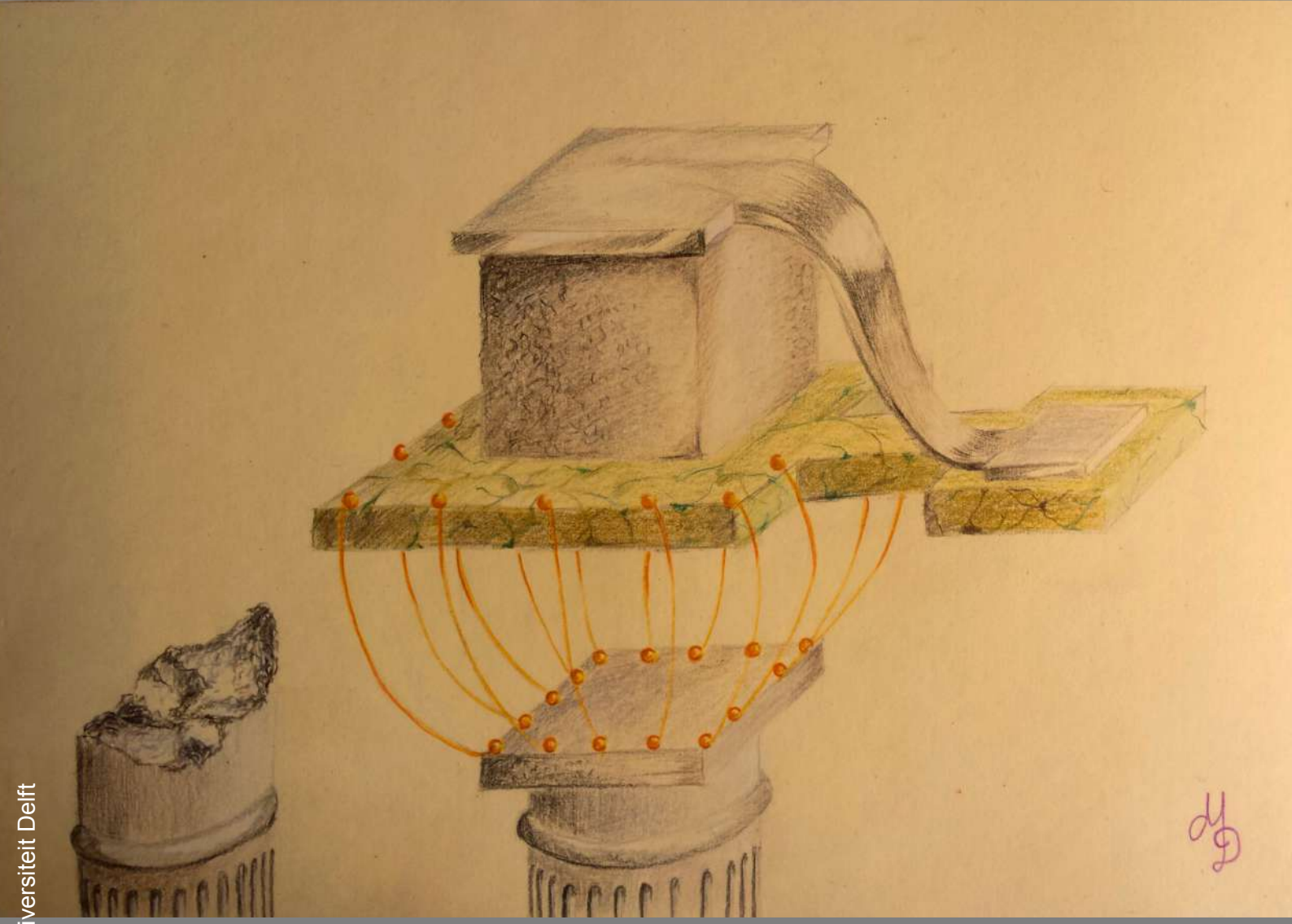


Technische Universiteit Delft

An Element-Level Sigma-Delta ADC for Ultrasound Imaging

Michele D'Urbino

TU Delft

Oldelft Ultrasound

CONFIDENTIAL

An Element-Level Sigma-Delta ADC for Ultrasound Imaging

by

Michele D'Urbino

to obtain the degree of Master of Science in Microelectronics
at the Delft University of Technology,
to be defended publicly on Friday May 19th, 2017.

Student number:	4393856	
Project duration:	January 18, 2016 – May 11, 2017	
Thesis committee:	dr. ir. M.A.P. Pertijs,	TU Delft, supervisor
	dr. ir. F. Sebastiano,	TU Delft
	dr. ir. M. Spirito,	TU Delft
	Ir. B. Lippe	Oldelft Ultrasound

TU Delft

Acknowledgements

Michele D'Urbino
Delft, May 2017

Dedicated to Carlo Segre

I would like to thank a few of the people that helped and supported me during the year and a bit that I dedicated to this work.
My gratitude goes first to my supervisor Michiel Pertijs, an outstanding person, engineer and scholar. His talent is only matched by his dedication, which will be an inspiration for me in the years to come. I particularly appreciate all the time he dedicated to me, as well as his willingness to try innovative but "risky" ideas, such as the one behind this project.
This idea was born from the mind of Boris Lippe, whom I thank for his intuition and the time he invested in my project. The gratitude extends to the entire Oldelft team, for kindly sponsoring and believing in my thesis work.
In particular, I'd like to thank Jacco Ponte, for his infinite patience and good will. Without his much appreciated work, the results of this project would have remained only theoretical.
I want to thank my daily supervisor Chao Chen for always being available and willing to set aside his own priorities in favor of mine, providing endless insightful suggestions that made my work significantly better. Moreover, Zhao Chen also deserves an acknowledgment for his frequent help, and for providing me with a starting point for many pieces of my design.

In general I'd like to express my gratitude to the whole Electronic Instrumentation lab, which is composed by friendly, yet highly skilled group of people. I learned something from each of you, a special mention goes to Douwe, Mingliang, Jan, Cagri, Lorenzo, Burak, Bahman and Eunchul. A huge thank you goes to Zu-Yao and Lukasz, that quickly and brilliantly solved all the endless little and big problems in the life of an engineer, and filled the gap left by my severe computer science illiteracy. I am sure you will not feel too sorry about my departure! Another thank you to Joyce, for her humor and hard work.
Moreover, I'd like to thank members of the defense committee, Fabio Sebastiano and Marco Spirito, for being willing to read and review my thesis and for all the things they taught me during these years.

"No man is an island", and I am no exception: I am grateful to Michela, who has always supported me in the best moments and the worst, for her patience and care. Infinite gratitude goes to my family, that supported my stay in Delft in every possible way. My parents for their constant concern about my well-being, my grandparents, including Laura, for truly being an inspiration and a model of humbleness, grace and virtue. The list goes on with Daniele, a friend above all, Mimma, Ermanno e Anna, for believing in me, Simone, Ilan, Manuel, Gadi, Yardena, Alice, Lillo, Erica, Franco, Dafne, Daniel, Adriana...
Many friends helped me through my Delftian adventure, either by being there physically or just remotely. In particular thanks to Constatino, Rosario, Marco, Monica, Juju, Anton, Stefanos, Andrea and many, many others.
Finally, I want to mention Kofi Makinwa, a man with endless curiosity and with an unique point of view, for welcoming me in the EI group and for the interest he took in my project.

Contents

List of Figures ix

List of Tables xiii

1 Introduction 3
- 1.1 Application . . . 3
  - 1.1.1 TEE Probes . . . 4
- 1.2 Transducer Characteristics & Models . . . 5
- 1.3 Proposal . . . 9
- 1.4 Derivation of Specifications . . . 10
- 1.5 Previous Art & Ultrasound ADCs . . . 12
  - 1.5.1 Ultrasound Sigma-Delta Modulators . . . 12
  - 1.5.2 Physical Loop Filters . . . 15

2 System-Level Design 17
- 2.1 Exploiting the Transducer . . . 17
- 2.2 Current-feedback structure . . . 18
- 2.3 Voltage-feedback structure . . . 20
  - 2.3.1 Design Challenges and Constraints . . . 21
- 2.4 Transducer's Capacitance Compensation . . . 22
  - 2.4.1 Negative Capacitance Circuits . . . 22
  - 2.4.2 Switched Feedback Configuration . . . 24
  - 2.4.3 Delayed Feedback Configuration . . . 24
  - 2.4.4 Compensation Capacitance Configuration . . . 25
- 2.5 Design Flow . . . 27
  - 2.5.1 Bandpass vs. Low-pass Modulators . . . 28
  - 2.5.2 Continuous-time vs. Discrete-time Modulators . . . 28
  - 2.5.3 Multi-bit vs. Single-bit Quantizers . . . 28
  - 2.5.4 Tracking Quantizer . . . 29
  - 2.5.5 Design Parameters . . . 32
  - 2.5.6 Architecture of Choice . . . 34
- 2.6 System-Level Simulation Results . . . 36
  - 2.6.1 Signal and Noise Transfer Functions . . . 36
  - 2.6.2 Step Response . . . 37
  - 2.6.3 Jitter and ELD sensitivity . . . 38
  - 2.6.4 Mismatch Effects . . . 39
  - 2.6.5 Time-Gain Compensation Implementation . . . 41
- 2.7 Conclusion . . . 42

3 Element-Level Design 43
- 3.1 First Stage . . . 43
  - 3.1.1 Topology Choice . . . 43
  - 3.1.2 Theoretical Behaviour . . . 44
  - 3.1.3 Medium V_T NMOS Input Pair . . . 44
  - 3.1.4 Simulation Results . . . 45

3.2 Second Stage . . . 47
3.2.1 Topology Choice . . . 47
3.2.2 Gain Derivation . . . 48
3.2.3 Simulation Results . . . 49
3.3 Tracking Quantizer . . . 50
3.3.1 Comparator . . . 50
3.3.2 Digital Logic . . . 54
3.3.3 DAC . . . 56
3.3.4 Simulation Results . . . 56
3.4 Other Blocks . . . 57
3.4.1 Biasing Block . . . 57
3.4.2 Tunable Compensation RC . . . 58
3.5 Pixel-Level Design . . . 59
3.6 Element-Level Layout . . . 60
3.7 Conclusion . . . 63

4 Chip-Level Design **65**
4.1 Shift-Register . . . 65
4.2 LVDS Outputs . . . 65
4.3 Clock Distribution . . . 67
4.4 Design for Test . . . 68
4.5 Power Domains . . . 69
4.6 Chip Layout & Fabrication . . . 70
4.7 Transducer Integration . . . 72
4.8 Conclusion . . . 74

5 Measurements **75**
5.1 Measurement Setup and PCB Design . . . 75
5.2 Transducer Impedance Characterization . . . 77
5.2.1 Measurements in Air . . . 77
5.2.2 Measurements in Gel . . . 77
5.3 Electrical Measurements without Transducers . . . 80
5.3.1 Open-Loop Spectra . . . 80
5.3.2 Transfer Functions & TGC Measurement . . . 80
5.4 Electrical Measurements with Transducers . . . 83
5.4.1 Closed-Loop Spectra . . . 83
5.4.2 Dynamic Range Measurement . . . 84
5.4.3 Power Consumption . . . 85
5.5 Acoustic Measurements . . . 85
5.5.1 Closed-Loop Spectra with an Acoustic Input . . . 85
5.5.2 Transducer and Modulator Transfer Functions . . . 87
5.5.3 Time-Domain Acoustic Measurements . . . 87
5.6 Performance Comparison . . . 89
5.7 Conclusion . . . 90

6 Conclusions **93**
6.1 Thesis Contribution . . . 93
6.2 Future Work . . . 94

Bibliography **95**

A Appendix **99**
A.1 PCB Design 99

List of Figures

1.1 Age-adjusted death rates for selected leading causes of death: United States, 1958–2014 [1] . . . 4
1.2 Artist impression of the TEE probe's working principle [2] . . . 5
1.3 Two classical transducer models: Mason (left) and KLM (right) . . . 6
1.4 Two near-resonance transducer models: low frequency (left) and high frequency (right) . . . 6
1.5 Impedance measurement performed in water on a 9×12 transducer array. . . . 7
1.6 Magnitude fitting using Z-View . . . 8
1.7 Phase fitting using Z-view . . . 8
1.8 Time-Gain Compensation graphical explanation. . . . 11
1.9 Beamforming in an array of element-level ΣΔ ADCs [3] . . . 13

2.1 Transducer impedance around resonance . . . 18
2.2 Representation of the current-feedback structure (left) and corresponding linear model (right) . . . 19
2.3 Spectrum of the current-feedback structure . . . 19
2.4 Spectrum of the compensated current-feedback structure . . . 20
2.5 Representation of the voltage-feedback structure (left) and corresponding linear model (right) . . . 21
2.6 Representation of the voltage-feedback structure with a realistic position for the DAC (left) and corresponding linear model (right) . . . 22
2.7 Schematic of the voltage-feedback structure without the ΣΔ feedback. . . . 23
2.8 Bode diagram of the thermal noise transfer function, at different values of the amplifier's open-loop gain. . . . 23
2.9 Schematic of the voltage-feedback structure with negative capacitance compensation. . . . 24
2.10 Schematic of the voltage-feedback structure with time windowing compensation. 25
2.11 Schematic of the voltage-feedback structure with delayed feedback compensation. . . . 25
2.12 Schematic of the voltage-feedback structure with added compensation capacitance. . . . 26
2.13 Block diagram of the proposed tracking quantizer. . . . 30
2.14 Visual explanation of the tracking quantizer principle and of Slope Overload Distortion. . . . 31
2.15 Simplified block-level schematic of the chosen architecture. . . . 34
2.16 System-level Simulink simulation for the final architecture . . . 36
2.17 System-level Simulink simulation of the modulator's step response, compared to that of the transducer. . . . 37
2.18 System-level Simulink simulation showing the SNR degradation caused by a jitter of $\sigma = 411ps$. . . . 38
2.19 System-level Simulink simulation showing the SNR behaviour while varying the Excess Loop Delay of the modulator. . . . 39

2.20 System-level Simulink simulation showing the SNR behaviour while varying the compensation capacitance C_c. 40
2.21 System-level Simulink simulation showing the SNDR variation when a mismatch is introduced in the feedback DACs (left) and in the quantizer's DACs (right) . . 41
2.22 System-level Simulink simulation showing the time-domain waveforms of the modulator's critical nodes for continuous TGC. 42

3.1 Schematic of the first stage OTA. 43
3.2 Transient simulation showing the operating point of M_1 versus time. 45
3.3 Simulations of the first stage in open loop, closed loop and resistive feedback mode. 46
3.4 Transimpedance gain of the first stage, while injecting a current at its input. . . 46
3.5 Schematic of the second stage OTA. 48
3.6 Gain settings of the TGC scheme and correspondent switching times. 49
3.7 2^{nd} stage frequency response (post-layout). 50
3.8 2^{nd} stage loop gain magnitude (post-layout) 51
3.9 Schematic of the proposed comparator. 51
3.10 Comparator decision dynamics vs. common-mode voltage 52
3.11 Comparator noise simulation 53
3.12 Comparator offset simulation 53
3.13 Detailed schematic of the logic block included in every element of the array. . . 54
3.14 Schematic representation of the tracking quantizer's DACs 56
3.15 Tracking quantizer transient response (post-layout, over corners) 57
3.16 In-pixel constant-Gm biasing block. 58
3.17 Schematic of the tunable capacitor C_c and resistor R_c. 58
3.18 Pie chart representing the current consumption in μA for every block in the pixel. 59
3.19 Spectra of the post-layout pixel-level simulations, without and with transient noise. 60
3.20 AC response of the loop filter. 61
3.21 Pixel Layout 62
3.22 Area contributions of every block. 63

4.1 Block diagram of the shift register daisy chain. 66
4.2 Schematic of the implemented LVDS driver and pre-driver. 67
4.3 Block diagram of the array's clock distribution strategy. 68
4.4 Block diagram of the PMOS buffer designed for test purposes. 68
4.5 Block diagram of the complete chip 71
4.6 Full Chip Layout 72
4.7 Micrograph of the fabricated ASIC. 73
4.8 Section view of the acoustic stack, divided into layers. 73
4.9 Micrograph showing the transducers on top of the ASIC. 74

5.1 Block diagram of the implemented measurement set-up. 78
5.2 Measurement of transducer impedance through an impedance analyzer. 79
5.3 Measurement of transducer impedance with and without applying gel. 79
5.4 Spectra of the open-loop structure's output for all TGC configurations. 80
5.5 First stage transimpedance measurement, compared to post-layout simulations. 81
5.6 Second stage measured transfer function, compared to post-layout simulations. 81
5.7 Measured time-domain trace showing the TGC functionality. 82
5.8 Spectra of the reconstructed modulator's output for all TGC configurations. . . 83

5.9 Spectrum of the modulator output bitstream, combined with a phase plot of that sample's measured transducers. 84
5.10 Measured Dynamic Range for different TGC configurations, $f_{in} = 4.4MHz$. . . 84
5.11 Block diagram of the set-up that allows for acoustic measurements. 86
5.12 Spectra obtained at different TGC configurations with an acoustic input. 86
5.13 Measured modulator and transducer frequency response with an acoustic input. 87
5.14 Comparison between the modulator output and the transducer response, for the same pixel. 88
5.15 Modulator output with an acoustic pulse as an input, at different distances between the acoustic transmitter and the ASIC. 88
5.16 Arrival delay between two elements, applying an acoustic pulse from a tilted position 89
5.17 Walden Figure of Merit vs. area of modulators published in ISSCC and VLSI [4]. 90
5.18 Comparison of this work with respect to others in the same SNR and BW range [4]. 91

A.1 Daughter-board PCB layout. 99
A.2 Mother-board PCB layout. 100

List of Tables

1.1 Summary of the target specifications 12
1.2 Performance comparison of this work with respect to similar designs. 14

2.1 Possible design choices for similar resulting SQNR performance 32
2.2 Design parameters 36

3.1 Time and gain settings of the implemented TGC steps. 49
3.2 Current consumption and maximum delay over corners for every block in the tracking quantizer. 56
3.3 Noise contributions of the critical transistors and resistors within the pixel. . . . 61
3.4 I/O connections for each pixel. 63

4.1 Shift-register bits. 66

5.1 Statistical data of the electrical transducer model parameters, measured in air (11 measured transducers). 77
5.2 Extracted electrical transducer model parameters, measured in acoustic gel (2 measured transducers). 80
5.3 Comparison between simulated current consumption and measured one 85

Abstract

This work presents an architecture capable of digitizing every channel of an ultrasound transducer array independently and simultaneously. This feature is achieved by exploiting the frequency response of the piezoelectric transducer in order to save area and reduce to a minimum the required building blocks of the ADC.

