulation.

[VI]Currently, the 5 mil threshold is used as a rule of thumb to separate two PCB manufacturing processes of etching and direct laser imaging. The latter will allow smaller trace widths, down to 1 mil[84], but at a much higher cost. The analysis presented here considers only the traditional, less expensive PCB manufacturing technology, with a 5 mil trace width limit, thus is, in the author's opinion, to be considered a worst case in terms of line resistivity for the applications in which the ARCADIA main demonstrator will be used.



Conductor width (W)	127 µm
Conductor thickness (T)	35 µm (1 oz)
Trace gap (G)	127 µm
Dielectric thickness (D)	1.2 mm
Dielectric constant	4.2

Table 4.1 – Parameters used for the stackup model needed by Cadence mtline component used in the transmitter simulations. The scheme reported is not to scale.

DAC code	Logic 0 [mV]	Logic 1 [mV]	Eye height [mV]	Eye width [ps]	DJ [ps]	RJ [ps]	Power consumption [mW]
000	-34.6 ±0.87	34.8 ±0.85	64.24	489.3	3.06	1.04	0.52
001	-68.1 ±1.57	68.29 ±1.44	127.3	487.1	3.58	1.32	0.94
010	-99.3 ±2.26	99.6 ±2.04	185.9	488.5	3.33	1.22	1.34
011	-131.0 ±3.70	130.9 ±3.53	240.2	488.6	1.75	1.59	1.74
100	-157.4 ±6.49	157.5 ±6.61	275.6	487.0	1.10	1.99	2.08
101	-185.3 ±9.09	185.3 ±8.86	316.7	486.9	0.63	2.25	2.45
110	-209.7 ±9.38	210.3 ±9.98	365.0	486.3	–	2.36	2.75
111	-233.2 ±7.35	233.8 ±7.96	421.1	487.2	3.44	1.45	3.10

Table 4.2 – Summary of the transmitter simulated parameters for a 15 cm transmission line using a 2 Gbit s^{-1} PRBS as input.

in fig. 4.17 and relevant eye parameters and power consumption are summarized in table 4.2; the extracted parameters are the voltages of the logic zero and logic 1, alongside their standard deviations, the eye height and width, as defined in fig. 4.11, deterministic jitter[VII] (DJ) and random jitter (RJ) and the total power consumption. Figure 4.18 shows the MOS corners and PVT corners simulations for the transmitter. The layout of the transmitter is shown in fig. 4.19, with a total area of $60 \times 35\,\mu m^2$.

4.3 Custom LVDS receiver design and simulations

The schematic of the implemented receiver is reported in fig. 4.20; the block has been implemented as a simple symmetric OTA followed by an inverter chain to square the amplified signal. The input pair has been chosen to be of the P-channel type to better function with the chosen common mode voltage of 450 mV, although the receiver has been designed to function with common modes ranging from 200 mV to 700 mV. The

[VII] In the case where TX_DS<2:0> were 110b, the simulation tool was not able to extract the deterministic jitter.

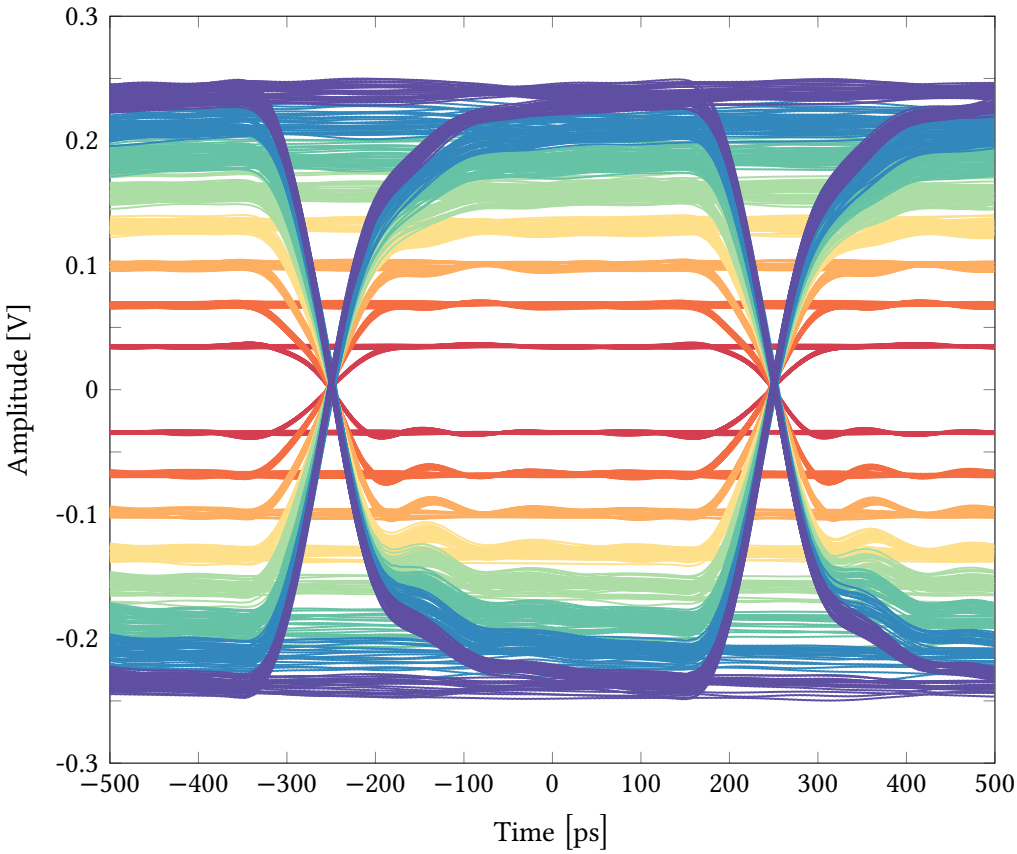
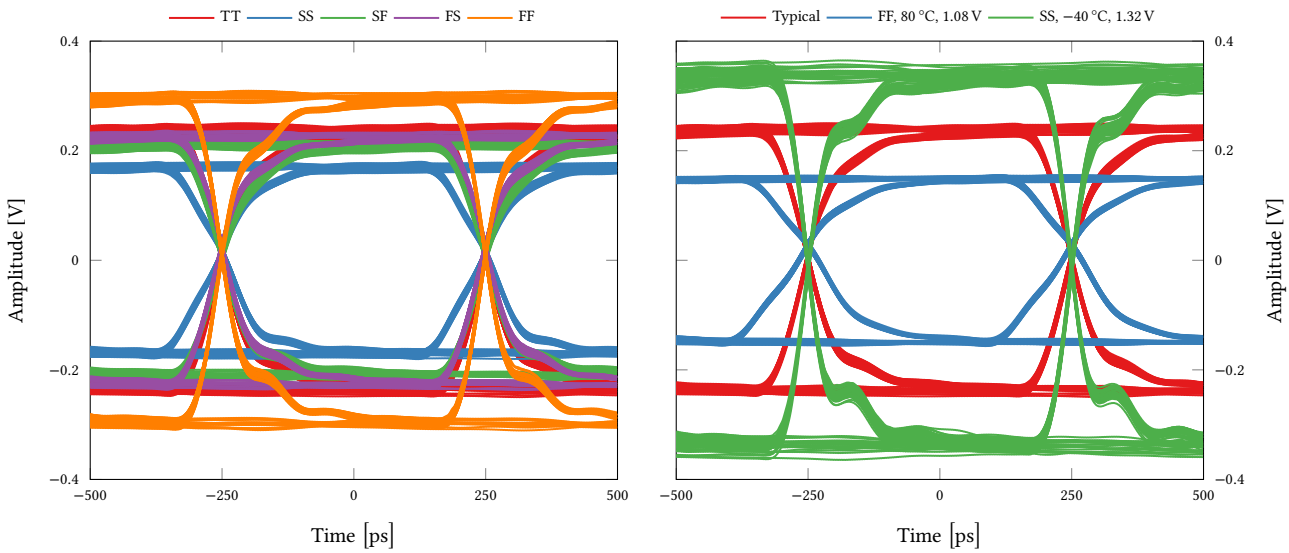


Figure 4.17 – Simulated eye diagram of the transmitter measured on the 100 Ω termination resistor across a 15 cm transmission line using a 2 Gbit s^{-1} PRBS as input.



(a) **(b)**

Figure 4.18 – Simulated eye diagram of the transmitter measured on the 100 Ω termination resistor across a 15 cm transmission line using a 2 Gbit s^{-1} PRBS as input: **(a)** MOSFET corners simulations, **(b)** PVT corners.

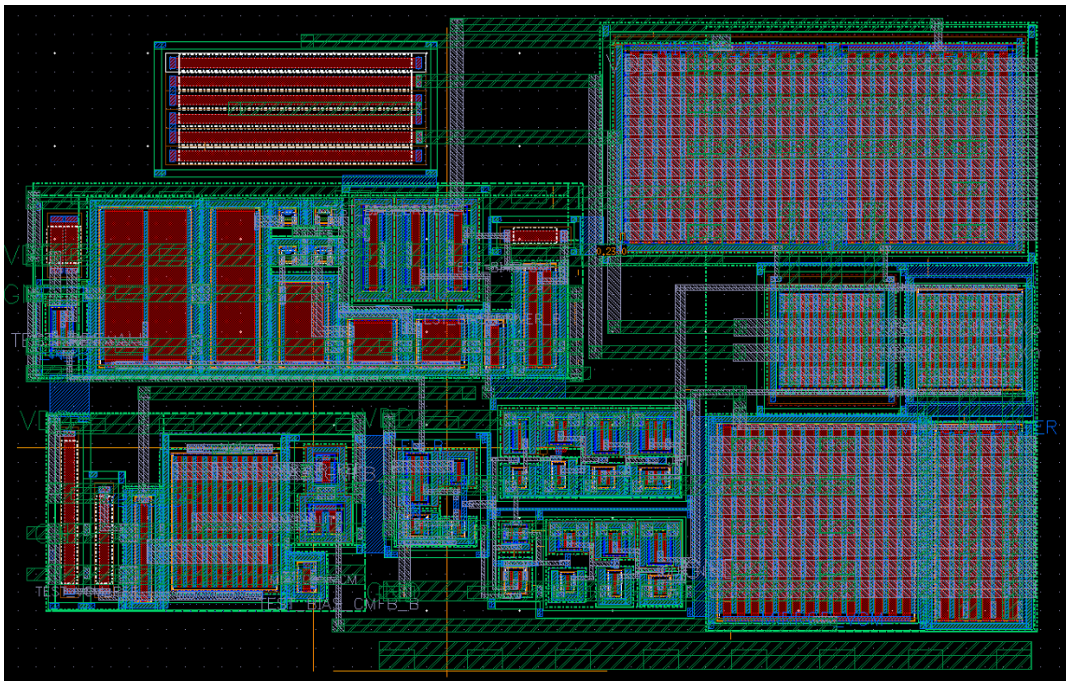


Figure 4.19 – Layout of the ARCADIA custom LVDS transmitter implemented in a 110 nm CMOS technology.

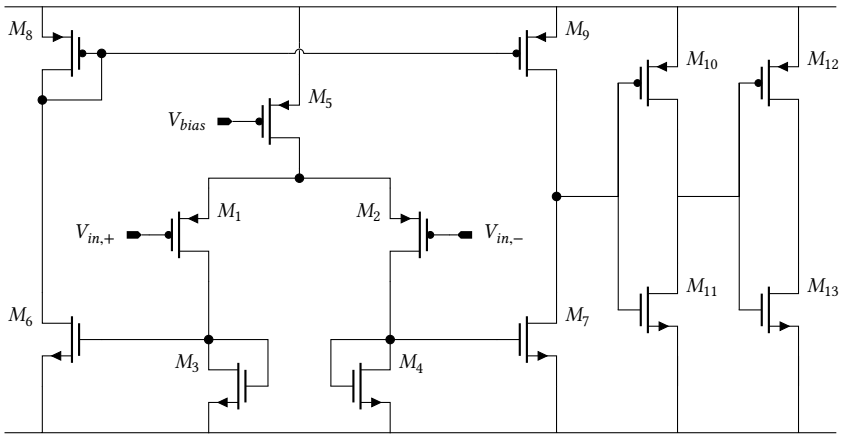


Figure 4.20 – Schematic of the implemented custom LVDS receiver used in the ARCADIA main demonstrator.

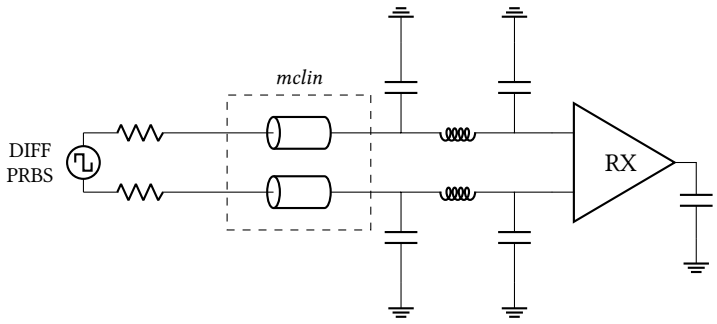


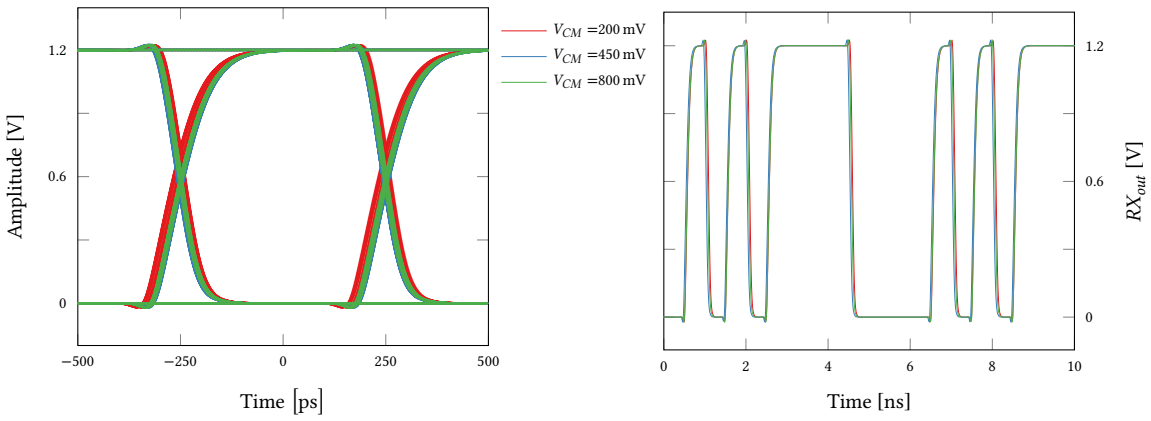
Figure 4.21 – Schematic of the test bench used for the ARCADIA receiver simulations.