The transducer is used as an electro-mechanical resonator, which is embedded in a band-pass continuous-time, multi-bit $\Sigma\Delta$ modulator. This converter relies entirely on the noise shaping provided by the transducer, which proves to be sufficient to reach the required specifications.

Each converter only includes an inverter-based transimpedance amplifier, a variable gain amplifier for time-gain compensation and a 3-bit tracking quantizer, which comprises two comparators and two DACs.

A prototype chip has been fabricated in TSMC $0.18\mu m$ technology, featuring 20 channels, with one ADC per channel, and 20 parallel high-speed LVDS transmitters to convey the bitstreams out to the measurement system.

The whole modulator, clocked at $200MHz$, achieves an area of $150\mu m \times 150\mu m$, a power consumption of $800\mu W$ and an SNR of 47dB in a 75% bandwidth around a transmit frequency of $5MHz$.

The measurements clearly show the desired noise shaping behaviour, thus proving that the proposed concept is valid.

Keywords: Band-pass $\Sigma\Delta$, Electro-mechanical Resonator, Tracking Quantizer, Inverter-based OTA.

1 Introduction

1.1. Application

In today's world, many engineers and physicians are continuously striving to create innovations to improve, enrich and lengthen our average life expectancy. Much progress has been made in this direction during the last decades, however there is still room for development in many fields related to health care. Figure 1.1 shows the statistics associated with the causes of death, the ultimate enemy to defeat. This data indicates that heart-related diseases are, and have been, by far the most common cause of death. Therefore an improvement in any technique aimed at curing, diagnosing or preventing such diseases can have a big impact on our average life expectancy. These health issues are especially dangerous for the elderly and they make up to almost 30% of their deaths [1]. Furthermore, it can be noticed that, unlike cancer, the percentage of deaths caused by heart failure has been decreasing slowly but steadily from the 1960s to today, especially in the last decades. This trend suggests that the innovations to standard medicine practice introduced by the recent technology advancements have had a big role in treating heart-related diseases, making it possible to discover and treat them faster, more precisely and more effectively. Extending the same trend to the future, one could conclude that, since there was such an improvement in the treatment of heart issues along the years, it is worth investing time, energy and finances to go even further, as it does seem possible.

One of the most popular diagnostic tools used in the context of heart-related diseases, along with electrocardiogram (ECG), is ultrasound imaging. ECG is mainly used to detect problems related with the generation and transmission of the electrical stimuli in the heart. As such, it plays a very important role in cases of atrial or ventricular fibrillation and other issues related with the heartbeat frequency. However, in order to have a better view of the mechanical deformations leading to heart problems, ECG falls short as it does not provide enough information to form an image, thus, ultrasound imaging is used instead. The vast majority of people is familiar with the ultrasound-based procedure that allows the physicians to diagnose congenital diseases or malformations in foetuses before they are born. The same technique can be employed to render a real-time, 3D image of the heart.

The basic working principle behind this kind of imaging revolves around the ultrasound transducer: at the beginning of each measurement, this element receives an electrical stimulus (usually a pulse) and responds creating an acoustic wave, which will then travel across the patient's tissues. The wave generates a reflection when it crosses the boundary between two different kind of tissues, e.g., the heart muscle and the atrium. By measuring the time

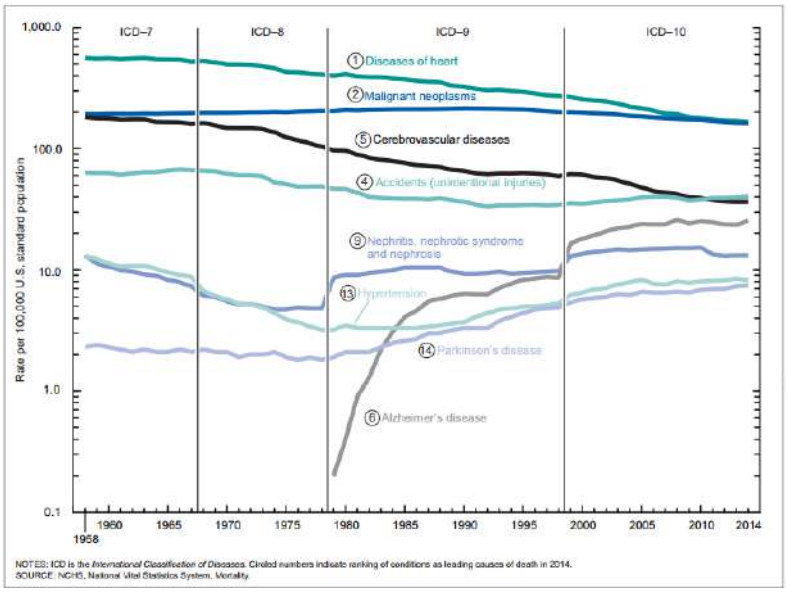


Figure 1.1: Age-adjusted death rates for selected leading causes of death: United States, 1958–2014 [1]

of flight of the forward and return path, the location of the boundary between tissues, an be determined and processed to create an image.

The most common medical procedure used for diagnostic ultrasound imaging is the transthoracic echocardiogram (TTE), which gives valuable information over the heart's overall working state. The advantage of the TTE approach is that it is completely non-invasive, however, because the signal has to cross through many layers of tissue, including the rib-cage, there is a significant attenuation in the reflected signal's amplitude. The attenuation is a function of the distance between the source and the target, as well as of the wave's frequency and the kind of tissue it passes through. For soft tissues, the attenuation coefficient is roughly [5]:

$$\alpha = 0.5 \frac{dB}{MHz \cdot cm} \tag{1.1}$$

Such an attenuation causes the output image to be noisy and thus its diagnostic power can be compromised.

1.1.1. TEE Probes

In order to increase the image quality with respect to TTE based systems, transesophageal echocardiogram (TEE) probes have been introduced. As the name suggests, these instruments have to be inserted into the patient's oesophagus. This allows for much clearer images, as the distance between the oesophagus and the heart (especially the upper chambers and valves) is drastically smaller than in TTE systems.
The design of such probes has proven to be a complex engineering challenge. The main constraints introduced by TEE probes are dictated by the biologic environment in which they have to operate:

- The probe's outer material(s) have to be bio-compatible.
- The probe's power consumption and consequent heating should not result in a temperature increase of more than 1^0C of the surrounding biologic tissue [6]. This threshold accounts for large engineering margins, and can be slightly exceeded without real risks

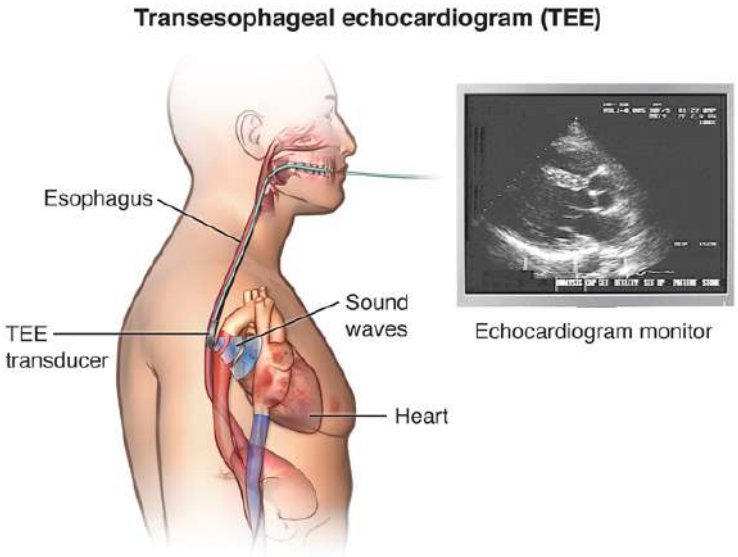


Figure 1.2: Artist impression of the TEE probe's working principle [2]

for the patient. However, it is a meaningful measure when dimensioning the power budget for the probe.

- The probe's girth has to be as small as possible, to avoid discomfort for the patient. This directly translates into a limitation in terms of cable count through the shaft.

In case of passive probes, the number of transducer elements (see section 1.2) is strictly linked to the cable count; therefore, for high resolution imaging (which requires a considerable number of transducer elements) a thick shaft has to be employed, thus introducing, quite literally, a bottleneck in terms of number of transducers.
Active probes, on the other hand, make use of CMOS technology and the multiplexing technique to break this trade-off, allowing for both high resolution imaging as well as low cable count.
Given the context, this work harnesses the properties of the ultrasound transducer to digitize each channel's output. This approach offers a robust digital communication between the TEE probe and the imaging system, while having a limited number of cables. It promises to reduce the cross-talk among channels and introduces the possibility of (partial) on-chip image compression and processing.

1.2. Transducer Characteristics & Models

A wide range of transducer elements have been employed in the ultrasound world, namely: capacitive micro-machined ultrasound transducers (CMUT), piezoelectric micro-machined ultrasonic transducers (PMUT) and bulk piezoelectric transducers. This work will focus only on the latter transducers. As the name states, these elements consist of two conductive layers separated by a piezoelectric ceramic dielectric material. In this work, Lead-Zirconium-Titanate (PZT) is employed, one of the most commonly used materials.
The resonant frequency and the transducer's sensitivity to applied acoustic pressure also depend on their geometry: the thickness must be tuned according to the wanted resonant frequency, while the width and length (usually equal) should be small compared to the wavelength.

Many models have been developed with the purpose of mimicking with discrete electrical elements the energy conversion between the electrical and the acoustic domain taking place in the transducer. The most precise models are the ones which relate the conversion between the two domains with a transmission matrix. However, the coefficients of such matrices often do not have a clear physical relation, making them hard to estimate and treat. The Mason model [7], shown in Figure 1.3 has been one of the most popular due to its lumped nature and relative ease of use. The acoustic and electrical domains are neatly separated by an ideal transformer, whereas the acoustic characteristics are modelled by ideal transmission lines.

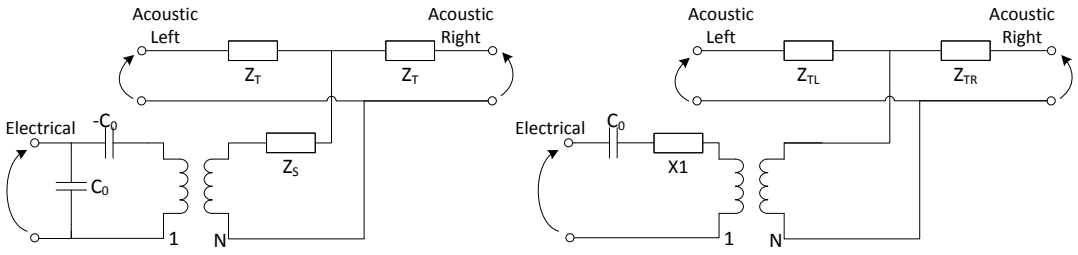


Figure 1.3: Two classical transducer models: Mason (left) and KLM (right)

Another version of Mason's model, the KLM model [8] (shown in Figure 1.3) was proposed, named after the three scientists who developed it, in this version, the negative capacitance proposed by Mason, thought of as "unphysical", is replaced with a time-variant admittance [9].

These two models provide an accurate electrical representation of the transducer by comparing it to a two-port network: the acoustic ports have as inputs/outputs velocity and force, while the electrical port has voltage and current. The Mason and KLM models are somewhat close to the physics of the device, but require an accurate estimation of parameters such as plate stiffness and density, permittivity and elastic constants, which can be a rather time-consuming task

The simpler "Butterworth-Van Dyke" model was chosen, shown in Figure 1.4, because of the

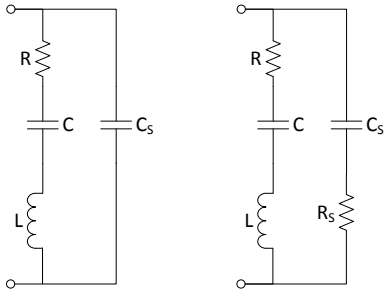


Figure 1.4: Two near-resonance transducer models: low frequency (left) and high frequency (right)

limited time scope of this thesis, together with the fact that the modelling of the transducer's acoustic impedance and characteristics are not of great importance in the scope of this thesis.

This model has been extensively used for electrical simulation and modelling, since it offers a simple electrical representation of the transducer behavior near its resonance frequency. As a matter of fact, this model can be derived from the previously discussed ones, taking into account the acoustic terminations on both sides of the element (one terminated by the backing and the other by the acoustic impedance of the chosen medium). However, this transducer model cannot be deemed suitable for a wide bandwidth [10] and is only reliable around the resonant frequency.
Figure 1.4 shows two versions of the Butterworth-Van Dyke model: the one on the right has an additional resistance in series with the parasitic capacitance C_S [11]. This was added to account for the very-high-frequency behaviour of the transducer. However, the value for this resistor is not precisely known, as the available measurement set-up cannot provide impedance estimations at such frequencies.

The chosen model is divided in two parts: the motional branch, which hosts the series resonator and accounts for the transducer's mechanical resonance mode, and the parasitic capacitance branch, that models the actual capacitance between the transducer's two conductive plates. This parameter, for which this design is centered, can be easily estimated with the well-known capacitance formula, where ϵ_0 and ϵ_r are respectively the vacuum dielectric permittivity and the relative one, A is the area of the parallel plates and t is the distance between them:

$$C_S = \epsilon_0 \epsilon_r \frac{A}{t} \tag{1.2}$$

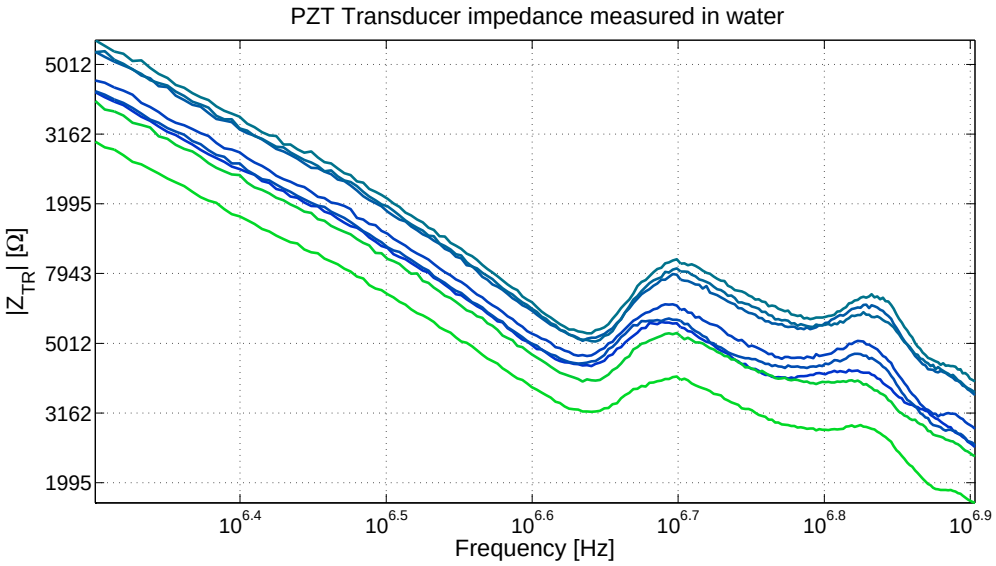


Figure 1.5: Impedance measurement performed in water on a 9×12 transducer array.

The other parameters, concerning the motional branch, can be estimated from previous measurements performed by the acoustics department in TU Delft, shown in Figure 1.5. The measured data was fitted to the model in Figure 1.4 making use of the Z-View software package. For this fitting, up to two additional motional branches have been added to the model, in order to account for the higher frequency resonance modes appearing in Figure 1.5. The magnitude and phase responses of the fitted models are shown in Figures 1.6 and 1.7. For simplicity and simulation time reasons, a single resonator has been chosen for the circuit sim-

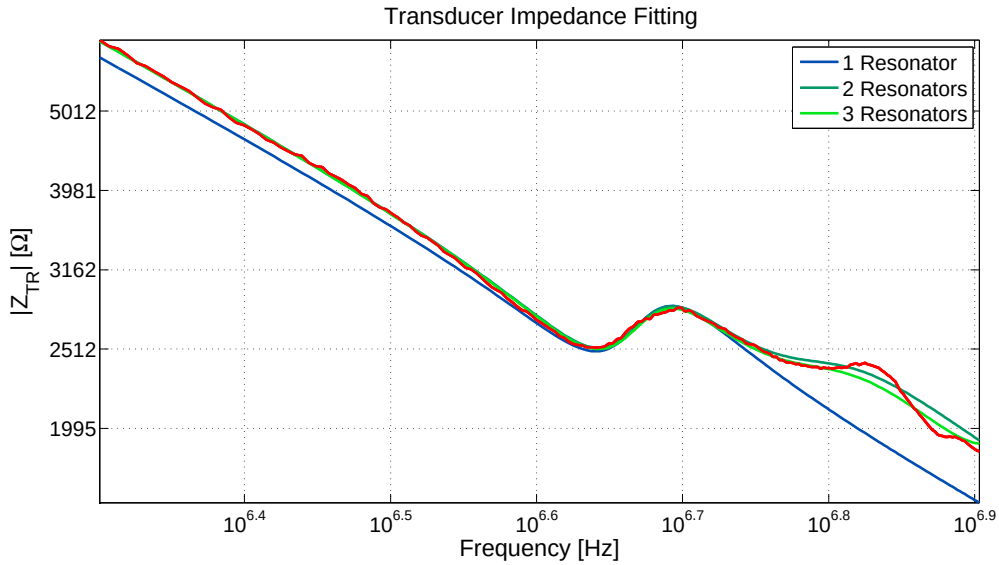


Figure 1.6: Magnitude fitting using Z-View

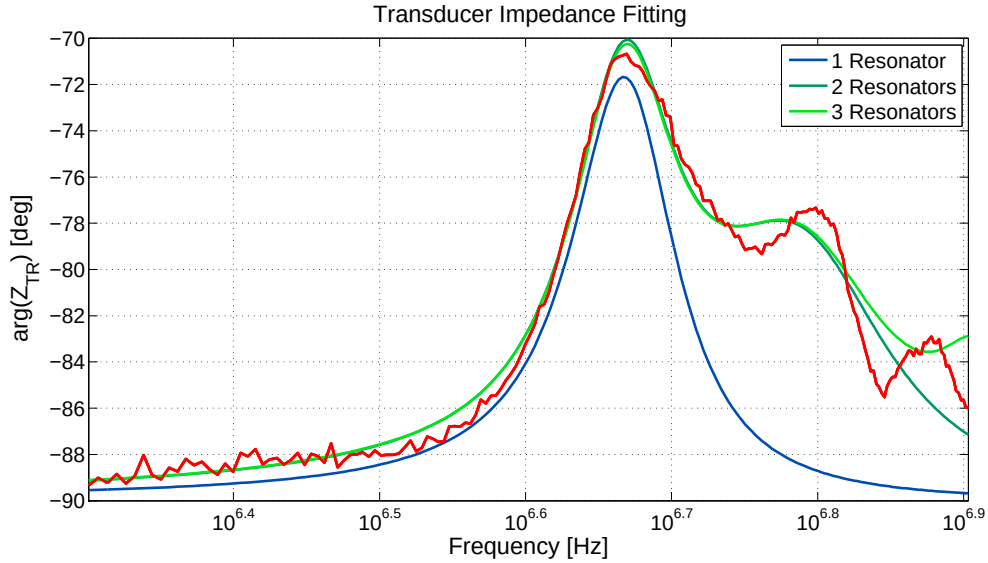


Figure 1.7: Phase fitting using Z-view

ulations, therefore the impedance of the transducer can be written as:

$$Z(s) = \frac{LCs^2 + RCs + 1}{sC_S(LCs^2 + RCs + 1 + \frac{C}{C_S})} \tag{1.3}$$

Where $R = 9448\Omega$, $C = 731fF$ and $L = 1.6mH$. These values, especially the resistance, can be expected to vary significantly depending on the environment. For example, the plots in Figure 1.5 are measurements in water, but performing the same measurement in air would yield a significantly lower resistance (around $2k\Omega$). This effect directly impacts two extremely important aspects for this project: the intrinsic thermal noise added by the transducer and the quality factor of the resonance. The first directly yields the required thermal noise performance and the second, defined as $Q = \frac{\omega_0 L}{R} = \frac{\omega_0}{Bandwidth}$, determines the intrinsic bandwidth of the system. The measurements in water, as opposed to those performed in air, were mostly

taken into consideration for this project, as they are the closest to the actual application that is being targeted. However, the system was designed such that it has some degree of robustness against the inevitable variability in the transducer's frequency response, introduced by fabrication tolerances and measurement limitations. Furthermore, this fitting gives a parasitic capacitance $C_S = 11pF$ much larger than the expected value computed in Equation 1.7. This is probably caused by an incomplete de-embedding of the measurement set-up's added capacitance. Such a large capacitance dominates the impedance curve also at resonance in Figures 1.6, 1.7 and 1.5, while in the real case, with $C_S \approx 2pF$, the resistance in the motional branch should prevail at the resonant frequency.

1.3. Proposal

The main novelty that this work wishes to introduce is the idea of exploiting the filtering properties of the ultrasound transducer in an attempt to reduce the hardware needed for the analog to digital conversion of the acoustic input and thus making it possible to fit an entire oversampled ADC underneath a standard transducer element. In particular, as explained in Section 2.1, the band-pass characteristic of the transducer will be used as the only noise shaping element of a Continuous-Time Band-Pass ΣΔ Modulator (CTBPSDM) specifically built around the transducer.
Having one ADC per channel allows for high frequency data being available about all of the transducers at the same time, without relying on multiplexing and reducing the functionality that has to be implemented in the periphery of the chip to a minimum. The latter feature helps to create very large transducer arrays, without being limited by the size and multiplexing speed of the ADCs placed outside of the core array.
Another benefit that can be derived by this novelty is that the CTBPSDM's loop filter is intrinsically matched to the frequency response of the transducer itself, without the need for centre frequency tuning. Furthermore, given that the whole filtering is performed by the transducer itself, the hardware required to implement the modulator can be kept simple, as will be discussed in Chapter 2. Many of these components are highly digital, therefore they can benefit from technology scaling, achieving lower power and functionality per unit area.
In addition to these advantages, the beamforming operation may consume a lower power in case oversampled ADCs are employed, as will be discussed in Section 1.5.1. Finally, the proposed element-level analog to digital conversion will allow for a meaningful cable count reduction in future TEE probes, thanks to a heavy use of digital multiplexing. For an estimation of the achievable cable reduction factor, some assumptions have to be made:

- A micro-coaxial cable (usually the cable of choice for the communication between the probe and the ultrasound system) has a maximum data rate of $1\frac{GS}{s}$. This assumption is rather conservative.
- A perfect decimation filter is employed, such that the final sampling frequency is exactly twice the Nyquist frequency, for example $7.5MHz$
- The output SNR of the modulator is $45dB$, therefore:

$$ENOB = \frac{SNR - 1.76}{6.02} = \frac{45 - 1.76}{6.02} = 7.18bits \approx 7bits \tag{1.4}$$

From these assumptions, the cable reduction factor can be computed:

$$n = \frac{Data\,rate}{2f_N \cdot ENOB} = \frac{1\frac{GS}{s}}{2 \cdot 7.5MHz \cdot 7bits} = 10.57\frac{channels}{cable} \approx 10\frac{channels}{cable} \tag{1.5}$$

Of course, perfect decimation is hard to achieve, but this calculation should provide a order-of-magnitude estimation of the potential of this system, without even introducing beamforming into the picture. Previous works, such as [12] do not reach such a high cable reduction factor as they rely on analog multiplexing, which is much less robust with respect to its digital counterpart.

1.4. Derivation of Specifications

Now that the core principle has been explained, the design specifications for the system can be derived.

- **Area:** The constraint that is easiest to formulate is the area in which the element-level $\Sigma\Delta$ ADC should fit. This is directly derived from the transducer pitch, which in turn depends on the chosen center frequency. In case of PZT elements tuned around $5MHz$, the wavelength can be calculated as:

$$\lambda = \frac{v}{f} = \frac{1500\frac{m}{s}}{5MHz} = 300\mu m \tag{1.6}$$

  The element pitch should be $\leq \frac{\lambda}{2} = 150\mu m$. However, the actual dimensions of the PZT element also depend on the thickness of the dicing kerf used while building the transducer array, which has a thickness of $30\mu m$, thus the element will measure $120\mu m \times 120\mu m$. The size of the piezoelectric element also directly determines the expected parasitic capacitance. In this case:

$$C_S = \epsilon_0 \epsilon_r \frac{A}{t} = 8.85\frac{pF}{m} \cdot 3200\frac{120\mu m \cdot 120\mu m}{200\mu m} \approx 2pF \tag{1.7}$$

- **Power:** The most limiting factor for the power consumption is the self-heating; the rules provided by the FDA [6] state that the oesophagus' tissue temperature should not increase by more than 1^oC.
  This translates into a total power consumption of the probe of $1W$; distributing this power budget among 1000 channels (this is roughly the number of elements present in commercial probes) yields $1\frac{mW}{channel}$. The latter has to be divided once again between the transmit and the receive circuitry: being that the one considered is a receive-only system, its power budget per channel should be around $500\frac{\mu W}{channel}$. However, as it can be deduced by the way that the previous number has been reached, this is not a hard boundary and there is quite some flexibility associated with it. In fact, many of the papers cited in this thesis declare a power consumption per channel even orders of magnitude higher with respect to this work's target.

- **Instantaneous Signal to Noise Ratio (SNR):** This is a comparison between the noise generated by the transducer-ADC system and the signal intensity at a given instant during the receive phase of an ultrasound cycle (See Section 1.1). This parameter ultimately determines the image quality obtainable by this system. It has been speculated [13] that the SNR of the single ADC should be higher than 50dB. However, the overall SNR of the image is influenced by the number of elements (N) in the array, as the noise introduced in any element is uncorrelated with that of others, while the signal is correlated [13]. The improvement can be computed as:

$$SNR_{total} = SNR_{element} + 3dB \cdot \log_2(N) \tag{1.8}$$

  Regardless, a requirement of $50dB$ SNR for each channel separately was agreed upon.

- **Dynamic Range (DR):** The DR is defined as the ratio in dB between the highest and the lowest signals that the ADC can process. Unlike traditional systems, where the SNR and DR are similar, in ultrasound systems the dynamic range can be very different from the SNR. As a matter of fact, it is possible to exploit the fact that echoes reflected by the nearest tissues will reach the transducers earlier and will face low attenuation (see Equation 1.1), while the waves scattered back by the farthest tissues will be delayed and weaker. Therefore the ADC needs to handle large inputs at the beginning of the receive phase and small ones at its end. In order to take advantage of this information, ultrasound systems usually implement a Time-Gain Compensation (TGC), which increases the gain of the analog front-end as time progresses during the receive phase. In order to get an estimation of the system's dynamic range, the gain range swept by the TGC can be directly added to the instantaneous SNR at the end of the receive phase, as illustrated in Figure 1.8. In this work, a range of around 20 to 30dB was agreed upon, so as to prove the feasibility of this feature in this new kind of ultrasound ADC.

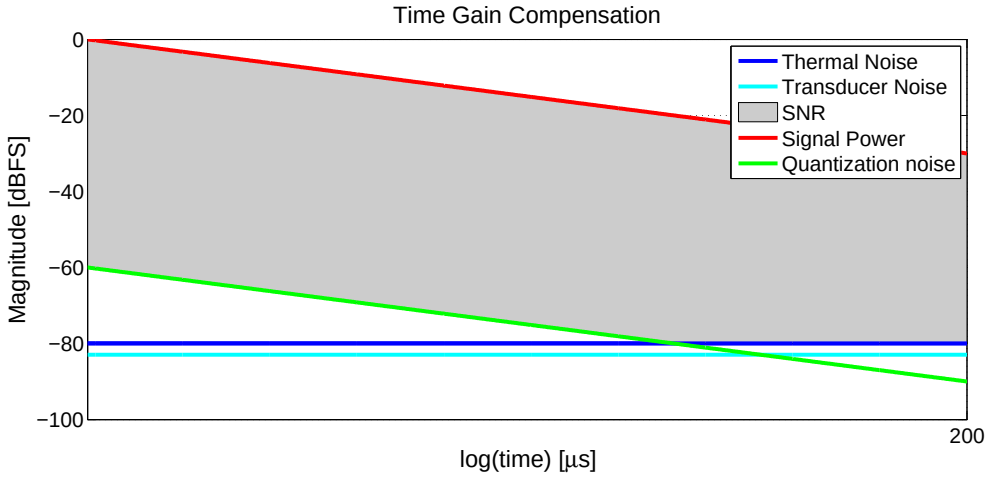


Figure 1.8: Time-Gain Compensation graphical explanation.

- **Bandwidth:** Ideally, for ultrasound applications, it is desirable to have a $-3dB$ fractional bandwidth (FBW) of 100% around the center frequency (f_c), meaning that the range of frequencies of interest is:

$$\left[f_c - \frac{FBW}{100}\frac{f_c}{2},\ f_c + \frac{FBW}{100}\frac{f_c}{2}\right] \tag{1.9}$$

However, in this case, the system bandwidth is intrinsically bounded to the frequency response of the transducer, which can be computed in the case of transducers immersed in water, as explained in Section 1.2:

$$FBW = \frac{100}{Q} = \frac{100 \cdot R}{\omega_0 L} = \frac{100 \cdot 9448\Omega}{2\pi 5MHz \cdot 1.6mH} = 18\% \tag{1.10}$$

This parameter is heavily dependent on the acoustic environment around the transducer, in fact, with the resistance value estimated in air (around $2k\Omega$), the Q factor is much higher and the fractional bandwidth is drastically reduced accordingly. Even in water, an 18% bandwidth is quite poor for the target application, therefore a bandwidth of 75% (from $3.125MHz$ to $6.875MHz$) was selected, considering the transducer's intrinsic bandwidth limitation.

Table 1.1: Summary of the target specifications

Area	$0.0225mm^2$
Power	$500\mu W$
Centre Frequency	$5MHz$
Instantaneous SNR	$50dB$
Dynamic Range	$80dB$
Bandwidth	$[3.125MHz, 6.875MHz]$
Distortion (THD)	$40dB$
Technology	TSMC $0.18\mu m$
FOM	$260\frac{fJ}{conv.}$

- **Distortion:** This work focuses on fundamental imaging, meaning that only the reflected waves at the transmit frequency will be considered, and harmonics will be disregarded. For this reason, distortion is not the main concern in this project. Furthermore, the second harmonic of the center frequency, provided that the latter is greater than $3.44MHz$, falls outside of the bandwidth of interest. In order to keep the continuity with the previous works from this group, a THD of $40dB$ was chosen as an upper boundary.

- **Technology:** Finally, the two technologies made available for this project were $0.18\mu m$ from XFAB and $0.18\mu m$ from TSMC. The latter was chosen because it has a much more reasonable turn-out time and more flexible and frequent tape-out dates, even though technically inferior in terms of capacitor density and overall range of available components.

The specifications hereby derived result in a Walden figure of merit, in the best possible scenario, of:

$$FOM_{Walden} = \frac{Power}{2BW \cdot 2^{ENOB}} = \frac{500\mu W}{2 \cdot 0.75 \cdot 5MHz \cdot 2^8} = 260\frac{fJ}{conv.} \tag{1.11}$$

This theoretical result is in line with the other works in this area. Table 1.1 summarizes the specifications discussed before.

1.5. Previous Art & Ultrasound ADCs

1.5.1. Ultrasound Sigma-Delta Modulators

There have been many publications in the ultrasound community dealing with ways to effectively compress the information received by the transducers, in an attempt to obtain a number of elements much larger than the number of cables needed for the data transmission between the probe and the imaging system. This asymmetry is desirable for two reasons: first, a high number of elements ensures a good spacial resolution and acoustic aperture; second, TEE probes can be made thinner because of a low cable count.
A very common approach is beamforming: this procedure allows for a concise communication between the chip/probe and the ultrasound system. The idea behind beamforming is combining the outputs of different transducers through a process of delay and sum in order to only gain information about a certain spatial direction. This operation can be done in the analog domain such as in [14]; however, doing the same in the digital domain can yield a higher flexibility, together with lower power consumption and complexity, depending on the available technology node. The advantage becomes even more evident in case of oversampled data converters (see Figure 1.9), as they intrinsically have a high delay resolution [15]. For this reason, together with other advantages of oversampled ADCs with respect to their traditional

Nyquist counterpart, many ΣΔ modulators aimed at ultrasound applications were developed.
One of the first attempts to extend the $BP\Sigma\Delta$ approach to the ultrasound field was proposed

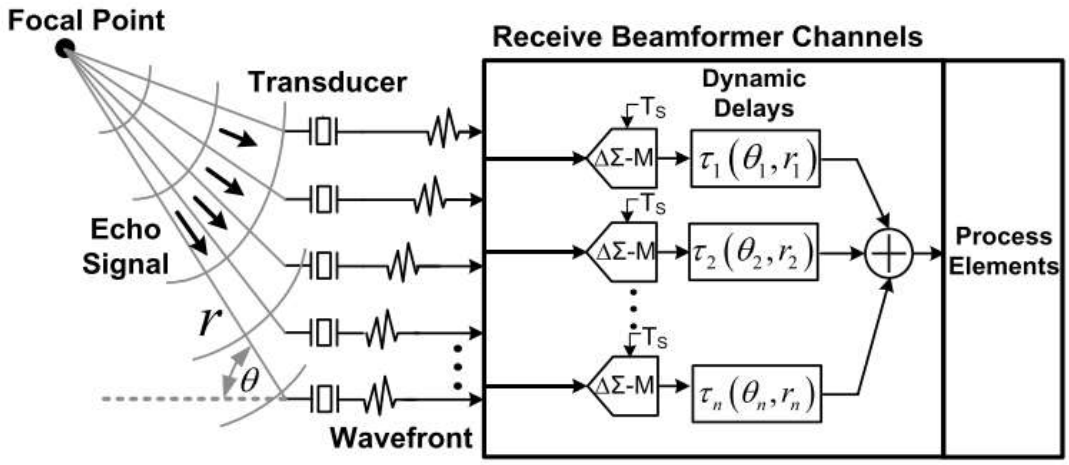


Figure 1.9: Beamforming in an array of element-level ΣΔ ADCs [3]

in [16]. This modulator shows very similar specifications to the ones discussed in Section 1.4: $f_s = 160MHz\ BW = 2.5MHz\ DR = 84dB\ SNR = 60dB$, however the test chip, which was fabricated using a $0.8\mu m\ BiCMOS$ technology and accommodates only one channel, is massive ($10mm^2$) and consumes $1W$.
A more recent study, discussed in [17], shows a power consumption per channel ($3.4mW$) close to this works's target. They employ a conventional 3^{rd} Order $CTLP\Sigma\Delta M$, with almost exactly the specifications described in Section 1.4, including the technology node. However, they do not specify the area, and the higher SNR ($60dB$) is paid by a higher power consumption. Additionally, in [18] a $3mW$, a 4-bit $CTLP\Sigma\Delta M$ is proposed to read out (previously amplified) signals coming from a CMUT array, achieving an area of $0.177mm^2$ and an $ENOB$ of 11.5 bits.

In the latest years, many groups around the world exploited the most recent deep submicron technologies to achieve better ADC FOMs. For example, in [19], combining a $28nm$ technology with a heavy use of digital electronics, a very high power efficiency is reached, employing a $CTLP\Sigma\Delta M$. However, even if the bandwidth and the dynamic range are superior ($BW = 18MHz$ and $DR = 78dB$), the area and power consumption are still too high. Furthermore, the cost of such a technology is extremely high.
The same winning combination of factors was used in [3] ($65nm$ technology and highly digital architecture in a $CTLP\Sigma\Delta M$), very similar to the example that was just discussed. This chip is specifically designed for ultrasound imaging and reaches superb specifications and efficiency, but the downside is the extremely convoluted architecture and the consequently large area and power consumption ($0.16mm^2$ and $7mW$).
Amongst the ADCs that can be found in literature, only a few are suitable, in terms of area, for simultaneous digitization of all transducer elements. In particular, between 2015 and 2017, five of such works were published.

- In [20], yet another 3^{rd} Order, 1-bit $CTLP\Sigma\Delta M$ was employed for ultrasound. In this work, a 12-bit resolution was reached in a $10MHz$ bandwidth, however, the efficiency is quite poor, as each modulator consumes $17mW$ of power, mostly in the digital domain. To this power consumption, one should also add that of the front-end ($13.1mW$) for fair comparison. The high digital power consumption is a problem that has to be tackled

Table 1.2: Performance comparison of this work with respect to similar designs.

	This work	[24]	[23]	[22]	[21]	[20]
Transducer Type	PZT	CMUT	PA	CMUT	PA	-
Technology $[nm]$	180	28	130	65	65	130
No. of Channels	20	16	64	-	8	128
Element Matched	Yes $(150\mu m)$	Yes $(250\mu m)$	No	No	No	No
Center Frequency $[MHz]$	5	5	5	-	260	2.5
Bandwidth $[MHz]$	3.125-6.875	10	8	10	20	10
Area/Channel $[mm^2]$	0.025	0.0625	0.1	0.0175	0.03	-
Power/Channel $[mW]$	0.5-1	17.5	6.32	0.59	13.1	30.1
SNR/Channel $[dB]$	50	60	48.5	45	54	65

also in this project, and arises from the combination of a high sampling rate with an older technology (in [20] $480MHz$ and $0.13\mu m$). The area of a single modulator is not stated.