receiver has been simulated using a differential PRBS 2 Gbit s^{-1} signal, the two levels being:

$$V_+ = V_{CM} + V_S \tag{4.9}$$

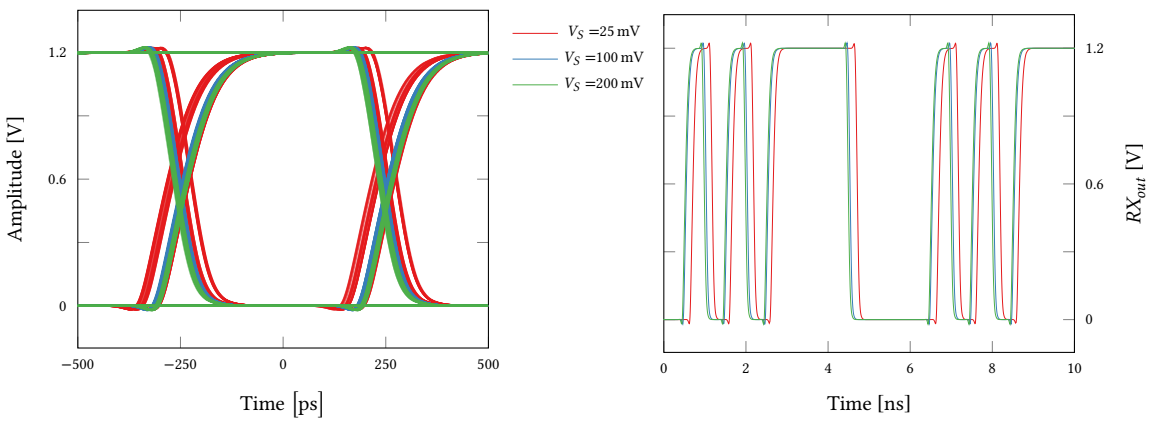
$$V_- = V_{CM} - V_S \tag{4.10}$$

where V_{CM} is the common mode and V_S the swing. This signal is fed onto a transmission line (*mclin*) and a π-network CLC modeling bonding inductance and pad capacitance, ending onto the 100 Ω termination resistor; the receiver has been loaded with a 50 fF capacitance. The 15 cm transmission line parameters are the same discussed for the transmitter and reported in table 4.1, and the consideration on the line distributed capacitance (60 pF m^{-1}) is the same as well. The scheme of the test bench is reported in fig. 4.21. The receiver simulation for common modes ranging from 200 mV to 800 mV is reported in fig. 4.22 and those obtained varying the input signal swing from ±25 mV to ±200 mV are reported in fig. 4.23. The receiver has been designed to accommodate swings down to ±25 mV to retain functionality in the chance of a chip-to-chip communication use case; in this circumstance, the previously described transmitter would be paired directly with the receiver, alongside its configurability down to a ±25 mV swing. The nominal use case will see the use of the ±200 mV swing. The results show that improvements could definitely be made to better the performance of the receiver at the lower swing values, but this has not been deemed critical for the first iteration of the design, where the receiver will be used in conjunction with commercial transmitters on the host PCB. For Monte Carlo simulations, aggregated data are, for the sake of simplicity, reported in the form of eye parameters in table 4.3. The total power consumption of the receiver is 158 µW and it has been realized in the 20 µm × 12 µm cell shown in fig. 4.24.



(a) **(b)**

Figure 4.22 – Simulation results for the receiver, using a 2 Gbit s^{-1} PRBS with ±100 mV swing and different common mode values. Figure **(a)** shows the eye diagram, figure **(b)** shows the time-domain response.



(a) **(b)**

Figure 4.23 – Simulation results for the receiver, using a 2 Gbit s^{-1} PRBS with 450 mV common mode and different signal swings. Figure **(a)** shows the eye diagram, figure **(b)** shows the time-domain response.

	Mean value	Standard deviation
Eye height [V]	1.173	0.011
Eye width [ps]	454	26.2
Deterministic Jitter [ps]	15.6	8.2
Random jitter [ps]	2.88	1.4

Table 4.3 – Eye diagram parameters extracted from the receiver Monte Carlo simulations over 500 iterations.

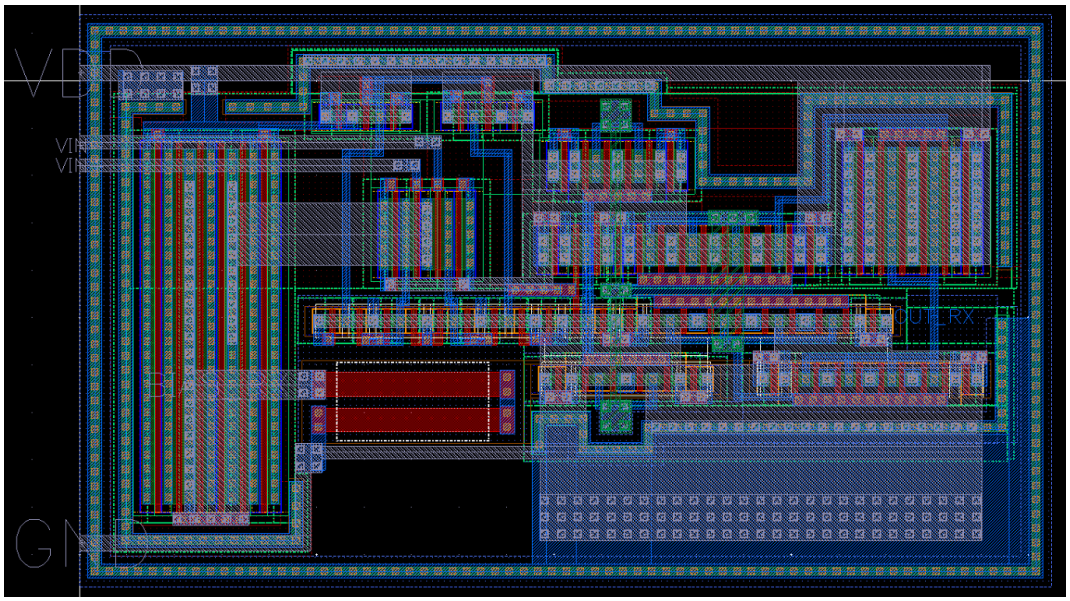
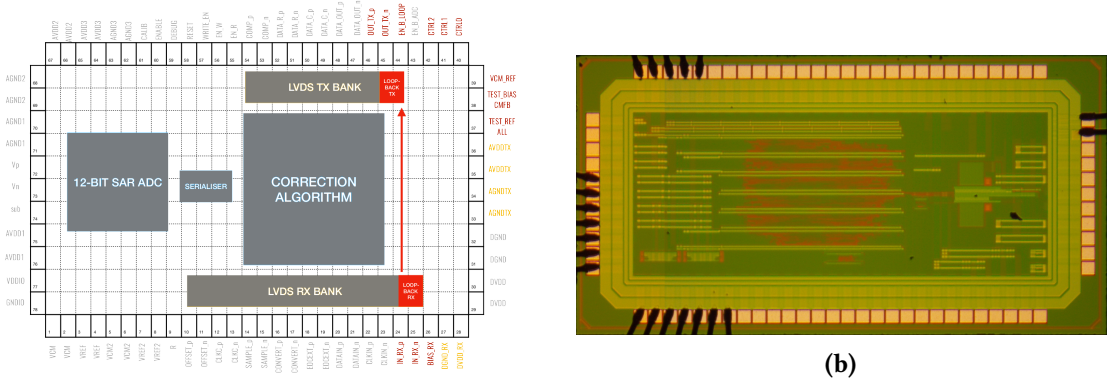


Figure 4.24 – Layout of the ARCADIA custom LVDS receiver implemented in a 110 nm CMOS technology.

4.4 Custom LVDS loopback structure measurements

The custom LVDS transmitter and receiver described in this chapter have been implemented in a test chip in order to verify their operation before the submission of the ARCADIA main demonstrator. The test chip includes a transmitter and a receiver for testing, alongside a SAR ADC developed by the collaboration; fig. 4.25 shows a block diagram view of the chip, together with its micrograph. The transmitter and receiver have been laid out in a loopback configuration[87], as per the schematic shown in fig. 4.26. A PCB has been designed to host the test chip, its schematic reported in fig. 4.27. Some of the passive components shown in the schematic are represented as general impedances allowing the network to be configured according to the testing needs; all footprints have been chosen to be standard 0402[VIII] to ensure low parasitic contributions due to the packages and guarantee that the footprint can host either a resistor or a capacitor. The two envisioned configurations are meant to obtain a DC coupling and an AC one: in the former, Z_1, Z_2 and Z_3 would be mounted using V_{ref} =3.3 V, providing the correct common mode for the incoming LVDS signal and the shifted down one towards the receiver. The latter would see Z_4 and Z_3 as resistors and Z_2 as a capacitor; the series capacitor guarantees the AC coupling, and the resistive divider including Z_4 and Z_3 sets the correct common mode. Resistor R_{AB} has been foreseen to allow testing with only the termination resistor mounted, without any additional impedance network, although it has to be noted that the exposed metal provided by the remainder of the network

[VIII] Imperial units, the footprint is roughly 0.04 in × 0.02 in.



(a) (b)

Figure 4.25 – Block diagram view **(a)** and micrograph **(b)** of the test chip used in the custom LVDS transmitter and receiver verification process.

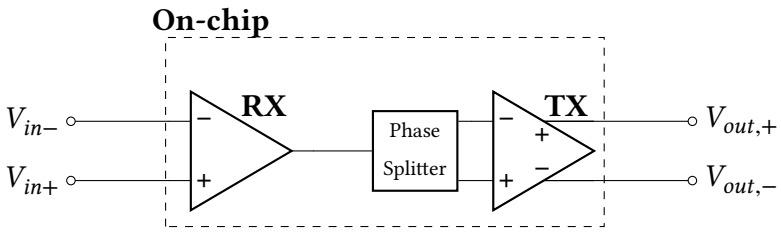


Figure 4.26 – Schematic of the implemented loopback configuration for the custom LVDS verification.

footprints will add to the parasitics contribution, but this should be negligible. A 0 Ω resistor will be mounted in place of Z_2 if the termination R_{AB} is used. DIP-switches have been added to control the $\overline{EN}$ and TX_{DS}<0:2> bits and a couple of bypass capacitors have been placed near the *VDD* bonding pad; the *VDD* supply is generated by an on-board LDO regulator. A 100 Ω termination resistor has been connected to the output of the transmitter using a Z_{diff} =100 Ω differential microstrip having trace widths and trace spacing of 127 µm.

4.4.1 Measurements results

Figure 4.28 shows some measurement results on the test chip for a 2 Gbit s^{-1} PRBS input. For the maximum driving strength, the eye height is 194 mV and the eye width is ∼360 ps, for the minimum one the eye height is roughly 10 mV and the eye width 280 ps and for the 001b configuration the eye height is ∼45 mV and the eye width is ∼310 ps. The results show quite a substantial drop in the performance of the loopback; unfortunately the choice of implementing the transmitter and receiver in a loopback configuration, albeit useful to save a couple of pads, limits the flexibility in the debugging phase. The 111b configuration shows quite a large jitter, most likely coming from some

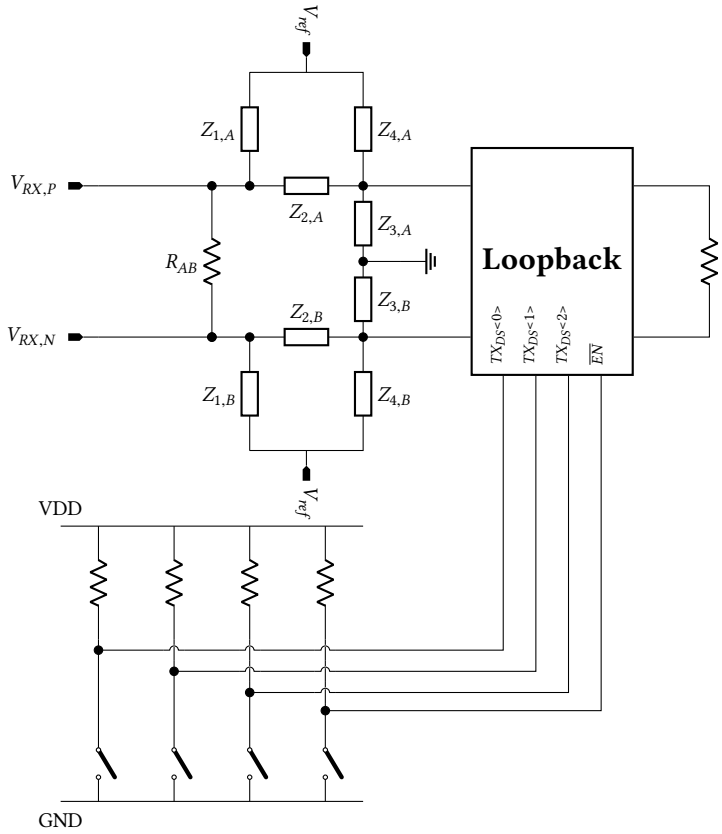
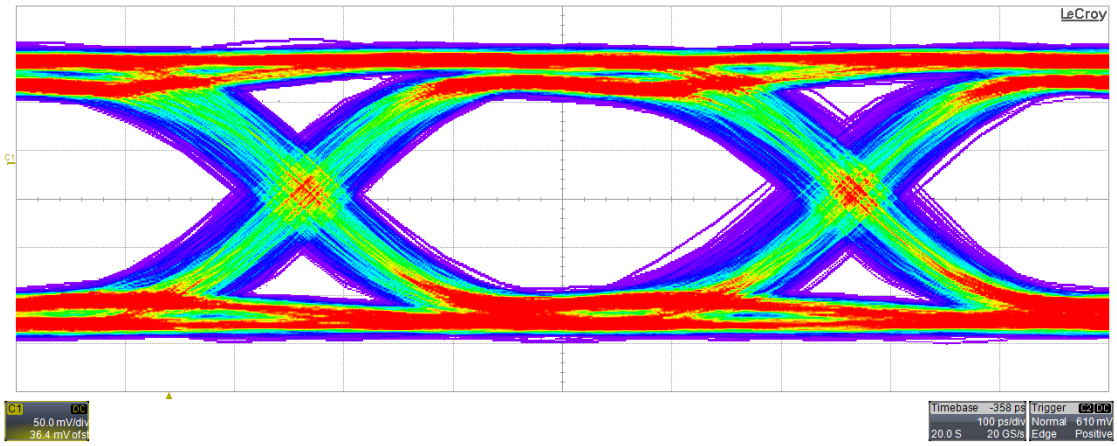
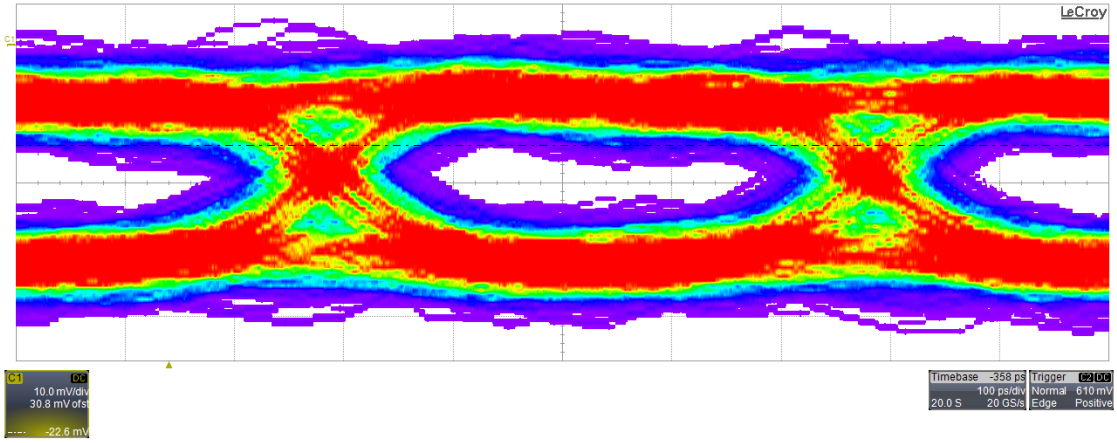


Figure 4.27 – Schematic of the PCB used in the measurements of the custom LVDS test chip.