- In [21] a 4^{th} Order $CTBP\Sigma\Delta M$, with a 5-level quantizer is used for a phased array. In order to stabilize such an architecture, three DACs and two delay elements clocked at $1.04GHz$ are necessary. Such a system would consume a prohibitive amount of power in older technology nodes such as ours, but this work uses a $65nm$ process. The SNDR of the single modulator is $54dB$ while the bandwidth is $20MHz$ around a carrier at $\frac{f_s}{4} = 260MHz$. The power consumption of each modulator is still quite high $(12.1mW)$, but the area is $0.03mm^2$, which would be suitable for a transmit frequency of $5MHz$, such as in this thesis.

- [22] is a simulation study which presents a 4^{th} Order, 1-bit $CTLP\Sigma\Delta M$ for a portable ultrasound scanner. This work bears significant similarities with the project discussed in this thesis: The area, power and SNR ($0.017mm^2$, $0.587mW$ and $45dB$) are in the right order of magnitude. The main differences are the technology node (here $65nm$) and the fact that the transducer is not used as (part of the) loop filter. However, translating this design into older and cheaper technology processes may yield a degradation in efficiency. Furthermore, this modulator is quantization noise-limited, which suggests that it could be further optimized in order to become thermal noise-limited. The provided figures are a result of post-layout simulations, no measurements on silicon have been presented.

- [23] focuses mainly on the digital post processing of the acquired ultrasound data, but it also shows an architecture with one analog front-end, including an ADC, per element. A non-uniformely sampled asynchronous SAR ADC is used to save power. The area, however is too high for a matrix integration $(0.1mm^2)$ and the power of the front-end $(6.32\frac{mW}{channel})$ exceeds this thesis' target specifications.

- Finally, in the very recent [24], the authors present a $250\mu m$ pitch, 16-element array of CMUTs with pixel-matched analog front-end and ADC. A $28nm$ technology was employed to cram the whole receive chain, composed of a TIA, a LPF, a VGA and an ADC in the confined space. The employed ADC is a 3^{rd} order, single-bit $DTLP\Sigma\Delta M$, clocked at 960MHz, with a peak SNR of 60dB and a bandwidth of $10MHz$. The power consumption per channel, which includes the beamforming operation, is still too high $(17.5\frac{mW}{channel})$.

Table 1.2 shows a comparison of the desired specifications with respect to those achieved by similar designs.

From the previous examples, it is clear that continuous-time ΣΔ modulators, both in the low-pass and band-pass flavours, have been employed as ultrasound ADCs because of their power and area efficiency, while at the same time providing sufficient image quality. However, in almost all cases these data converters are shared between multiple elements. This brief literature review also shows how the available technology node heavily limits the achievable specifications. In this work, however, making use of the transducer in a new way, an acceptable power and area efficiency can be reached with an affordable standard $0.18\mu m$ technology. Therefore, scaling further into deep sub-micron processes will surely allow for even better figures of merit.

1.5.2. Physical Loop Filters

In literature, a few examples can be found where a ΣΔ modulator structure was created without much added circuitry simply by using the available sensor's frequency response to gain a noise shaping characteristic.

- In [25] an electro-mechanical ΣΔ structure was built using two PVDF piezoelectric elements, one used as a sensor and the other as an actuator. The feedback is therefore provided in the mechanical domain, thus creating a force balance loop, rather than in the electrical domain. The downside of such a structure is that it relies on two elements, that can be stacked vertically. This poses many technical and fabrication challenges which ultimately decrease the yield. Furthermore, the sensor's frequency response, modeled with Mason's model (See Section 1.2), is exploited for its low frequency behaviour (before the resonant frequency), therefore implementing a $CTLP\Sigma\Delta M$, while the most interesting part of the transducer's frequency response is really around its resonance.
- Another $CTLP\Sigma\Delta M$ was proposed in [26], built around an electrochemical sensor. The "input signal" in this case is chemical, therefore extremely slow. This application makes use of a high time constant already present in the system as an integrator for the modulator. The circuitry needed around the sensor is therefore highly digital and consumes comparable power with respect to sensor systems with an analog output.
- [27] shows a smart wind sensor, which, in order to obtain a digital output, exploits a thermal ΣΔ modulator, using the thermal capacitance of silicon as an integrator. Instead of measuring the temperature gradient caused by the air flow, this system cancels it, using resistive heaters as the feedback elements and thermopiles as sensors and summing nodes.
- Finally, [28] proposes a $CTBP\Sigma\Delta M$ where a mechanical SAW resonator was used as a replacement for LC tank circuits in the loop filter, because of its superior Q factor at high frequencies. This work is intended as a general purpose high-frequency ADC, therefore the input is purely electrical. The resonator has two available nodes and can therefore be easily included in a ΣΔ as a loop filter; on the contrary, in this work, only one node is available (see Section 2.1). In their design process, the authors encountered similar problems to the ones discussed in Section 2.4, and proposed similar solutions.

2

System-Level Design

In this Chapter, the various possible system architectures that were evaluated will be discussed. We will start from the transducer and follow a path that leads to the final architecture that was brought to the next step: schematic-level design (Chapter 3). Along the way, the major trade-offs and issues that were encountered will be highlighted.

2.1. Exploiting the Transducer

As previously discussed, in order to define an optimal system architecture for this ultrasound ADC, one should start from the transducer, as it not only provides the input signal, but also constitutes the loop filter of the $\Sigma\Delta$ modulator.

The electrical input signal, caused by the incident acoustic wave on the transducer, can be modelled as an ideal voltage source in series with the resistor in the Butterworth-Van Dyke model, or as an ideal current source in parallel to the same resistor. In those two electrically identical representations, the input signal is strategically placed in the same position as the system's main noise source, which is the model's resistor noise.

Ideally, it would be desirable to provide the modulator feedback signal in such a way that it can be directly subtracted from the input signal, effectively implementing a summation node. For this purpose, feedback in the form of an acoustic signal would be very desirable, but that would require an additional piezoelectric actuator to implement the conversion from the modulator's electrical output to the acoustic feedback, which poses fabrication challenges [25] beyond the scope of this thesis. Therefore, only feedback in the electrical domain can be provided, however the model's internal nodes (for example the node between the voltage source and the resistor) do not have a physical equivalent.

As a matter of fact, only one node is available for this purpose in the chosen transducer model, even though the piezoelectric element itself is a two-terminal device. This is because one side of the transducer is always connected to a ground foil that is in common among all the transducers in the array. The ground foil takes its name from the fact that it should be connected to a fixed potential, which can be ground. This node being common for all of the transducers, it cannot be used to provide a different feedback signal to each modulator. This situation is highly unusual, as conventional loop filters have separate input and output nodes.

All things considered, there are three ways to use the only available node:

- Offer a high impedance to the node, read out the voltage and provide feedback in the current domain
- Offer a low impedance to the node (through a virtual ground), read out the current and provide feedback in the voltage domain

- Offer an intermediate (possibly matched) impedance to the node.

The third alternative requires an explicit resistance, in series (in case of current read-out) or in parallel (for voltage read-out) with the transducer. This resistance would contribute with as much noise as the transducer, in case it is matched with that of the transducer, without yielding any obvious advantage. For this reason, the third option was discarded. In Sections 2.2 and 2.3, an in-depth analysis of the first two options will be provided, together with the reason why the second one was selected.

2.2. Current-feedback structure

In the current-feedback case, the parasitic capacitance C_s (see Figure 1.4) would be exploited as an integration element for the first stage of a low-pass ΣΔ converter. The voltage output caused by the feedback current flowing into the transducer will be:

$$V_{transducer}(s) = I_{feedback}\frac{LCs^2 + RCs + 1}{sC_s\left(LCs^2 + RCs + 1 + \frac{C}{C_s}\right)} \tag{2.1}$$

The numerator shows a resonant behaviour at the intended frequency, the first part of the denominator (sC_s) determines the desired integration and finally the second part of the denominator $\left(LCs^2 + RCs + 1 + \frac{C}{C_s}\right)$ adds the anti-resonance at a slightly higher frequency with respect to the resonant one. A system-level implementation of a complete modulator was sim-

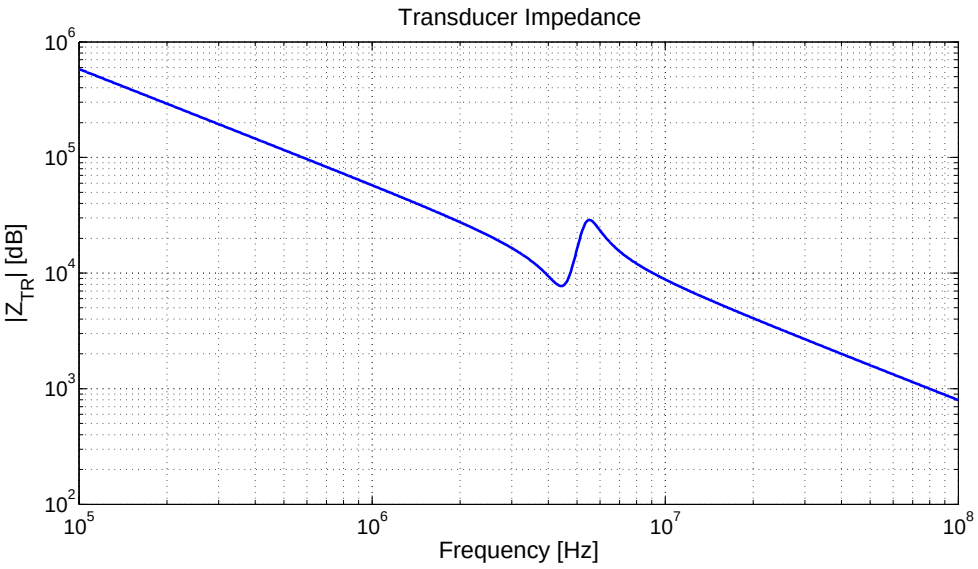


Figure 2.1: Transducer impedance around resonance

ulated, using the current-feedback solution. The electrical representation of the system and the corresponding linear model are shown in Figure 2.2. A second stage was added in order to reach the desired SQNR and, in order to keep the modulator stable, a feed-forward path was introduced, together with a local resonator for optimal zero placement. The results obtained from simulating the previously described system on Mathworks' Simulink (using some basic electrical components from the Simscape library) are shown in Figure 2.3. The blue line shows the spectrum at the modulator's digital output, highlighting the bandwidth of interest and the main tone. The dashed line shows the behaviour that the linear model predicts for the noise transfer function (NTF), while the dot-dash trace represents the signal transfer function

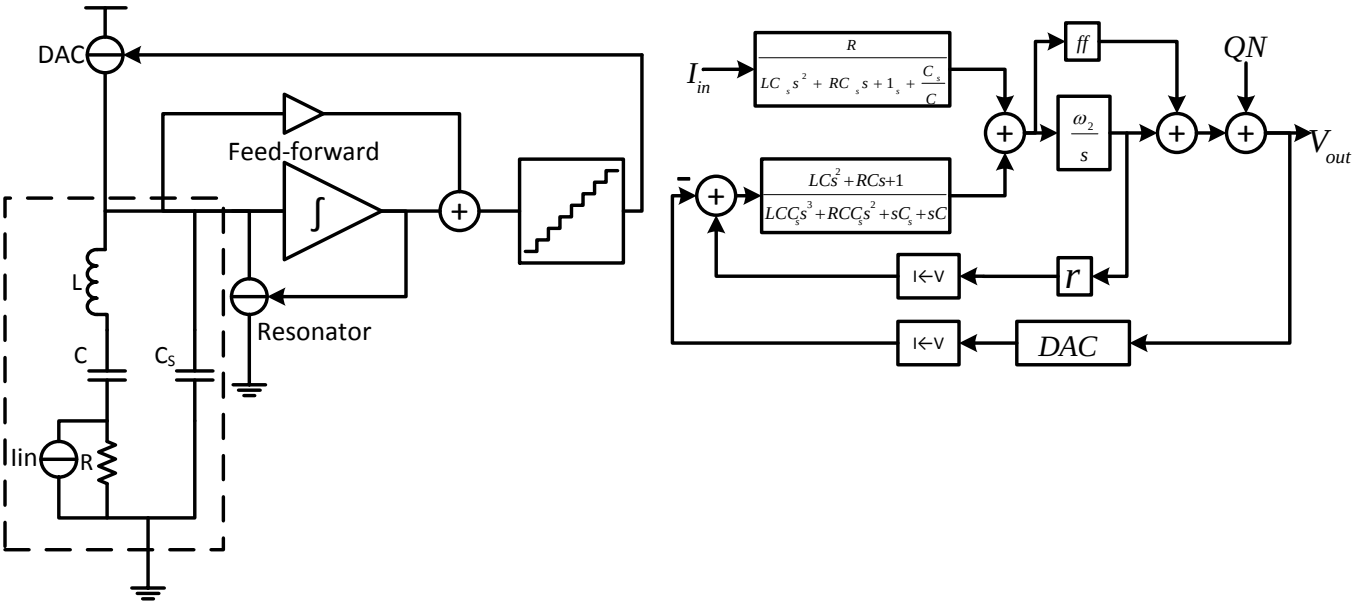


Figure 2.2: Representation of the current-feedback structure (left) and corresponding linear model (right)

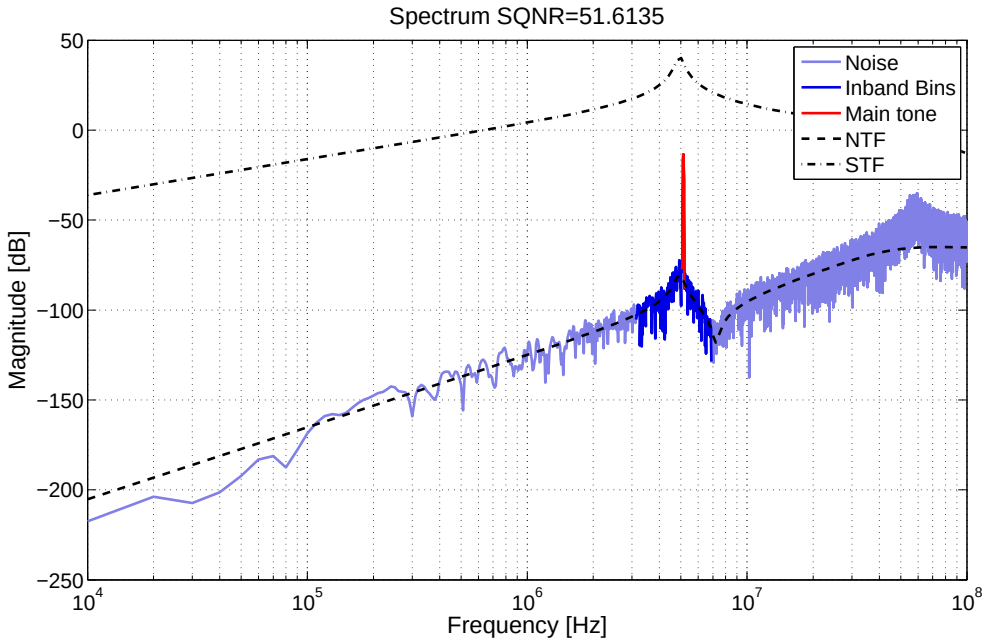


Figure 2.3: Spectrum of the current-feedback structure

(STF). The linear model is obtained by substituting the quantizer with an additive node, which injects quantization noise, assumed white, in the system. This prediction is in agreement with the system level simulation, up to around $40MHz$.
One can immediately notice how inconvenient the NTF looks; in fact, it seems that the quantization noise in the bandwidth of interest is not suppressed, therefore making it challenging to reach the SQNR specifications.
This phenomenon is easily explained: the idea on which the current-feedback architecture is

based is exploiting the large parasitic capacitor provided by the transducer as an integration device. However, at resonance (where the impedances of L and C cancel each other), this capacitor has an impedance of:

$$Z_{C_s}|_{f=5MHz} = \frac{1}{sC_s} = \frac{1}{2\pi \cdot 5MHz \cdot 2pF} \approx 16k\Omega \tag{2.2}$$

This impedance is too high with respect to the transducer's intrinsic resistance (9448Ω), therefore one can conclude that the integrator behaviour for the first stage only shows itself at frequencies higher than the bandwidth of interest, being that the integrator is too "leaky" around resonance. In this case, most of the noise shaping comes from the second stage, therefore exploiting the transducer this way does not guarantee many advantages over more traditional architectures.
A way to compensate for this unwanted shape of the NTF is replacing the second stage with a more complex filter, designed such that it cancels out the unwanted poles and zeroes added by the transducer's resonance. A resonant band-pass filter was chosen for this purpose, the results are clearly visible in Figure 2.4: the quantization noise spectrum does not show any upwards peak around resonance, therefore the SQNR is drastically improved. However, building such a resonant band-pass filter introduces many accuracy and matching constraints, which are hard to meet in an area such as the one described in Section 1.4. To conclude, this architecture does not seem to bear many advantages over traditional architectures: it saves some area because of the "external" integration capacitance, but it yields a sub-optimal STF and NTF, a complex modulator and a convoluted compensation scheme.

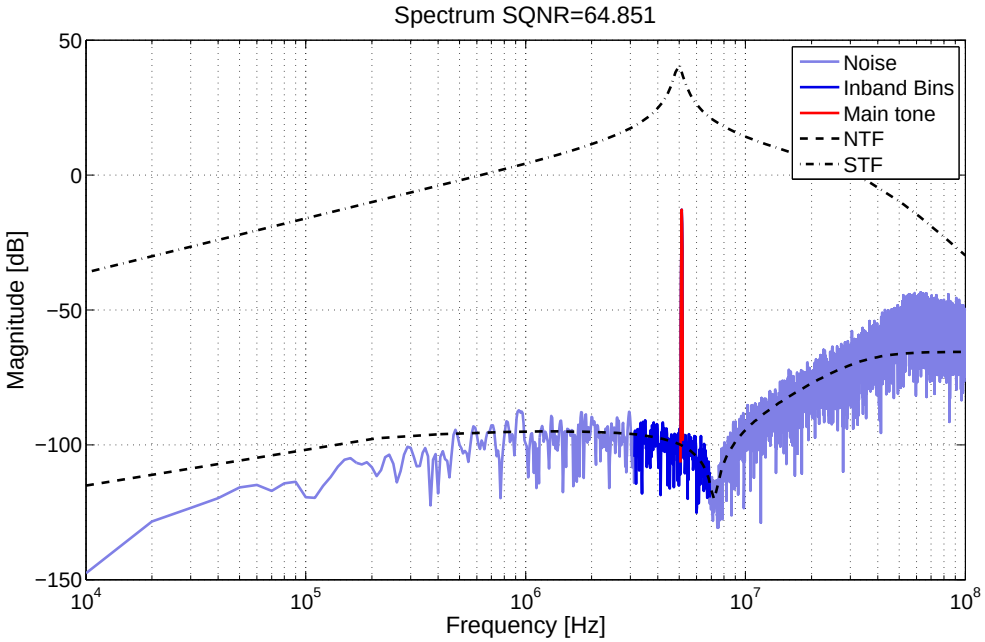


Figure 2.4: Spectrum of the compensated current-feedback structure

2.3. Voltage-feedback structure

As was argued in Section 2.2, around the resonant frequency, the dominant impedance offered by the transducer is that of the RLC branch. Therefore, naturally this would be the component that should be used as a loop filter. In order for the voltage-feedback structure to work, a low impedance node is needed to attract the current signal from the transducer, as well as the

current that results from the voltage feedback divided by the transducer impedance.
The easiest and best known technique to create a low impedance node is using the virtual ground of an amplifier in a negative feedback configuration. The feedback signal coming from the DAC has to be directly interfaced with the transducer's only available node; this can be achieved in two ways, depicted in Figures 2.5 and 2.6. The first schematic seems simpler at first, however it relies on a low impedance, floating voltage DAC, which is very hard to implement using known circuit techniques.
On the other hand, the system in Figure 2.6 is more realistic in terms of implementation, as the DAC in this case does not need to provide a low impedance and one of its terminals is connected to ground. However, it does have some downsides:

- The DAC voltage has to be processed by the amplifier before it can reach the inverting terminal. Since the amplifier has a limited bandwidth, the resulting signal at the inverting terminal will experience delay at frequencies higher or lower than the amplifier's bandwidth.
- From the DAC's point of view, the system resembles an OPAMP in a non-inverting configuration, therefore, at the output of the amplifier, there will be an additional factor, which corresponds to the DAC voltage, not shaped and not gained up. This factor, although not necessary, acts as a local feedback, as it can be seen in Figure 2.6, improving the modulator's stability and tolerance to excess loop delay (ELD).

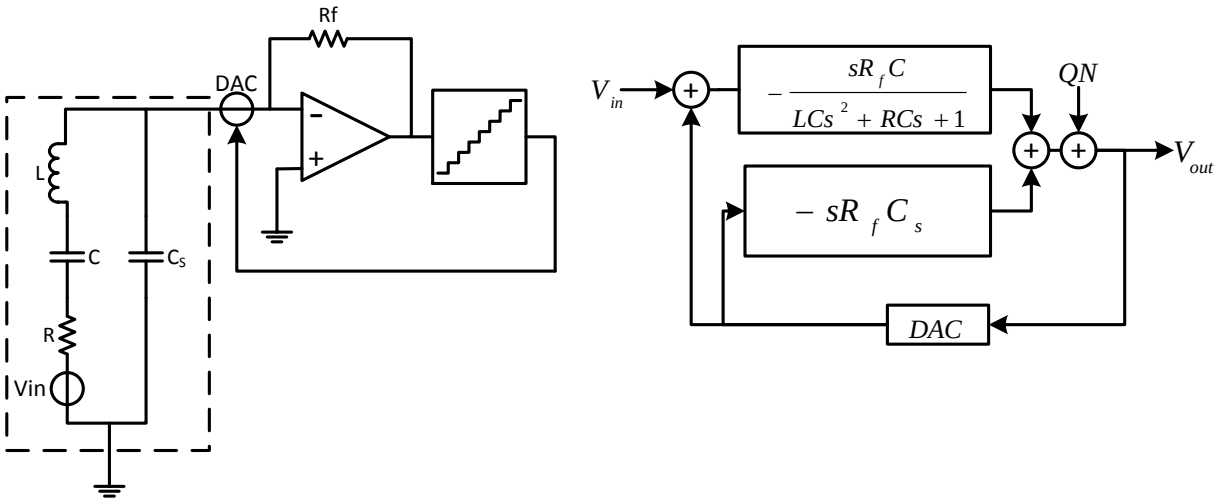


Figure 2.5: Representation of the voltage-feedback structure (left) and corresponding linear model (right)

2.3.1. Design Challenges and Constraints

The chosen architecture, along with its many advantages, also introduces some unwanted effects, which should be compensated for at system level. The most relevant issues derive from the transducer's parasitic capacitance C_s:

- C_s creates an unwanted "branch" in the loop filter of the modulator (see Figure 2.6), which goes in parallel with the main noise-shaping transfer function. The additional path adds a differentiated version of the feedback signal to the output of the amplifier. The feedback signal looks like a staircase (as the outputs of many multi-bit DACs), therefore

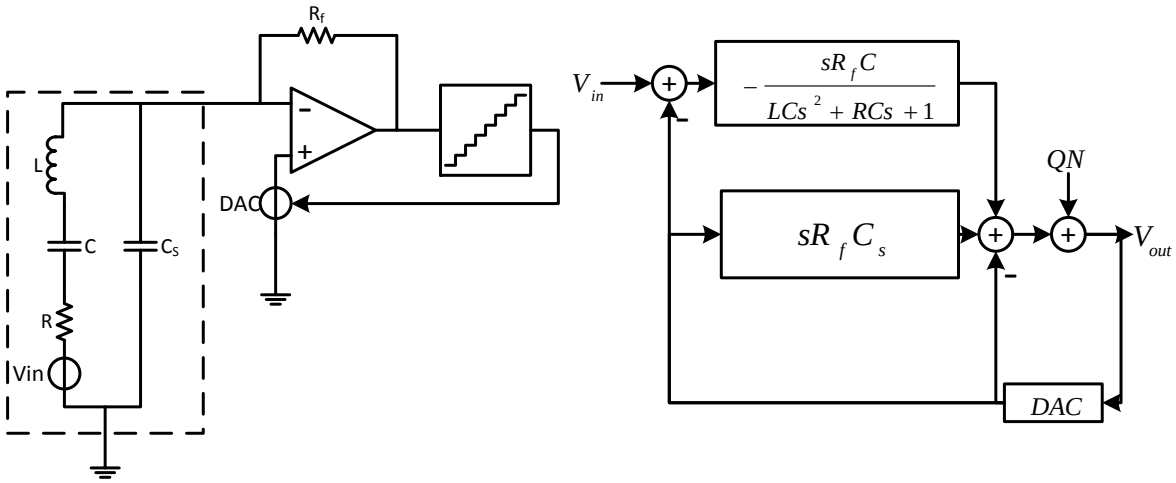


Figure 2.6: Representation of the voltage-feedback structure with a realistic position for the DAC (left) and corresponding linear model (right)

its spectral content also includes very high frequencies. If this signal gets differentiated, the result is a series of positive and negative spikes which will overlap with the desired output of the amplifier.

- Let us examine the transfer function of the noise introduced by the amplifier and by the DAC, after opening the ΣΔ loop (this situation is shown in Figure 2.7). The result is shown in Figure 2.8: the transfer function converges to the specified value for A, the open-loop amplifier gain. This is because, at high frequencies, the big capacitor C_s behaves like a short, effectively forcing the amplifier to work in an open-loop fashion. A high value for the amplifier gain guarantees that the input impedance of the trans-impedance amplifier is low and that the DAC voltage on the positive terminal is precisely copied at the negative one, however, at the same time, it increases the thermal noise gain at high frequencies. This noise, which should be the main noise contributor of the entire system, can end up having an rms value much greater than that of the intended signal, thus it will dither the quantizer, making its output only weakly dependent on the acoustic input signal.

An easy way to solve the second problem that was brought up is adding a low-pass filter between the amplifier and the quantizer. However, by doing that, it is very easy to compromise the modulator stability, as a low-pass filter shapes the quantization noise in an unwanted manner. More importantly, the current spikes generated by C_s and converted in voltage by the feedback resistor will also have to go through this hypothetical low-pass filter. This causes the output of the amplifier to contain a series of slow (depending on the cut-off) exponential decays with a much higher swing than the small signal coming from the transducer.
Therefore, adding a low pass filter in order to attenuate the thermal noise of the amplifier and the DAC, amplified by the relative transfer function, has to be combined with a way to eliminate the previously discussed current spikes. This will be the main topic of the next Section.

2.4. Transducer's Capacitance Compensation

2.4.1. Negative Capacitance Circuits

A way to compensate for the transducer's parasitic capacitance is adding a "negative capacitance" of the same value in parallel with it. The resulting shunt impedance, in case of no

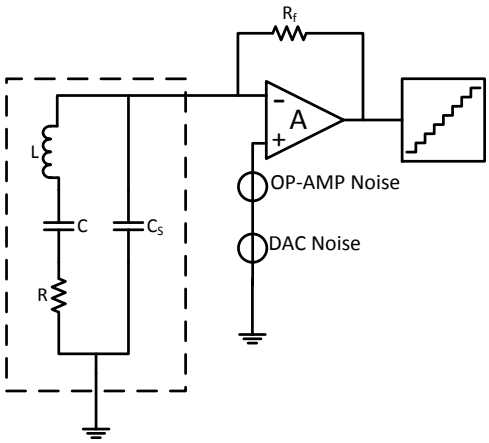


Figure 2.7: Schematic of the voltage-feedback structure without the ΣΔ feedback.

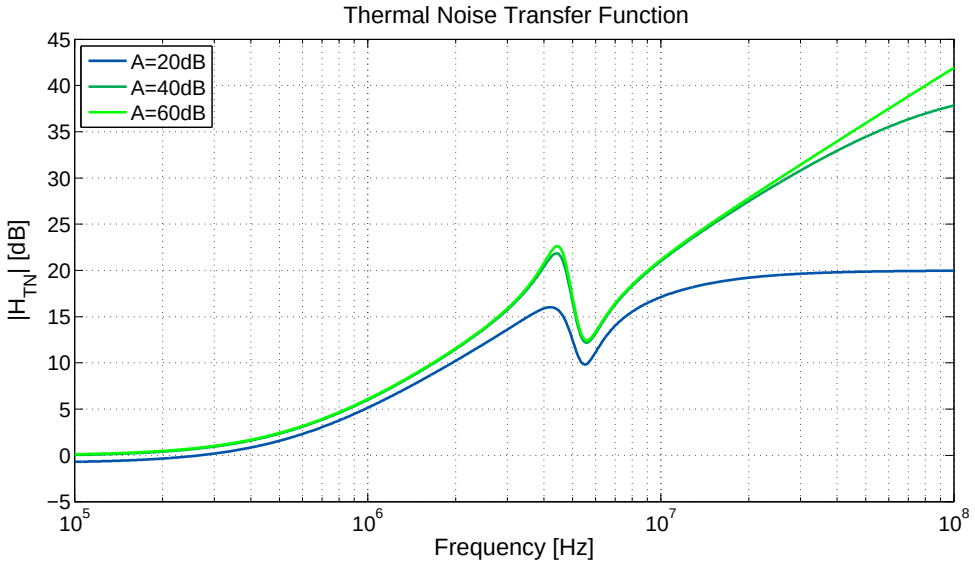


Figure 2.8: Bode diagram of the thermal noise transfer function, at different values of the amplifier's open-loop gain.

mismatch, will be $Z_{eq} = (sC_s + s(-C_s))^{-1} = \infty$.
Of course, negative capacitances cannot be obtained using passive components, but a circuit with a negative input capacitance can be achieved using positive feedback, as explained in Figure 2.9. An auxiliary amplifier is added, in order to provide the loop gain needed for the positive feedback. The input impedance of such a configuration can be shown to be $Z_{in} = -\frac{R_1}{R_2}\frac{1}{sC_c}$. The capacitance C_c can therefore be scaled down, in order to save silicon area. This solution automatically solves both the problems described in the previous section, however the price to pay is too high: in fact, the auxiliary amplifier should be extremely fast, in order to source the current needed to compensate the one that arises from the parasitic capacitance, before it flows in the feedback of the first stage amplifier. This solution also requires good matching between the parasitic capacitance and the corresponding negative capacitance: in case the latter becomes bigger in absolute value than the former, even at a frequency that is out of the bandwidth of interest, the modulator could become unstable.

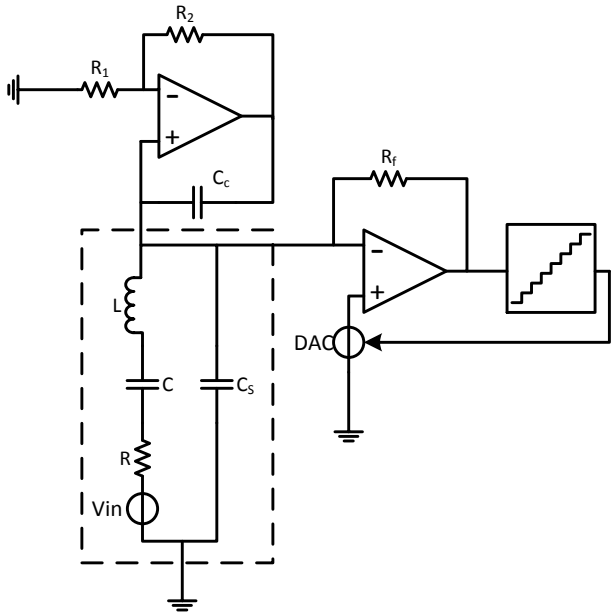


Figure 2.9: Schematic of the voltage-feedback structure with negative capacitance compensation.

2.4.2. Switched Feedback Configuration

Another way to address the problem of the current spikes is illustrated in Figure 2.10. Here, the feedback resistor gets shorted when the DAC changes its value. In this case, the current needed to charge up C_s to the new level is still provided by the first stage amplifier, however, it will not cause a voltage drop across the inverting input and the output of the amplifier. This solution does not need any matched components. The main shortcoming for this configuration is that the output of the first stage will not be continuous any more, as, at each clock cycle, the output will be reset to the DAC voltage; in other words, the transimpedance gain becomes zero for a certain time. This means that the amplifier's $3dB$ point needs to be significantly larger than the clock frequency, in order to be able to guarantee an acceptable settling behaviour in a fraction of the clock period. Such a speed (a bandwidth in the order of $1GHz$) is only reachable by spending an unacceptable amount of power. Additionally, the amplifier needs to charge a potentially large capacitance in a very short time, thus causing slew rate issues.

Furthermore, this switching scheme does not intrinsically solve the noise gain problem and, in fact, only moves its solution further out of reach. Since the output of the first stage is not continuous, its spectral content becomes much wider than that of the signal of interest, therefore, also the low-pass filter needs to be switched, together with any additional circuitry between the first stage and the quantizer. Thus, the resulting loop filter would show the disadvantages of both a continuous-time and a discrete-time modulator, as the sampling operation happens at the input of the loop filter, rather than the output, like standard $CT\Sigma\Delta M$.

2.4.3. Delayed Feedback Configuration

A delayed-feedback configuration (shown in Figure 2.11) was also taken into consideration to address the previously stated problem. In this architecture, the feedback is not directly provided to the first stage: instead, in ϕ_1, the transducer is pre-charged to the new level of the modulator's feedback DAC, while in ϕ_2 the transducer is newly connected to the first

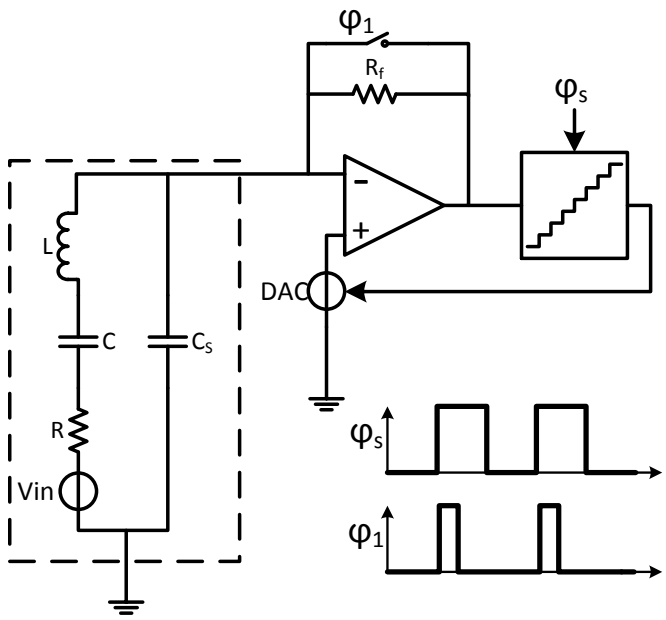


Figure 2.10: Schematic of the voltage-feedback structure with time windowing compensation.

stage, while the latter also receives the new voltage level from the DAC. At the beginning of ϕ_2, the voltages at the input terminals of the amplifier will be equal, if one does not take into consideration the non-idealities introduced by the switching operation. The benefit of this structure is the fact that the first stage does not have to provide the current needed to charge up the transducer's capacitance, as that will be given by the DAC instead. However, the $\frac{kT}{C}$ noise associated with the three added switches will appear at the two most critical nodes of the circuit: the input terminals of the amplifier.

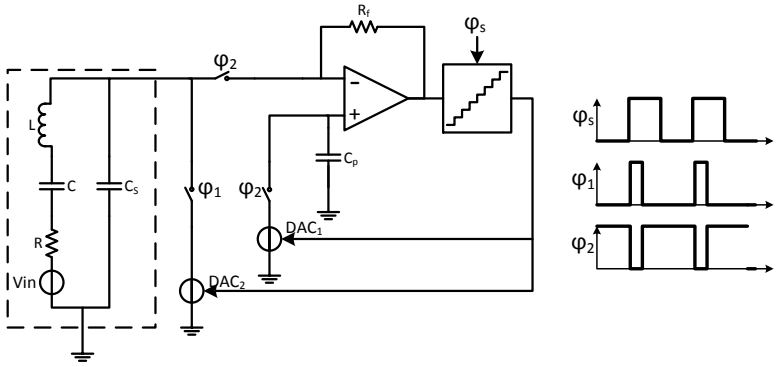


Figure 2.11: Schematic of the voltage-feedback structure with delayed feedback compensation.

2.4.4. Compensation Capacitance Configuration

The chosen way to compensate for the parasitic capacitance C_s was adding another explicit capacitance (C_{comp} in Figure 2.12), connected to C_s on one side and to a second DAC, synchronous but scaled with respect to the main one, on the other side. A similar "anti-resonance

cancellation" technique was employed in [29]. The function of this added capacitance is to provide the current needed for C_s to be charged to the DAC's new level. In this solution, this current does not come from the amplifier, therefore it does not cause a voltage drop across the feedback resistance. No switching is involved for this method, and no added noise either, except for the second DAC's noise, which can be kept low without a penalty in power consumption.