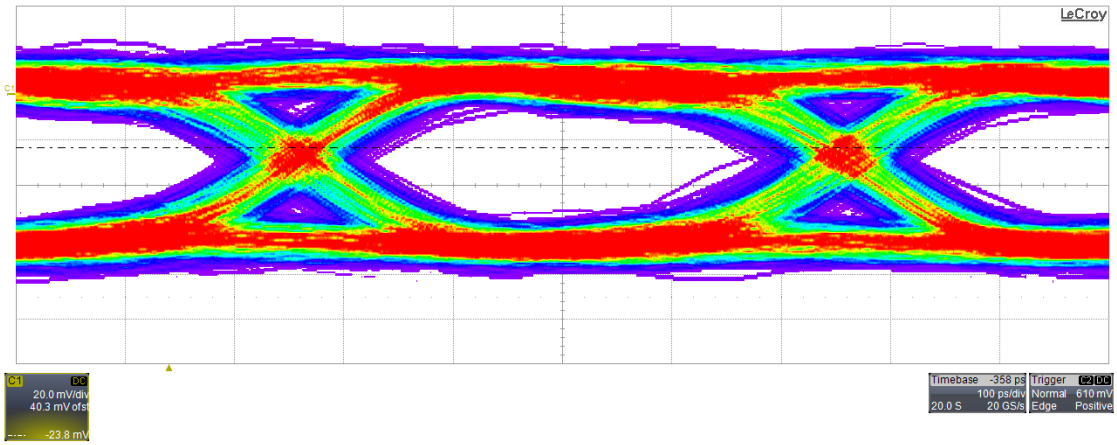
bandwidth limitations: the first contribution investigated was the load capacitance of the receiver which, due to an oversight in the loopback implementation, was larger than expected. The long distance between the receiver and the transmitter in the test chip led to a load capacitance of around 150 fF, but further simulations by loading the receiver with this capacitance resulted in degraded performance, although not comparable with the results seen. Nevertheless, the eye opening for the maximum driving strength is enough to allow the bit sampling without errors; the same cannot be said for the minimum driving strength configuration (TX_{DS}<0:2> at 000b), where the eye is too closed to allow the use of this configuration. The following one (TX_{DS}<0:2> at 001b) should be usable at 2 Gbit s^{-1}. The issue with the 000b configuration does not seem to lie in bandwidth limitations, and it is likely due to the high variability around the two logic levels; fig. 4.29 shows the minimum driving strength configuration at a lower frequency of 1 Gbit s^{-1}, suggesting the fact that, since the eye is not improved, the limitations of this configuration come mostly from the eye height.



(a)

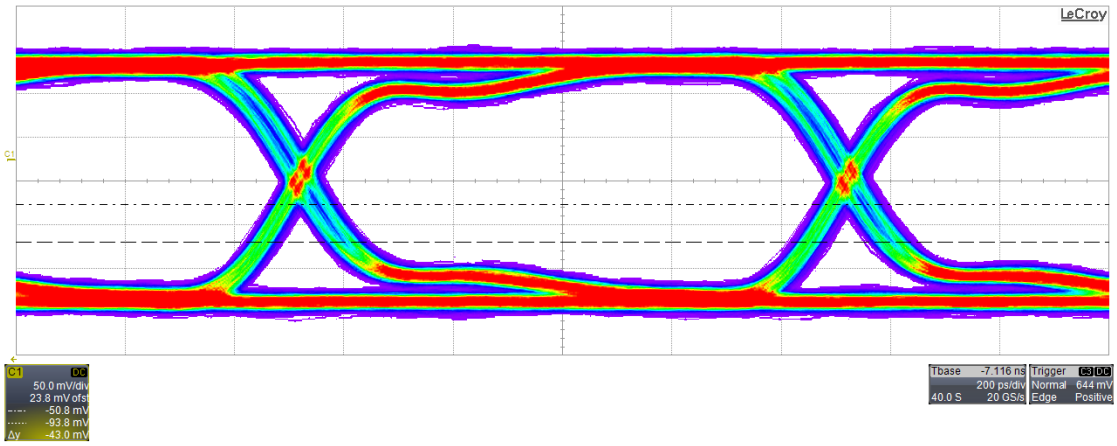


(b)

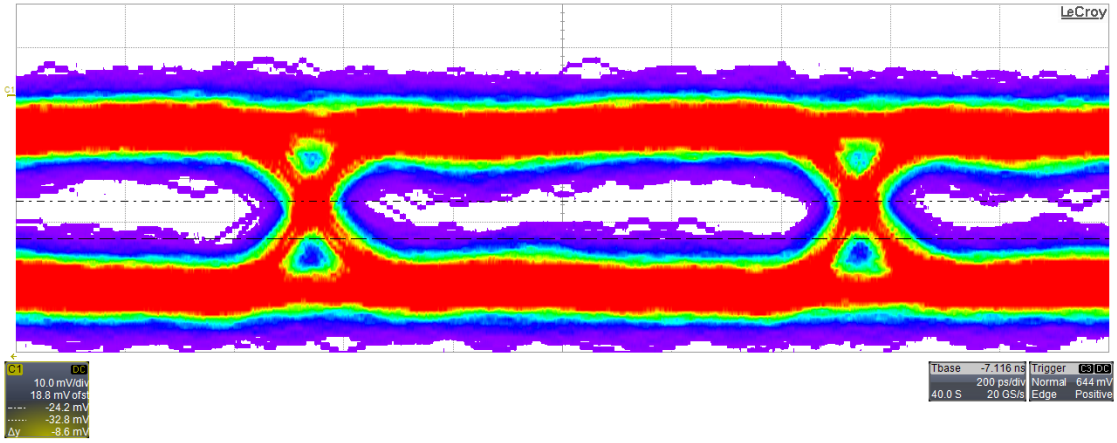


(c)

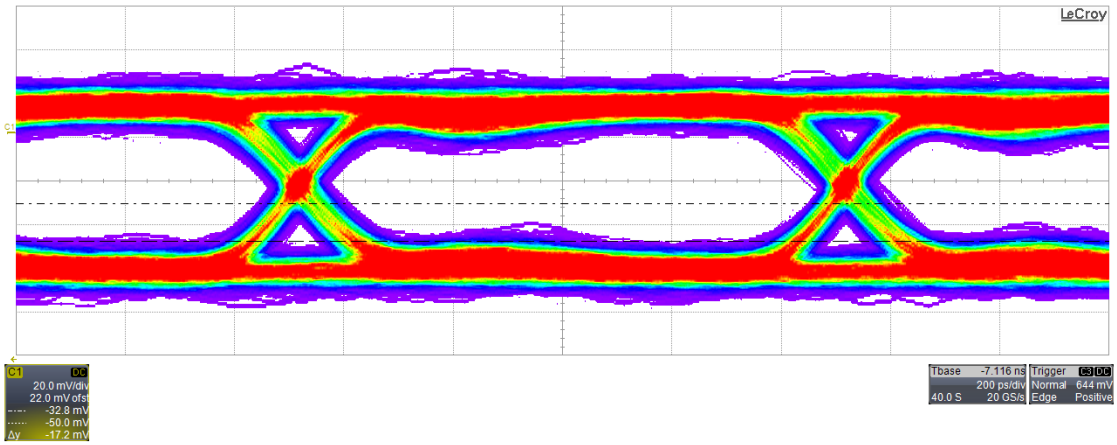
Figure 4.28 – Custom LVDS loopback test results for a 2 Gbit s^{-1} PRBS input; **(a)** shows the eye diagram at the maximum driving strength (TX_{DS}<0:2> are 111b), **(b)** shows the minimum driving strength (TX_{DS}<0:2> are 000b) and **(c)** shows the driving strength right after the minimum (TX_{DS}<0:2> are 001b). All timescales are the same at 100 ps per division, vertical scales vary.



(a)



(b)



(c)

Figure 4.29 – Custom LVDS loopback test results for a 1 Gbit s^{-1} PRBS input; **(a)** shows the eye diagram at the maximum driving strength (TX_{DS}<0:2> are 111b), **(b)** shows the minimum driving strength (TX_{DS}<0:2> are 000b) and **(c)** shows the driving strength right after the minimum (TX_{DS}<0:2> are 001b). All timescales are the same at 200 ps per division, vertical scales vary.

5

Conclusions

In this thesis work, the design of two IP blocks and the characterization of the technology used to develop them has been presented. The work has been carried out within the context of the ARCADIA collaboration, funded by the Italian national institute of nuclear physics (INFN). The goal of the collaboration is to develop a large-scale pixel matrix demonstrator with a fully depleted, monolithic sensor in a standard CMOS technology for use in high energy physics applications, astrophysics experiments and medical imaging.

This work has presented the results from the characterization of the 110 nm CMOS technology used by ARCADIA, starting from basic characteristics extraction such as I_D-V_{GS} and I_D-V_{DS} curves, g_m and g_{ds} from a set of devices included in a purposely designed test vehicle. Additional static parameters have been reported, such as the characteristic current and the subthreshold slope, whose values have been found to be around 300 nA, 1.25 for NMOS devices and 110 nA, 1.3 for PMOS devices. The intrinsic gain has been shown as a function of the inversion coefficient, comparing it with data from other technology nodes; the intrinsic gain for the devices with the shortest channel length is around 15.72, the value being in good agreement with results from newer and older technologies. Lastly, for what concerns static parameters, the threshold voltage has been extracted with three different methods: the Ratio Method, the Transconductance Change Method and the Second Derivative Logarithmic Method. Results for the three methods are in good agreement with each other, yielding a mean value of 458 mV and 420 mV for the threshold voltage of NMOS and PMOS devices, respectively. Following, results of the noise characterization have been reported, showing measured noise voltage spectra for the available devices and aggregated data on white and $1/f$ noise. White noise results show good agreement with the theoretical values, with a mean error of less than 5 % for all devices but the ones with channel length of 120 nm; the low amount of white noise data available for the latter has been deemed the cause of the worse results seen in their case, with a mean error of 16 % with respect to the theoretical values. A higher

bandwidth of the measurement setup should lead to these results being comparable to the ones obtained with longer channel devices. The noise characterization reported concludes with the K_f parameter extraction results: data show good agreement with the expected trend of higher K_f values for short channels, moving then into a constant value for longer channels.

Regarding the developed IP blocks, the more direct contribution of this work to the ARCADIA project has been the development of a bandgap reference and of a custom sub-LVDS transmitter and receiver pair. The bandgap reference has been implemented as a current-mode bandgap with subthreshold-driven MOSFETs as active devices to increase the radiation hardness of the circuit. Simulation results show a temperature variation of 34 PPM/°C and a power consumption of 75 µW. The simulated line regulation is 2.1 % V^{-1} and the PSRR at 1 kHz is −38 dB. An additional block has been added to the bandgap core to generate a 2.4 µA current for the front-end bias; simulations show very similar results to the bandgap performance, and the variation of the current is within specifications after trimming. Both blocks have been laid out with a total area of $100 \times 80\,\mu m^2$; the bandgap core, without the reference current generation, covers an area of $75 \times 80\,\mu m^2$.

The second IP block reported in this thesis is a custom sub-LVDS transmitter and receiver pair for data rates up to 2 Gbit s^{-1}. The transmitter has been designed with a common mode of 450 mV and selectable differential output swings, from a minimum value of ±25 mV, up to ±200 mV, making the power consumption selectable from a minimum of 0.52 mW up to a maximum of 3.1 mW. The receiver has been designed for common modes ranging from 200 mV up to 700 mV and for operation up to 2 Gbit s^{-1}, with a power consumption of 158 µW. The two blocks occupy an area of $60 \times 35\,\mu m^2$ and $25 \times 12\,\mu m^2$, respectively. Measurements conducted on a test chip implementing the two blocks in a loopback configuration show operation up to 2 Gbit s^{-1} for all the driving strength configurations but the lowest one.

The presented blocks have been implemented in the ARCADIA Main Demonstrator 1, which is currently being tested as of December 2021.

Bibliography

[1] N. Wermes, "Pixel detectors … where do we stand?" *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 924, pp. 44–50, Apr. 2019.

[2] G. Contin, L. Greiner, J. Schambach, M. Szelezniak, E. Anderssen *et al.*, "The STAR MAPS-based PiXeL detector," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 907, pp. 60–80, Nov. 2018.

[3] W. Kucewicz, A. Bulgheroni, M. Caccia, K. Domanski, P. Grabiec *et al.*, "Fully depleted monolithic active pixel sensor in SOI technology," in *Proceedings of the 2004 IEEE Symposium Conference Record Nuclear Science.* IEEE, Oct. 2004.

[4] Y. Arai, T. Miyoshi, Y. Unno, T. Tsuboyama, S. Terada *et al.*, "Development of SOI pixel process technology," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 636, no. 1, pp. S31–S36, Apr. 2011.

[5] T. Miyoshi, Y. Arai, T. Chiba, Y. Fujita, K. Hara *et al.*, "Monolithic pixel detectors with 0.2μm FD-SOI pixel process technology," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 732, pp. 530–534, Dec. 2013.

[6] W. Snoeys, G. A. Rinella, H. Hillemanns, T. Kugathasan, M. Mager *et al.*, "A process modification for CMOS monolithic active pixel sensors for enhanced depletion, timing performance and radiation tolerance," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 871, pp. 90–96, Nov. 2017.

[7] I. Perić, "A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology," *Nuclear Instruments and Methods in Physics Research*

Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 582, no. 3, pp. 876–885, Dec. 2007.

[8] M. Mager, "ALPIDE, the monolithic active pixel sensor for the ALICE ITS upgrade," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 824, pp. 434–438, Jul. 2016.

[9] T. Corradino, G.-F. D. Betta, L. D. Cilladi, C. Neubüser, and L. Pancheri, "Design and characterization of backside termination structures for thick fully-depleted MAPS," *Sensors*, vol. 21, no. 11, p. 3809, May 2021.

[10] S. Panati, J. Olave, A. Rivetti, L. Pancheri, F. Cossio *et al.*, "MATISSE: A versatile readout electronics for monolithic active pixel sensors characterization," in *Proceedings of the 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, Oct. 2017.

[11] M. Ajello, W. B. Atwood, M. Axelsson, R. Bagagli, M. Bagni *et al.*, "Fermi large area telescope performance after 10 years of operation," *The Astrophysical Journal Supplement Series*, vol. 256, no. 1, p. 12, Sep. 2021.

[12] P. S. Marrocchesi, "Performance of the charge detectors of CALET with relativistic ions," in *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*. IEEE, Oct. 2013.

[13] N. Higashio, T. Takashima, I. Shinohara, and H. Matsumoto, "The extremely high-energy electron experiment (XEP) onboard the arase (ERG) satellite," *Earth, Planets and Space*, vol. 70, no. 1, Aug. 2018.

[14] M. Aguilar, J. Alcaraz, J. Allaby, B. Alpat, G. Ambrosi *et al.*, "The alpha magnetic spectrometer (AMS) on the international space station: Part i – results from the test flight on the space shuttle," *Physics Reports*, vol. 366, no. 6, pp. 331–405, Aug. 2002.

[15] T. Miller, R. Schaefer, and H. B. Sequeira, "PRIDE (passive radio [frequency] ice depth experiment): An instrument to passively measure ice depth from a europan orbiter using neutrinos," *Icarus*, vol. 220, no. 2, pp. 877–888, Aug. 2012.