In the right part of Figure 2.12, another version of the same solution is proposed: in this case, the two DACs are combined into a single element, which will take care of providing the feedback to the modulator, as well as performing the capacitance compensation as explained above.
The modulator feedback is obtained by a simple capacitive divider between C_{ref} and C_c, while a copy of C_{ref} connected to the inverting input of the amplifier provides the compensation current. This configuration is similar to the well-known Wheatstone bridge. The capacitance C_c should be tuned so that its value is the same as C_s, this way, the bridge is balanced and the two amplifier inputs will respond to the DAC signal moving together, thus avoiding to create a differential signal that could reach the output. The tunability of C_c and of the reference voltages of the DAC provides enough flexibility to face the variability of the transducer characteristics.

Unfortunately, this compensation scheme does introduce some vulnerabilities: the DAC voltage and the compensation capacitor need to be sized carefully. As a matter of fact, if the additional compensation current is larger in absolute value than the one which it should cancel, the feedback becomes positive and the modulator will quickly become unstable. In other words, if C_c is larger than C_s, when DAC changes its value, the amplifier's non-inverting input will see a lower voltage variation with respect to the inverting input. This causes the amplifier output to have an opposite polarity with respect to the DAC voltage, which causes the positive feedback.

Furthermore, this topology only solves the current spike issue that was discussed in Section 2.3.1, but does not affect the noise issue related to the capacitance C_s. As a matter of fact, this effect will be exacerbated by adding C_{ref}, which effectively sits in parallel with C_s. For this reason, high frequency noise should be extrinsically filtered before reaching the quantizer.

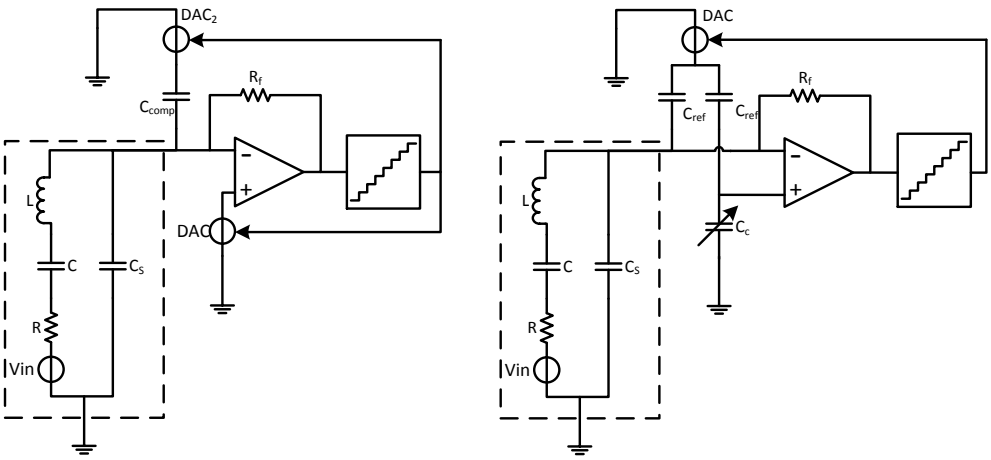


Figure 2.12: Schematic of the voltage-feedback structure with added compensation capacitance.

2.5. Design Flow

As explained in the previous sections, the basic idea on which this project is based is reasonably simple. Nonetheless, its implementation introduced many challenges to overcome and many trade-offs to ponder. In particular, the peculiarities of this modulator make it unfit for most of the techniques commonly used in order to derive and optimize the loop filter coefficients. The usual synthesis of a continuous-time, band-pass modulator follows the steps described below:

- Choose the oversampling ratio and therefore the sampling frequency and the nature and number of bits of the quantizer, so that the SQNR specification can be reached.
- Select one of the standard modulator topologies, such as feedforward and feedback.
- Synthesize an optimal NTF for the required specifications, in the form of a discrete-time, low-pass transfer function.
- Convert the low-pass modulator into its respective band-pass version, using the well-known transformation $z \rightarrow -z^2$, which conserves many of the characteristics of the original converter, most importantly stability and SQNR.
- Obtain the continuous-time implementation (in terms of loop filter coefficients), applying the Impulse Invariant Transform, or equivalent, to the discrete-time transfer function.

The main reason why this procedure could not be applied is that the design space has many intrinsic constraints that are forced by the transducer's coefficients: as a matter of fact, the loop filter transfer function will be dominated by that of the transducer, especially in the case of a first-order modulator. These constraints are defined for a $CTBP\Sigma\Delta M$, but should be somehow accounted for at the beginning of the synthesis chain, in a discrete-time, low pass environment.
Furthermore, the $z \rightarrow -z^2$ transformation can only be applied if $f_c = \frac{f_s}{4}$, where f_c is the center frequency and f_s the sampling rate. This requirement introduces a heavy limitation in the OSR selection. In this work, f_c and the desired bandwidth are specified by the application, as they are related to the resonance frequency and the Q factor of the piezoelectric element. A sampling frequency of $f_s = 4f_c = 20MHz$ would limit the OSR to $OSR = \frac{f_s}{2 \cdot BW} = \frac{20MHz}{2 \cdot 3.75MHz} = 2.7$, which would leave no space in the frequency domain for the noise shaping to suppress quantization noise.
Finally, the modelling of a tracking quantizer (see Section 2.5.4) in the Z domain can be quite bothersome, as one has to account for the peculiar kind of non-linearity that is typical for this quantizer.

For these reasons, a different, more generic design flow was chosen for the task of selecting an architecture: the major trade-offs were evaluated one by one trying to implement a working prototype on Matlab/Simulink for each situation and evaluating the obtained results and the relative costs, mainly in terms of power and area. The parameter tuning, needed to reach the best performance, was done through a simple genetic algorithm, which automatically randomizes the variables of the system starting from the last best result obtained, seeking an optimum set of parameters. Boundaries were added on the variables that define the critical specifications of the system, such as power consumption and silicon area.
The rest of Section 2.5 will focus on the main arguments that were used in order to converge towards a specific solution within a very wide design space: a continuous-time, first order, 3-bit band-pass $\Sigma\Delta$ modulator, employing a tracking quantizer.

2.5.1. Bandpass vs. Low-pass Modulators

The reason why a band-pass architecture was preferred over a low-pass one is described in detail in Section 2.3. In practice, this is the only implementation which allows us to advantageously use the transducer as a loop filter. However, it is beneficial to briefly describe the features of interest in a band-pass structure. Such a modulator, especially in its continuous-time flavour, has been heavily featured as the converter of choice for the direct digitization of RF signals in telecommunication applications ([30], [31]): in this field, converters require a generally low bandwidth centered around a high frequency carrier. As a matter of fact, the $BP\Sigma\Delta M$ is able to focus its performance on a specific band of interest. On the other hand, Nyquist-rate ADCs and $LP\Sigma\Delta Ms$ usually have a bandwidth that extends from $0Hz$ to the upper edge of the band of interest, effectively wasting power over a bandwidth that does not contain any desirable signal [32]. In this aspect, the ultrasound world is similar to that of telecommunications, however, while in the latter the bandwidth is much lower than the center frequency, in the ultrasound case these two frequencies are comparable, mitigating the advantage described above.

2.5.2. Continuous-time vs. Discrete-time Modulators

Above all, the reason why a CT implementation was selected is that the filtering characteristics of the transducer are naturally continuous-time. The arguments provided below also prove that a CT modulator is the preferable candidate for this application. First of all, one of the main advantages of CT architectures with respect to their DT counterparts is that, in many cases, they do not require an anti-aliasing filter. This can be easily understood, considering that, in case of CT modulators, the sampling operation only takes place once the signal has reached the quantizer. Thus, the loop filter effectively acts as an anti-aliasing filter too. On the contrary, DT loop filters rely on explicit anti-aliasing filtering to suppress out-of-band spurious components. This advantage becomes especially important for low OSR, high frequency modulators, such as the one described in this thesis, because the requirements of an adequate anti-aliasing filter would become very strict. [33] Additionally, the amplifiers needed for the switched-capacitor circuits in DT modulators need to have a higher bandwidth with respect to the clock frequency, so that they can safely settle within one clock period. This means that these amplifiers will have to burn more current and, more importantly, that the noise bandwidth of the system will be high compared to CT modulators, causing high frequency noise to fold in the bandwidth of interest.

DT modulators, on the other hand, show an improved linearity performance, as they are not limited by the intrinsic non-linearity of integrated resistors or g_m stages used in CT loop filters. However, linearity is not the most stringent specification to meet in this work.
Another desirable feature of DT modulators is the fact that their transfer function in the Z domain is independent of frequency, while in their CT counterpart, the design has to be centered at a specific sampling frequency. The tolerance of DT modulators to delay in the feedback path, as well as jitter, is also superior. Finally, the accuracy of the DT loop filter coefficients is defined with a capacitance ratio, therefore they can be very precise. Unfortunately, the limiting factor for the accuracy of the modulator coefficients in this design is that of the components in the transducer's electrical model.

2.5.3. Multi-bit vs. Single-bit Quantizers

A multi-bit quantizer was selected for this design, mainly because it helps in reducing the quantization noise to be suppressed by the noise shaping, thus enabling the modulator to

reach a higher SQNR. A multi-bit quantizer also yields many other advantages over a single-bit comparator:

- The error signal that enters the modulator has a much lower root mean squared value, as the LSB becomes exponentially smaller for every bit added to the quantizer: $\epsilon_{rms} = \frac{V_{ref}}{(2^N-1)\sqrt{12}}$.
- Because of the reduced swing of the error signal, the implementation of the loop filter becomes less constrained in terms of swing, and the distortion is limited.
- The modulator behaves in a less chaotic manner, because the magnitude of the feedback is proportional to that of the input of the quantizer, and not only to its polarity, therefore it tracks more closely the modulator's input.
- The quantizer gain is well defined and linear (if the modulator is not overloaded). This feature improves the stability performance of the modulator, as it allows for a higher input amplitude, compared to the DAC reference voltage(s).
- The jitter sensitivity is reduced [33], especially if combined with a Non-Return to Zero DAC. This happens because the influence of each feedback step, and its timing, becomes less dominant in the modulator performance.

All of the advantages listed above have been exploited in this design: the swing of the internal nodes of the modulator is kept low; the fact that the converter is less chaotic enables the use of a tracking quantizer, saving area and power; the stability is less dependent on the parameters of the input signal and on the transducer characteristics. Finally, the modulator is not very sensitive to jitter, as shown in Section 2.6.3.
The advantages of single-bit architectures are mostly the reduced complexity and the inherent linearity of a 2-level comparator and DAC. However, the ultrasound image quality is typically not distortion-limited, therefore the use of a single-bit quantizer is not justified. The increased complexity of a multi-bit quantizer has to be dealt with carefully, in order to avoid a proportional increase in power dissipation and silicon area. This problem is discussed in detail in Section 2.5.4.

2.5.4. Tracking Quantizer

The obvious pitfall of introducing in any modulator a multi-bit quantizer is that the increase in resolution must be paid with a higher power, area and complexity. The standard and most used quantizer architecture for all kinds of oversampled ADCs is the flash topology. Their popularity is due to the fact that $\Sigma\Delta$ modulators rely on a discrete-time feedback that must be provided, when possible, immediately after the quantizer reaches a decision, with virtually no delay. Flash ADCs only need one clock edge (plus the comparators' delay) per conversion, thus making them fit for this application. However, the power and area consumptions for such converters are proportional to 2^N, where N is the number of bits. Even a simple 3-bit flash quantizer needs 7 comparators, clocked at every cycle. If the sampling frequency is high, the power consumed by the comparators can be easily compared to that of the low-noise, high-accuracy analog blocks in the modulator, which is undesirable at best.

A solution to this problem is employing SAR quantizers instead [34]. This allows for a substantial reduction in area and power , together with superior scalability, but the downside is that SAR ADCs lack in the most important feature for a quantizer embedded in a $\Sigma\Delta$ loop: small latency. In fact, for each conversion they need a number of clock cycles equal to the

converter's number of bits. This means that, in order for the modulator to be stable, the SAR quantizer should be clocked at a speed of N times higher than the main $\Sigma\Delta$ loop. Such a speed would not be practical, and it would yield serious implementation issues in a $0.18\mu m$ technology. Furthermore, the comparator would be clocked three times per cycle, and should have a very low propagation delay, thus a high power consumption; also the SAR logic, at the speed in question, would have a far from negligible power consumption.

A way to achieve multi-bit quantization with a minimal increase in power consumption with respect to a single comparator is by means of a tracking ADC ([35], [36], [37], [38]): this converter architecture combines the advantages of the previously mentioned architectures. It only needs one clock cycle per conversion and the number of comparators is fixed with respect to N. The working principle of a tracking quantizer is based on the fact that each conversion is to be interpreted with respect to the previous cycle, rather than in absolute terms, similarly to Delta modulation. A practical way to achieve this is adding a feedback loop around the comparators, through some DACs, which updates their references based on the last outputted value.

In the specific implementation that was adopted in this design, captured in Figure 2.13, only

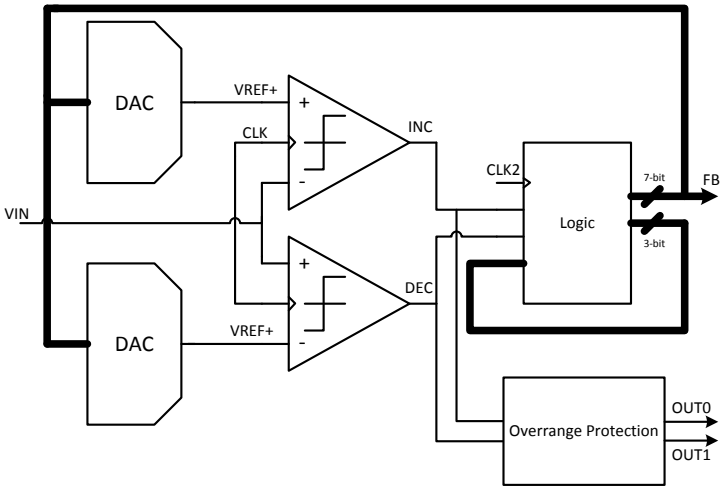


Figure 2.13: Block diagram of the proposed tracking quantizer.

two comparators were employed: the purpose they serve is to keep the input signal between an upper and a lower reference voltage, which get updated every cycle, as show in the top part of Figure 2.14. After the comparators make their decision, if they are both low, the references will not be updated. On the other hand, if the top comparator reaches a high value, it means that the input voltage was determined to be higher than the upper reference. The opposite is true for the bottom comparator. From the outputs of these two comparators (depicted on the bottom half of Figure 2.14) it is possible to reconstruct the quantized input signal Q, provided that the input stays within the quantizer's full scale range:

$$Q = \sum_{n=1}^{N} INC(n) - DEC(n) \tag{2.3}$$

If this is not the case, a simple overrange protection circuit will make sure that $OUT0$ and $OUT1$ stay low after the quantizer has reached respectively the highest and the lowest values. A 3-bit logic cell, similar to an up-down counter, calculates the multi-bit output of the modulator

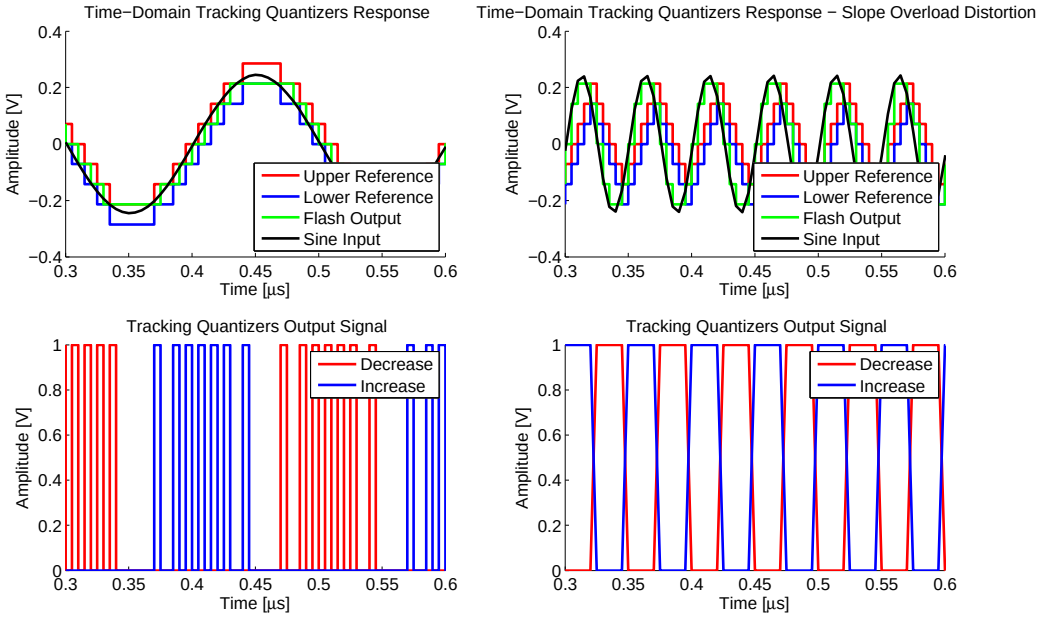


Figure 2.14: Visual explanation of the tracking quantizer principle and of Slope Overload Distortion.

(based on the comparators' decision and the stored value of the previous cycle) which is then fed back to the DACs and to the modulator's input node. A careful eye could notice that the second DAC in the figure is redundant, as its digital input is shared with the first and the output is always one LSB apart with respect to that of the first DAC. However, both DACs made the final design, for the sake of simplicity and better matching.

Slope Overload Distorsion (SOD)

The price to pay for such an efficient modulator is that, in the way it is implemented in this design, it can only tolerate a change in the input signal of one LSB per cycle. This poses a limitation on the maximum slope that the input signal may have, and causes the output to be distorted if this requirement is not met. This kind of distortion is called Slope Overload Distortion, and its characteristic shape is shown on the right part of Figure 2.14. Of course, increasing the number of bits without changing the sampling frequency will result in a slower response, and SOD will happen at lower frequencies. In mathematical terms, the maximum slope of the signal should be lower than the LSB divided by the sampling period:

$$\max\left[\frac{d}{dt}A\sin\omega t\right] < \frac{LSB}{t_s} \rightarrow \omega A < \frac{V_{ref}}{(2^N-1)}f_s \rightarrow 2^N < \frac{V_{FS}f_s}{2\pi f_{in}^{max}A} + 1 \tag{2.4}$$

Assuming the input amplitude to be maximum, and equal to $\frac{V_{FS}}{2}$ and $OSR = \frac{f_s}{2f_{in}^{max}}$ this equation can be further simplified into $2^N < \frac{2}{\pi}OSR + 1$. This result further justifies the choice of a tracking ADC as the quantizer in a $\Sigma\Delta$ loop, which is, by definition, oversampled. However, the designer should choose an NTF that does not allow excessively chaotic behaviour, else SOD will affect the noise shaping characteristic. In this design, system-level simulations of the full modulator, including many imperfections, were performed with a standard flash quantizer and with a tracking ADC. The results in terms of SNR were found to be the same, and the probability of the flash quantizer to change its value of more than one step per cycle was calculated to be 0.005%.