[16] G. Anelli, M. Campbell, M. Delmastro, F. Faccio, S. Floria *et al.*, "Radiation tolerant VLSI circuits in standard deep submicron CMOS technologies for the LHC experiments: practical design aspects," *IEEE Transactions on Nuclear Science*, vol. 46, no. 6, pp. 1690–1696, Dec. 1999.

[17] R. C. Hughes, E. P. EerNisse, and H. J. Stein, "Hole transport in MOS oxides," *IEEE Transactions on Nuclear Science*, vol. 22, no. 6, pp. 2227–2233, Dec. 1975.

[18] V. Bezhenova and A. Michalowska-Forsyth, "Aspect ratio of radiation-hardened MOS transistors," *Elektrotechnik und Informationstechnik*, vol. 135, no. 1, pp. 61–68, Jan. 2018.

[19] V. Re, M. Manghisoni, L. Ratti, V. Speziali, and G. Traversi, "Total ionizing dose effects on the noise performances of a 0.13 μm CMOS technology," *IEEE Transactions on Nuclear Science*, vol. 53, no. 3, pp. 1599–1606, Jun. 2006.

[20] V. Re, M. Manghisoni, L. Ratti, V. Speziali, and G. Traversi, "Impact of lateral isolation oxides on radiation-induced noise degradation in CMOS technologies in the 100-nm regime," *IEEE Transactions on Nuclear Science*, vol. 54, no. 6, pp. 2218–2226, Dec. 2007.

[21] F. Faccio, H. J. Barnaby, X. J. Chen, D. M. Fleetwood, L. Gonella *et al.*, "Total ionizing dose effects in shallow trench isolation oxides," *Microelectronics Reliability*, vol. 48, no. 7, pp. 1000–1007, Jul. 2008.

[22] T. V. Nordstrom and C. F. Gibbon, "The effect of gate oxide thickness on the radiation hardness of silicon-gate CMOS," *IEEE Transactions on Nuclear Science*, vol. 28, no. 6, pp. 4349–4353, Dec. 1981.

[23] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, 3rd ed. Oxford University Press, 2010, ch. 8.2.7, p. 382.

[24] C. Enz, F. Krummenacher, and E. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog Integrated Circuits and Signal Processing*, vol. 8, no. 1, pp. 83–114, Jul. 1995.

[25] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, 3rd ed. Oxford University Press, 2010, ch. 4.6, p. 173.

[26] V. Misra and M. Öztürk, *The Electrical Engineering Handbook*, 1st ed. Elsevier, 2005, ch. 3.3.4, p. 118.

[27] E. Vittoz, *Design of analog-digital VLSI circuits for telecommunications and signal processing*. Prentice Hll, 1994, ch. 5.3.1, p. 62.

[28] F. Maloberti, *Analog Design for CMOS VLSI Systems*. Kluwer Academic Publishers, 2003, ch. 1.6.3, p. 31.

[29] G. Anelli, "Analog design in ULSI CMOS processes," in *Proceedings of the 10th Workshop on Electronics for LHC and Future Experiments.* CERN, 2004, p. 55.

[30] A. Ortiz-Conde, F. G. Sánchez, J. Liou, A. Cerdeira, M. Estrada, and Y. Yue, "A review of recent MOSFET threshold voltage extraction methods," *Microelectronics Reliability*, vol. 42, no. 4-5, pp. 583–596, Apr. 2002.

[31] S. Jain, "Measurement of threshold voltage and channel length of submicron MOSFETs," *IEE Proceedings I, Solid State and Electron Devices*, vol. 135, no. 6, p. 162, Dec. 1988.

[32] G. Ghibaudo, "New method for the extraction of MOSFET parameters," *Electronics Letters*, vol. 24, no. 9, p. 543, Apr. 1988.

[33] H.-S. Wong, M. White, T. Krutsick, and R. Booth, "Modeling of transconductance degradation and extraction of threshold voltage in thin oxide MOSFET's," *Solid-State Electronics*, vol. 30, no. 9, pp. 953–968, Sep. 1987.

[34] K. Aoyama, "A method for extracting the threshold voltage of MOSFETs based on current components," in *Simulation of Semiconductor Devices and Processes.* Springer Vienna, 1995, pp. 118–121.

[35] H. Nyquist, "Thermal agitation of electric charge in conductors," *Physical Review*, vol. 32, no. 1, pp. 110–113, Jul. 1928.

[36] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, 3rd ed. Oxford University Press, 2010, ch. 8.5.2, p. 419.

[37] M. Manghisoni, L. Gaioni, L. Ratti, V. Re, and G. Traversi, "Assessment of a low-power 65 nm CMOS technology for analog front-end design," *IEEE Transactions on Nuclear Science*, vol. 61, no. 1, pp. 553–560, Feb. 2014.

[38] A. van der Ziel, "Noise in solid-state devices and lasers," *Proceedings of the IEEE*, vol. 58, no. 8, pp. 1178–1206, Aug. 1970.

[39] L. Vandamme and H. de Werd, "1/f noise model for MOSTs biased in nonohmic region," *Solid-State Electronics*, vol. 23, no. 4, pp. 325–329, Apr. 1980.

[40] K. Hung, P. Ko, C. Hu, and Y. Cheng, "A unified model for the flicker noise in metal-oxide-semiconductor field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 37, no. 3, pp. 654–665, Mar. 1990.

[41] M. Kirton and M. Uren, "Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency (1/f) noise," *Advances in Physics*, vol. 38, no. 4, pp. 367–468, Jan. 1989.

[42] C. Leyris, F. Martinez, M. Valenza, A. Hoffmann, J. Vildeuil, and F. Roy, "Impact of random telegraph signal in CMOS image sensors for low-light levels," in *Proceedings of the 32nd European Solid-State Circuits Conference.* IEEE, Sep. 2006.

[43] M. Manghisoni, L. Ratti, V. Re, V. Speziali, and G. Traversi, "Low-noise design issues for analog front-end electronics in 130 nm and 90 nm cmos technologies," in *Proceedings of the 12th Workshop on Electronics for LHC and Future Experiments.* CERN, 2007, pp. 483–487.

[44] E. Simoen, B. Cretu, W. Fang, M. Aoulaiche, J. M. Routoure *et al.*, "Low-frequency noise spectroscopy of bulk and border traps in nanoscale devices," *Solid State Phenomena*, vol. 242, pp. 449–458, Oct. 2015.

[45] K. Peng and Y. Xu, "Design of low-power bandgap voltage reference for IoT RFID communication," in *Proceedings of the IEEE 3rd International Conference on Integrated Circuits and Microsystems.* IEEE, Nov. 2018.

[46] C.-L. Lee, R. M. Sidek, F. Z. Rokhani, and N. Sulaiman, "A low power bandgap voltage reference for low-dropout regulator," in *Proceedings of the IEEE Regional Symposium on Micro and Nanoelectronics.* IEEE, Aug. 2015.

[47] X. Huang, J. Zhang, L. Liu, W. Huang, Y. Zhang, and L. Yu, "A precision 2.5v bandgap voltage reference with excellent initial accuracy of 0.25% for high resolution ADCs," in *Proceedings of the 9th IEEE International Conference on ASIC.* IEEE, Oct. 2011.

[48] T. Tetzlaff, M. A. A. B., and U. Witkowski, "Estimation of LED junction temperature based on forward voltage method for digital hardware implementation," in *Proceedings of the European Modelling Symposium.* IEEE, Nov. 2016.

[49] B. Razavi, *Design of analog CMOS integrated circuits*, 2nd ed. McGraw-Hill Education, 2017, ch. 12.3.1, p. 513.

[50] R. Widlar, "New developments in IC voltage regulators," *IEEE Journal of Solid-State Circuits*, vol. 6, no. 1, pp. 2–7, Feb. 1971.

[51] K. Kuijk, "A precision reference voltage source," *IEEE Journal of Solid-State Circuits*, vol. 8, no. 3, pp. 222–226, Jun. 1973.