Table 2.1: Possible design choices for similar resulting SQNR performance

Sampling Frequency (f_s)	200MHz	120MHz	170MHz	80MHz	350MHz	750MHz
Oversampling Ratio (OSR)	26.6	16	22.6	10.7	46.6	100
Filter Order (L)	1	1	2	2	1	1
Quantizer Bits (B)	3	4	1	3	2	1
Estimated SQNR	56.3dB	56.2dB	56.6dB	57.2dB	56.2dB	56.6dB

2.5.5. Design Parameters

Input Amplitude

The modulator's input amplitude can be derived from the noise associated with the resistance in the electrical model of the transducer. As discussed in Section 1.4, the desired minimum signal to noise ratio is $50dB$ in a $3.75MHz$ bandwidth. Let us also assume that the total input referred noise of the ADC should be equal to that of the resistor, so that, independently, they would yield a $53dB$ signal to noise ratio. The minimum input signal amplitude, modelled as a sinusoidal voltage source in series with the transducer's resistor ($R = 9448\Omega$), can be computed as:

$$A_{in}^{min} = 10^{\frac{53dB}{20}} \sigma_n \sqrt{2} = 10^{\frac{53dB}{20}} \sqrt{4k_B TR \cdot BW}\sqrt{2} = 15.3mV \tag{2.5}$$

The maximum input amplitude that the modulator should be capable of processing is the minimum one multiplied by the TGC range, which is assumed to be $25dB$:

$$A_{in}^{max} = 10^{\frac{25dB}{20}} A_{in}^{min} = 272.5mV \tag{2.6}$$

Sampling Rate, Modulator Order and Quantizer Bits

In order to determine the oversampling ratio, and therefore the sampling frequency for the modulator, a few assumptions have to be made. The signal to quantization noise ratio ($SQNR$) should be between $5dB$ and $10dB$ greater than the target SNR, so that the resolution of the modulator can be fully thermal-noise limited. In order to get a rough estimation of the $SQNR$ that can be achieved with a specific set of design parameters, Equation 2.7 was used from [39], where L is the modulator order and B is the number of bits for the quantizer.

$$SQNR = 10\log\left(\frac{3(2^B-1)^2(2L+1)OSR^{2L+1}}{2\pi^{2L}}\right) \tag{2.7}$$

In Table 2.1, a few sets of parameters, achieving an acceptable SQNR, are shown.
It is important to notice how the oversampling, independently of the additional noise shaping, aids the SQNR. As a matter of fact, an ADC with only 3 bits and no noise shaping (such as the quantizer in this design) sampled at $200MHz$ yields a theoretical SQNR of:

$$SQNR = 6.02B + 1.76 + 10\log\left(\frac{f_s}{f_N}\right) = 6.02B + 1.76 + 10\log\left(\frac{200MHz}{2\cdot 3.75MHz}\right) = 34dB \tag{2.8}$$

This resolution is much higher than the $20dB$ achieved by the same ADC clocked at Nyquist-rate.

The last two columns of Table 2.1 have to be discarded as they require a very high sampling frequency, which would make the quantizer (even if only composed of one comparator) much more power hungry than the analog part of the modulator. The second column was also not thoroughly investigated, as a 4-bit quantizer cannot be implemented with a tracking ADC, because the constraint described by Equation 2.4 is not satisfied. We are left with three possible implementations, which were carefully simulated at system level:

1. First order, 3-bit with tracking quantizer, $200MS/s$
2. Second order, 3-bit, $80MS/s$.
3. Second order, 1-bit, $170MS/s$.

First of all, it is worth to notice that a second order $CTBP\Sigma\Delta M$ would require the implementation of a second resonator, for example the one proposed in [40]. In order for it to be effective, it needs to have a moderately high Q factor (5-10), so that it significantly lowers the quantization noise in the bandwidth of interest. This requires high gain $(40 - 50dB)$ around the resonance frequency in the amplifier embedded in the resonator, which will therefore have a UGBW above $500MHz$. This requires the amplifier's g_m to be big in order to accommodate such a large bandwidth, therefore the second stage could consume more power than the first, depending on the loading. For example, the resonators implemented in [40], for a modulator with similar specs, consume over $1mW$ each.
Stability should also be guaranteed in a second order system, which would require at least a feed-forward path, a summation node just before the quantizer, and a feedback DAC around the quantizer, as without the latter, the system does not tolerate any ELD. On the other hand, a second stage would intrinsically filter out much of the high frequency noise added by the first stage, therefore relaxing the noise specifications for it. However, this advantage does not necessarily reduce the power consumption of the first stage, because its g_m should be high enough to reach its bandwidth requirements.

Let's consider in particular case 2: here, the main advantage is a reduced output sample rate, and therefore a lower digital power. However, in this case, a tracking quantizer should not be employed, even if Equation 2.4 is almost met. This is because this converter does not support a chaotic behaviour (see Section 2.5.4), as it cannot keep up with high frequency, high amplitude changes in its input. This means that the quantization noise cannot be shaped up to very high frequencies and instead the NTF will flatten a bit before the Nyquist frequency. Since there is less than a decade between the resonance frequency and the Nyquist frequency, the quantizer needs to be chaotic, i.e. toggle at almost every clock cycle. If this doesn't happen, the predicted SQNR cannot be reached. An acceptable performance is achieved at $f_s = 150MHz$, but this sample rate is very similar to that of case 1), thus making the advantage of the second stage negligible. Also, using a flash quantizer would yield a total data rate comparable to case 1), as it would generate 3 bitstreams at $80MHz$ each, serializable to $240MHz$, against the 2 "1.5 bit" bitstreams at $200MHz$ of 1). This is also the most expensive option in terms of design time as it is the most complex to implement.

Let's now consider case 3: The main advantage is having a single bit output. This makes the DACs and the successive digital electronics easier to implement and reduces the digital power significantly. Also, since there is only one threshold in the quantizer, the loop gain in theory doesn't need to be dynamically adjusted by the time-gain compensation mechanism, although the system should be designed to be stable for a wide range of input amplitudes, which is not trivial. We can also see this as the quantizer not having a defined gain. This gain uncertainty, together with the phase uncertainty of a single bit quantizer and the low OSR requires a sampling frequency that is higher than the predicted one in Table 2.1 in order to reach the given SQNR. System level simulations suggest a sampling rate around $250MHz$. Again, this is comparable to case 1) and requires a faster quantizer, instead of two slower ones. Another disadvantage is that having only one bit, the error signal entering the sigma-delta loop has a high swing, as discussed in Section 2.5.3.

Ultimately, solution 1 was selected and implemented in silicon. The high sampling frequency is needed to reach the desired SQNR, however, since it is not a very chaotic system, the digital output only changes on 37% of the clock cycles on average, therefore reducing the digital power and adding space for some compression in future implementations. Further power and area are saved as this solution does not require a second resonator. The main disadvantage here is the extrinsic filtering that needs to be provided in order to attenuate the thermal noise of the first stage. This filtering does not introduce any power or area penalties, but it tames the naturally chaotic behaviour of the modulator, sacrificing SQNR. For this reason, there is a direct trade-off between SNR and SQNR, as they cannot be optimized separately, thus the SQNR will be rarely discussed, in favour of its more realistic version, which includes thermal noise. In this implementation, the noise shaping function is entirely up to the transducer, making such a system intrinsically matched with the piezoelectric element's characteristics, but also basing the modulator's performance, particularly its stability, on the transducer modelling and on the repeatability of its fabrication steps. For these reasons, this design has a high risk factor, and should be made as insensitive as possible to transducer variations.

2.5.6. Architecture of Choice

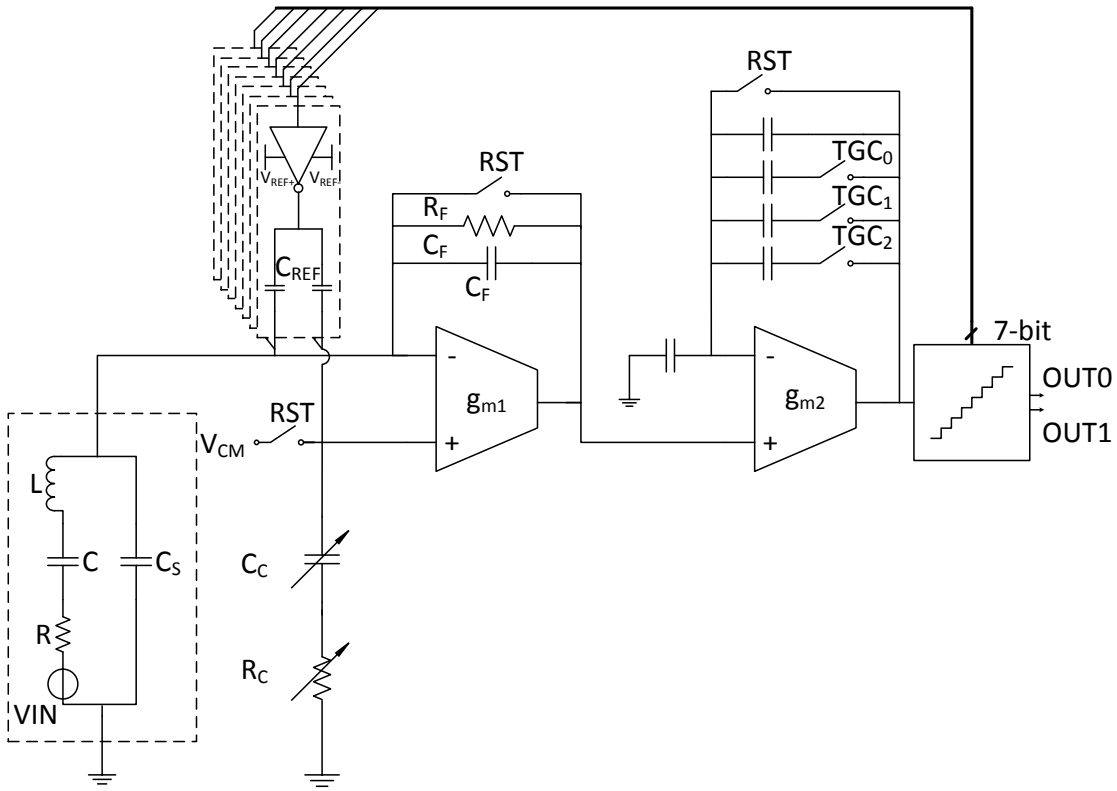


Figure 2.15: Simplified block-level schematic of the chosen architecture.

In this Section, more details will be given on the block-level implementation of the proposed design, shown in Figure 2.15, before explaining the transistor-level topologies of each fundamental block in Chapter 3. The starting point is shown in Figure 2.12. The DAC is implemented through a group of seven inverters: each one has two supplies, V_{REF+} and V_{REF-}, shared between all inverters and tunable from outside the chip. The advantage of this configuration is

that it only requires two reference voltages, rather than one per level, making the routing of the array much easier while still enabling multi-bit feedback. Furthermore, the linearity of such a DAC, which determines the distortion performance of the entire modulator, is only limited by mismatch in the capacitor banks. The inputs of the inverters are thermometer-coded 1.8V logic signals coming from the quantizer. In order to size the capacitance C_{ref} one can assume that when the logic level at the input of one inverter changes, the output will move from V_{REF+} to V_{REF-}, or vice versa. As a response to that, the DAC voltage has to change by one LSB:

$$V_{DAC}^{LSB} = (V_{REF+} - V_{REF-})\frac{C_{ref}}{C_s + 7C_{ref}} \rightarrow C_{ref} = \frac{C_s}{\frac{V_{REF+}-V_{REF-}}{V_{DAC}^{LSB}} - 7} \tag{2.9}$$

V_{DAC}^{LSB} can also be written as $\frac{2A_{in}k_{DAC}}{2^B-1}$, where A_{in} is the transducer's input voltage, while k_{DAC} is a coefficient, between 1 and 2, which accounts for the fact that the feedback voltage of a ΣΔ modulator should be slightly higher than its maximum input amplitude.

A tunable resistor R_c was added in series with the tunable capacitor C_c, in order to compensate for the transducer's high frequency behaviour (see Section 1.2).
The first stage has been implemented as an Operational Transconductance Amplifier (OTA), in order to save the power needed for the output buffer of an OPAMP, as a low output impedance or a superior driving capability is not required.
A feedback capacitor C_F was connected in parallel with the already discussed R_F. This is not to be intended as an integration capacitance, as its impedance in the bandwidth of interest is much higher with respect to R_F; the main functions of C_F are placing a known pole in the transfer function of the first stage, while providing a low impedance path for the high frequency currents that could arise from an imperfect capacitance compensation.
Furthermore, a second stage was cascaded after the first one. This stage serves multiple purposes:

- It provides the first stage with a low input capacitance to drive.
- It filters the high frequency noise, in order to avoid the negative effects discussed in Section 2.3.1.
- It adds a gain between the first stage and the quantizer, so that the latter may have a manageable full scale range, even when the input signal is very small, effectively reducing the impact of the comparator noise on the modulator's performance.
- It provides a variable gain, which can be tuned dynamically to accommodate for the changes in the input signal's amplitude during the receive phase, as explained in Section 1.4.
- Since the second stage provides much of the amplification needed before the quantizer, the first stage can have a lower gain, therefore reaching a higher bandwidth for the same power consumption.

Finally, the architecture now has a reset phase, where the common mode voltage V_{CM}, typically at $\frac{V_{DD}}{2}$ is forced on the non inverting pin of the first stage, while both stages' feedback is shorted by a switch, configuring them as unity-gain buffers. The reset phase also disables the clocks entering the tracking quantizer. This phase is needed to force the circuit in a known, stable initial condition, with all the floating nodes charged at the correct voltage, before actually starting the receive phase.

Table 2.2: Design parameters

First Stage	Second Stage	1^{st} Stage Feedback	Sampling	Modulator Feedback
$g_m = 3\frac{mA}{V}$ $R_{OUT} = 23k\Omega$ $C_{OUT} = 250fF$	$g_m = 0.4\frac{mA}{V}$ $R_{OUT} = 132k\Omega$ $C_{OUT} = 120fF$	$R_F = 100k\Omega$ $C_F = 100fF$	$f_s = 200MHz$ Bandwidth=75% $f_c = 4.98MHz$	$k_{DAC} = 1.05$ $V_{REF+} - V_{REF-} = 31mV$ $ELD = \frac{1}{2f_s}$

2.6. System-Level Simulation Results

2.6.1. Signal and Noise Transfer Functions

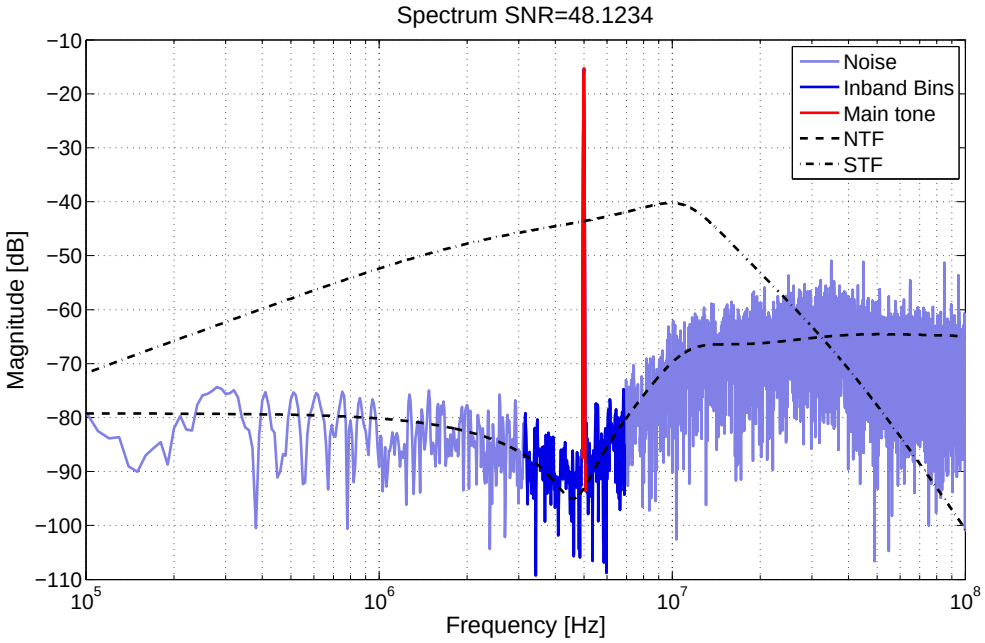


Figure 2.16: System-level Simulink simulation for the final architecture

A Simulink model of the system described in Figure 2.15 has been implemented and simulated (see Figure 2.16); in this representation, many non-ideal effects were introduced in order to evaluate their effect on the performance of the modulator, such as ELD and noise associated with R, R_F, g_{m1} and g_{m2}. Furthermore, the highest gain setting was selected, in order to consider the worst case scenario in terms of noise. The OTAs were modelled with a finite g_m and an output load formed by a shunt resistor and capacitor. The second stage in this simulation is modelled as a simple open loop transconductance. The values used in this simulation are shown in Table 2.2, and were found using the genetic algorithm described in Section 2.5.

The band-pass characteristic is evident. One might also notice how the quantization noise flattens, and actually starts to decrease, near $\frac{f_s}{2}$. This happens because the bandwidths of the first and the second sages do not reach $\frac{f_s}{2}$ as they would in a normal modulator. Also the tracking quantizer introduces an additional bandwidth limitation, as discussed in Section 2.5.4. As Figure 2.16 depicts, the behaviour shown by the Simulink model of the modulator in terms of noise transfer function matches quite well with the theoretical model extrapolated using the linear approximation. It is also interesting to notice how the STF does not show the narrow band-pass frequency response that is characteristic of ultrasound transducers (which is highlighted by the NTF shape around resonance), as instead, it stays rather flat across the

whole bandwidth of interest and further. This effect seems to suggest that the modulator's electrical feedback could modify the intrinsically narrow frequency response of the transducer and widen its bandwidth, which is a very desirable feature in ultrasound systems. Another way to see this effect is by realizing that the modulator works in order to minimize the voltage drop (and therefore the current) across the RLC branch by providing a quantized, discrete-time version of the input signal on the other side of the RLC branch. Pushing this to the extreme, let us assume an infinite sampling frequency and resolution: in this case, there will never be any voltage drop across the RLC branch, so the narrow-band filtering of the input signal that usually takes place will not happen any longer. This phenomenon considers the bandwidth of the system as purely defined by electrical quantities. However, the real resonance occurs in the mechanical domain, which is only partly coupled with the electrical one. Therefore, the described effect could not be representative of the reality, because of the transducer model limitations concerning the conversion between the two domains.