[52] M. K. Adimulam and K. K. Movva, "A low power CMOS current mode bandgap reference circuit with low temperature coefficient of output voltage," in *Proceedings of the Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics*. IEEE, dec 2012.

[53] J. Srour, C. Marshall, and P. Marshall, "Review of displacement damage effects in silicon devices," *IEEE Transactions on Nuclear Science*, vol. 50, no. 3, pp. 653–670, Jun. 2003.

[54] V. Gromov, A. J. Annema, R. Kluit, J. L. Visschers, and P. Timmer, "A radiation hard bandgap reference circuit in a standard 0.13 μm CMOS technology," *IEEE Transactions on Nuclear Science*, vol. 54, no. 6, pp. 2727–2733, Dec. 2007.

[55] Y. Cao, W. D. Cock, M. Steyaert, and P. Leroux, "A 4.5 MGy TID-tolerant CMOS bandgap reference circuit using a dynamic base leakage compensation technique," *IEEE Transactions on Nuclear Science*, vol. 60, no. 4, pp. 2819–2824, Aug. 2013.

[56] T. Vergine, M. D. Matteis, S. Michelis, G. Traversi, F. D. Canio, and A. Baschirotto, "A 65 nm rad-hard bandgap voltage reference for LHC environment," *IEEE Transactions on Nuclear Science*, vol. 63, no. 3, pp. 1762–1767, Jun. 2016.

[57] H. Bohuslavskyi, A. G. M. Jansen, S. Barraud, V. Barral, M. Casse *et al.*, "Cryogenic subthreshold swing saturation in FD-SOI MOSFETs described with band broadening," *IEEE Electron Device Letters*, vol. 40, no. 5, pp. 784–787, May 2019.

[58] K. Ueno, T. Hirose, T. Asai, and Y. Amemiya, "A 300 nW, 15 ppm/°C, 20 ppm/V CMOS voltage reference circuit consisting of subthreshold MOSFETs," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 7, pp. 2047–2054, Jul. 2009.

[59] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, 3rd ed. Oxford University Press, 2010, ch. 4.11, p. 189.

[60] H. shuo Ho, I. Y. ku Chang, and J. Y. min Lee, "The temperature dependence of the electron mobility degradation mechanisms in n-channel metal-oxide-semiconductor field effect transistors with ZrO2 and Sm2O3 gate dielectrics," *Applied Physics Letters*, vol. 91, no. 17, p. 173510, Oct. 2007.

[61] R. Wang, J. Dunkley, T. DeMassa, and L. Jelsma, "Threshold voltage variations with temperature in MOS transistors," *IEEE Transactions on Electron Devices*, vol. 18, no. 6, pp. 386–388, Jun. 1971.

[62] H. F. A. Amir, F. P. Chee, and S. Salleh, "Effects of high energy neutrons and resulting secondary charged particles on the operation of MOSFETs," in *Proceedings*

of the 2014 International Conference on Computational Science and Technology (ICCST). IEEE, Aug. 2014.

[63] T. Ma and P. Dressendor, *Ionizing Radiation Effects in MOS Devices and Circuits.* John Wiley & Sons, Apr. 1989.

[64] B. Schmidt, "The high-luminosity upgrade of the LHC: Physics and technology challenges for the accelerator and the experiments," *Journal of Physics: Conference Series*, vol. 706, Apr. 2016.

[65] F. Maloberti, *Analog Design for CMOS VLSI Systems.* Kluwer Academic Publishers, 2003, ch. 5.8.2, p. 265.

[66] K. O, N. Park, and D.-J. Yang, "1/f noise of NMOS and PMOS transistors and their implications to design of voltage controlled oscillators," in *Proceedings of the IEEE Radio Frequency Integrated Circuits (RFIC) Symposium.* IEEE, 2002.

[67] E. Boufouss, P. Gerard, P. Simon, L. A. Francis, and D. Flandre, "High temperature and radiation hard CMOS SOI sub-threshold voltage reference," in *Proceedings of the 2013 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S).* IEEE, Oct. 2013.

[68] G. Traversi, L. Gaioni, M. Manghisoni, M. Pezzoli, L. Ratti *et al.*, "A rad-hard bandgap voltage reference for high energy physics experiments," in *Lecture Notes in Electrical Engineering.* Springer International Publishing, 2020, pp. 19–24.

[69] D. Gruber, G. Hilber, and T. Ostermann, "A voltage reference with on-chip trimmable temperature coefficient and offset voltage," in *Proceedings of the 18th International Conference Mixed Design of Integrated Circuits and Systems*, Jun. 2011, pp. 231–236.

[70] L. Ratti and A. Manazza, "Optimum design of DACs for threshold correction in multichannel processors for radiation detectors," *IEEE Transactions on Nuclear Science*, vol. 59, no. 1, pp. 144–153, Feb. 2012.

[71] Telecommunications Industry Association, *Electrical Characteristics of Low Voltage Differential Signaling (LVDS) Interface Circuits*, Std., ref. no.: TIA-644.

[72] IEEE, *Low-Voltage Differential Signal (LVDS) for Scalable Coherent Interface (SCI)*, Std., ref. no.: 1596-1992.

[73] P. Clayton R., *Transmission Lines in Digital and Analog Electronic Systems: Signal Integrity and Crosstalk.* John Wiley & Sons, Aug. 2010, ch. 1.2, p. 10.

[74] M. Sadiku and L. Agba, "A simple introduction to the transmission-line modeling," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 8, pp. 991–999, Aug. 1990.

[75] C. Poole and I. Darwazeh, *Microwave Active Circuit Analysis and Design.* Academic Press, Nov. 2015, ch. 2.6, pp. 72–74.

[76] H. Zhang, S. Krooswyk, and J. Ou, *High Speed Digital Design: Design of High Speed Interconnects and Signaling.* Morgan Kaufmann Publishers, Sep. 2015, ch. 1, p. 14.

[77] A. Kumar, S. Mehta, S. Paul, H. S. Gupta, and R. M. Parmar, "Indigenous development of SERDES interface for miniaturization," in *Proceedings of the 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking.* IEEE, Dec. 2012.

[78] B. P. Lathi and Z. Ding, *Modern digital and analog communication systems*, 4th ed. Oxford University Press, 2009, ch. 3.7.3, p. 105.

[79] International Telecommunication Union, *Definitions and terminology for synchnonization networks*, Std., recommendation G.810 (08/96).

[80] F. Herzel and B. Razavi, "A study of oscillator jitter due to supply and substrate noise," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 1, pp. 56–62, Jan. 1999.

[81] E. Säckinger, *Analysis and Design of Transimpedance Amplifiers for Optical Receivers.* John Wiley & Sons, Inc., Oct. 2017, ch. C.1, p. 424.

[82] B. Razavi, *Design of integrated circuits for optical communications.* Boston: McGraw-Hill, 2003, ch. 2.6.4, p. 29.

[83] R. Kłeczek, "The design of low power 11.6mW high speed 1.8gb/s stand-alone lvds driver in 0.18-µm cmos," in *Proceedings of the 17th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES*, 2010, pp. 337–342.

[84] S.-Y. Tseng and J.-W. Chen, "Maskless lithography laser source target light dose reversed mapping," in *Proceedings of the 2020 International Computer Symposium (ICS).* IEEE, Dec. 2020.

[85] E. Holzman, "Wideband measurement of the dielectric constant of an FR4 substrate using a parallel-coupled microstrip resonator," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 7, pp. 3127–3130, Jul. 2006.

[86] J. Aguilar, M. Beadle, P. Thompson, and M. Shelley, "The microwave and RF characteristics of FR4 substrates," in *Proceedings of the IEE Colloquium on Low Cost Antenna Technology.* IEE, Aug. 1998.

[87] S. S. Lee, "External loopback testing on high speed serial interface," in *Proceedings of the 5th Asia Symposium on Quality Electronic Design.* IEEE, Aug. 2013.

Acknowledgements