2.6.2. Step Response

Figure 2.17 shows a comparison between the step response of the transducer and that of the complete system. Note that the vertical scale is different for the two charts. The modulator's step response is proposed in two representations: as the pure digital output, reconstructed with a virtual DAC and its band-pass filtered version, using a FIR filter with 224 coefficients and the same bandwidth as the modulator ($3.125MHz - 6.875MHz$). Here, the transducer response seems to take much longer to settle with respect to the full system, this can be explained observing again Figure 2.16: the STF, which is the transfer function that the step sees, is much broader than that of the pure transducer, thus the effective Q factor becomes lower, the bandwidth higher and the oscillations will experience a stronger damping factor. This effect can help telling apart echoes with similar arrival times, additionally, it reduces the attenuation experienced by the harmonics of the main transmit frequency, enabling a better SNR in case harmonic imaging is an application of interest.

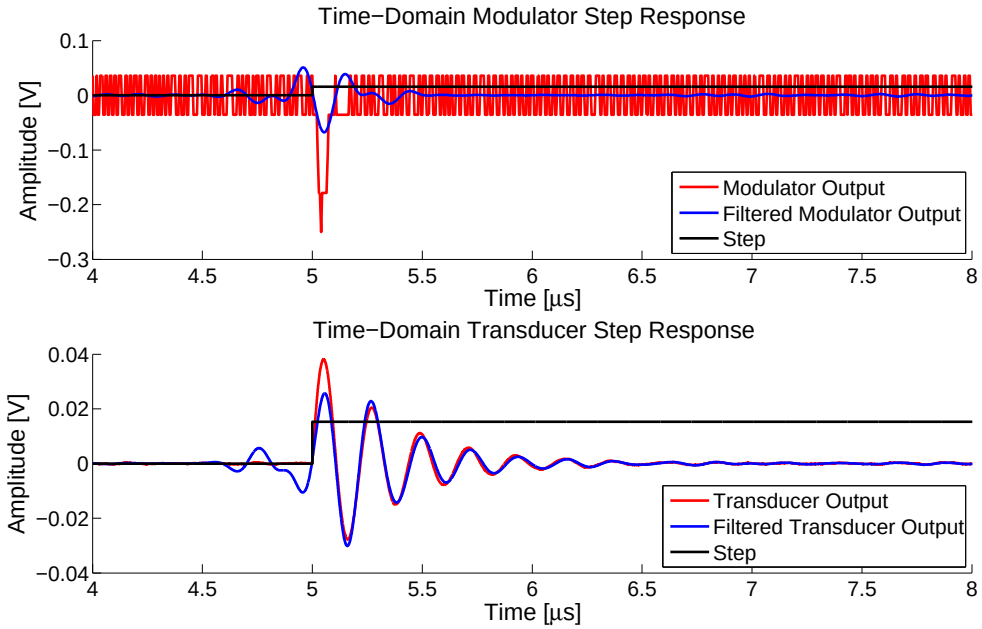


Figure 2.17: System-level Simulink simulation of the modulator's step response, compared to that of the transducer.

2.6.3. Jitter and ELD sensitivity

The implemented modulator shows a very low sensitivity to jitter, that is, the cycle-to-cycle variation associated with the clock arrival time. As a matter of fact, the system can tolerate a nominal jitter almost as high as $\sigma_j = \frac{1}{10 f_s} = 500ps$ without a significant performance drop, as shown in Figure 2.18. The reason for this effect is that the target SNR is low, thus the noise added by the jitter is almost insignificant with respect to the modulator quantization noise and thermal noise. We can use an equation from [33] to estimate the in-band noise added by the jitter in a ΣΔ modulator featuring a Non-Return-to-Zero, multi-bit DAC in the feedback path, as in this case:

$$IBN_{\sigma_j} \approx \frac{V_{range}^2}{(2^B - 1)^2} \left(\frac{\sigma_j}{T_s} \right)^2 \frac{2}{OSR} \tag{2.10}$$

Where V_{range} is the full-scale range of the DAC and B is the number of bits in the quantizer. Assuming a full scale, sinusoidal input and plugging in the parameters of this specific modulator, the allowed clock jitter can be estimated, which is roughly in line with the parameter obtain by simulation:

$$\sigma_j = \sqrt{OSR} \frac{2^B - 1}{2 f_s 10^{\frac{SNR}{20}}} = 286ps \tag{2.11}$$

Again, Equation 2.10 shows how a multi-bit quantizer can help reducing the modulator's sensitivity to jitter, and in general to timing variations. The stability of the system also does not

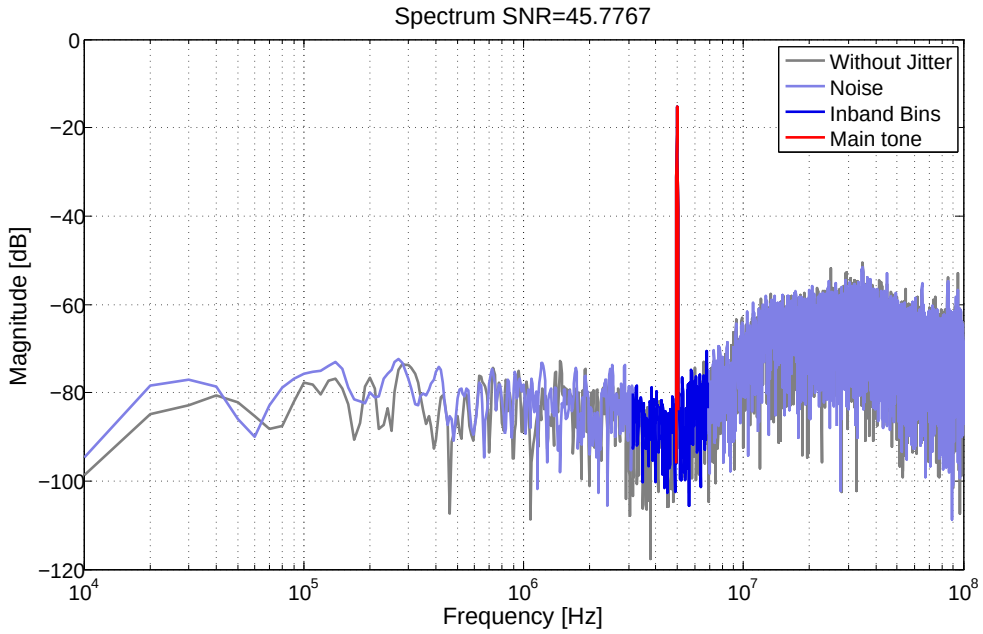


Figure 2.18: System-level Simulink simulation showing the SNR degradation caused by a jitter of $\sigma = 411ps$.

seem to be affected by Excess Loop Delay, up to a simulated propagation delay of $1.8T_s$. Such an extreme tolerance can be explained as such: the block diagrams in Figure 2.6 clearly show a local feedback branch that goes from the quantizer output to its input, with a sign inversion (without considering the "transparent" second stage, which does not affect the noise shaping significantly). This technique is widely used to compensate for ELD [41], and in this design it comes as a by-product of the selected feedback configuration. The theory behind this ELD compensation method states that, in order to counter an instability caused by a high ELD, a

high-speed feedback has to be added around the quantizer, such that it bypasses the loop-filter at high frequencies. This way, even when the loop filter adds to the signal before the quantizer a high phase-shift, the quantizer will still see some sort of feedback, thus keeping its stability. In this work, although the limited bandwidths of the two stages reduce the effectiveness of this high-speed path, the phase-shift caused by the transducer, which is the actual loop-filter, is still bypassed.

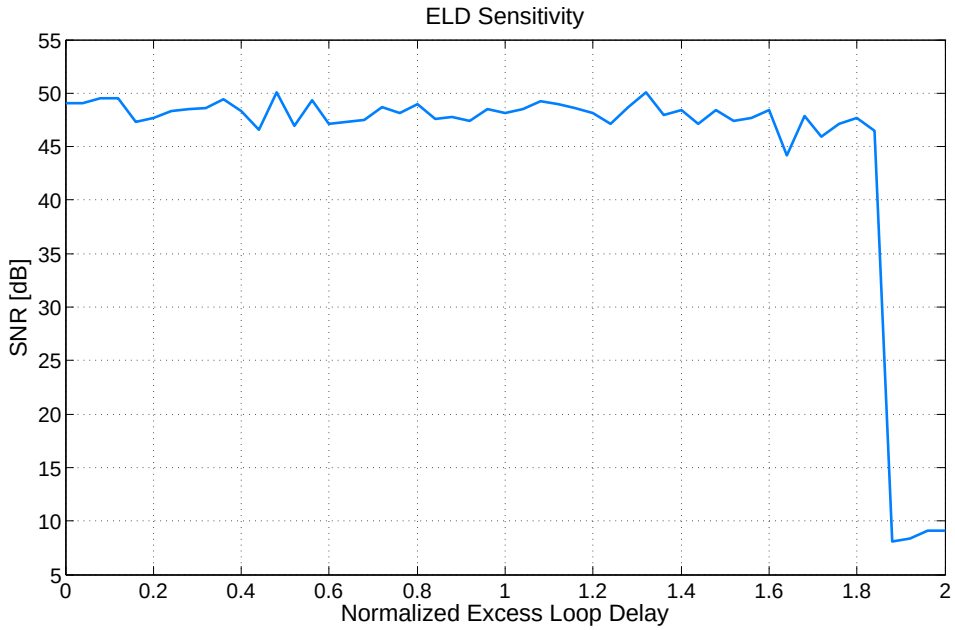


Figure 2.19: System-level Simulink simulation showing the SNR behaviour while varying the Excess Loop Delay of the modulator.

2.6.4. Mismatch Effects

There are many parts of the proposed system that are prone to component variations, and that can therefore possibly give rise to distortion or even instability. The most sensitive blocks in terms of mismatch are:

1. The compensation capacitor C_c and resistor R_c, which have to match C_s and R_s as much as possible, and thus are made tunable.
2. The capacitors C_{ref} which provide the modulator feedback and the capacitance compensation function.
3. The DACs included in the tracking quantizer (see Figure 2.13).
4. The input differential pairs of the first and second stages, as well as those of the comparators.

The other parameters in the loop filter, such as R_F, C_F and the gain of the second stage are not critical, and can accommodate for quite a large variation without compromising the modulator performance. The matching requirement for 1. has been described in detail in Section 2.4.4. The effect of a mismatch in C_c is shown in Figure 2.20: from this chart one can conclude that the best SNR performance is reached when C_c is roughly within the interval $C_s - 200fF < C_c < Cs + 50fF$, where the nominal parasitic capacitance is assumed to be

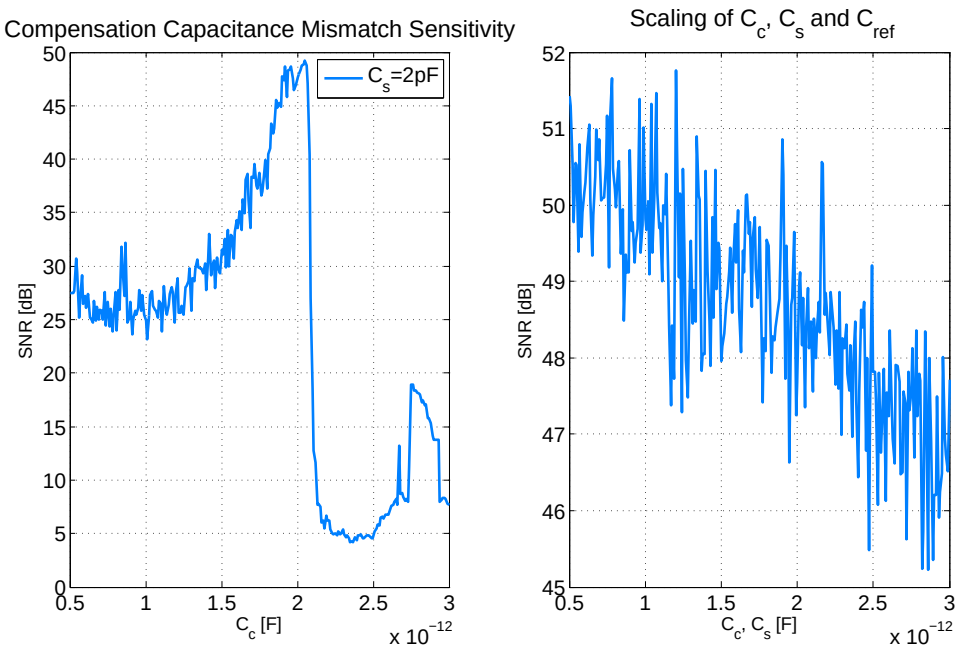


Figure 2.20: System-level Simulink simulation showing the SNR behaviour while varying the compensation capacitance C_c.

$C_s = 2pF$. On the left of this interval, the SNR performance gradually drops because of the imperfect compensation. In this case, the compensation current is not sufficient to completely cancel that which arises from C_s, thus there will be an undesired voltage drop across R_f in addition to the intended one, which will reduce the SNR.
On the right of the allowed region of operation the situation is even worse, as the modulator suddenly becomes unstable due to the now positive feedback path provided by C_c (see Section 2.4.4). This extreme sensitivity to variations in C_c can be used in the measurement phase to estimate the parasitic capacitance C_s within the transducer.

The right part of Figure 2.20, shows how the SNR is affected by a change in C_s; in this simulation, C_c was varied together with C_s and the values for the feedback capacitors C_{ref} were scaled in order to keep the capacitor ratio unchanged. With these assumptions, the SNR variation is only due to a change in the thermal noise transfer function: a higher C_s (and consequently C_c), will yield a higher thermal noise gain and thus reduce the modulator's SNR (see Section 2.3.1).

In the left part of Figure 2.21, the effect of the mismatch among the feedback capacitors C_{ref}, which implement the feedback DACs, is analysed. These components are likely the ones that define the distortion performance of the ADC, as the effect of their mismatch is not attenuated by the noise shaping. To first order, however, the distortion will not affect the in-band SNDR, as the distortion peaks will fall outside of the bandwidth of interest, for a transmit frequency of $5MHz$.
Nonetheless, if a large mismatch is present among the capacitors C_{ref}, the large-signal transfer function of the loop filter can be jeopardized, or even lead to instability, and the effect of it will be visible even in the bandwidth of interest. This seems to be happening when the mismatch is higher than 1%. Luckily, the expected matching for most of the capacitors in CMOS technology is better than 1%, thus the SNR performance of the modulator should not be limited by distortion.

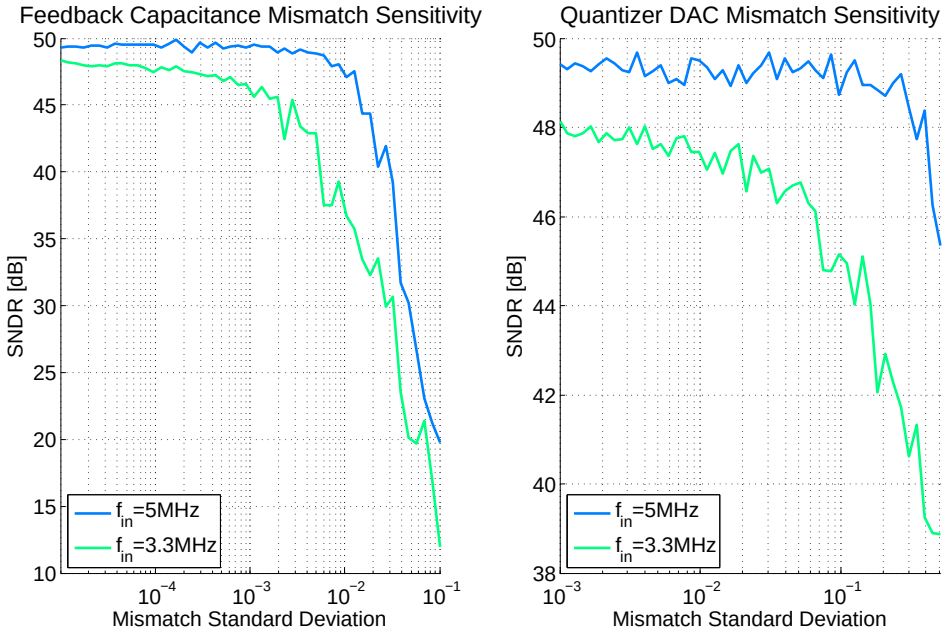


Figure 2.21: System-level Simulink simulation showing the SNDR variation when a mismatch is introduced in the feedback DACs (left) and in the quantizer's DACs (right)

The DACs embedded into the tracking quantizer show an impressive tolerance to mismatch, as shown in the right part of Figure 2.21 (notice that the x axis is scaled differently in the two subplots): even a mismatch with a standard deviation of 10% does not seem to reduce the achievable SNDR. The reason is that the distortion introduced by these DACs enters the modulator at the same point as the quantization noise added by the quantizer, thus the NTF will suppress its in-band components.
If hypothetically one would apply an input signal with a frequency such that its second-order distortion is within the bandwidth (for example 3.3MHz), the SNDR will be reduced by the in-band distortion peak(s) caused by either the feedback DACs or those within the tracking quantizer, as shown in Figure 2.21.

Finally, the effect of a mismatch in the input pairs of the first and second stage will only result in a very small DC offset: the two stages are AC coupled, so that the offset of the first stage will not influence the modulator output, while the second stage is periodically auto-zeroed. The influence of the comparators' offset will be discussed in detail in Chapter 3.

2.6.5. Time-Gain Compensation Implementation

This modulator should be able to face a large variation in the input voltage amplitude, as discussed in Section 1.4. In order to accommodate for that, the system has to be adapted dynamically as the input spans from its maximum to its minimum. First of all, the reference voltages V_{REF+} and V_{REF-}, acting as the supplies for the inverters in Figure 2.15 need to be scaled with the input signal. This means that, if the input amplitude follows an exponential, because of the nature of the attenuation of the acoustic input wave, the reference voltages will have to follow.
Additionally, the gain of the loop filter within the bandwidth of interest also needs to be scaled, as it is desirable to have the same amplitude at the input of the quantizer, regardless of the input signal. Thus, at the beginning of the receive phase, a high gain won't be needed, while

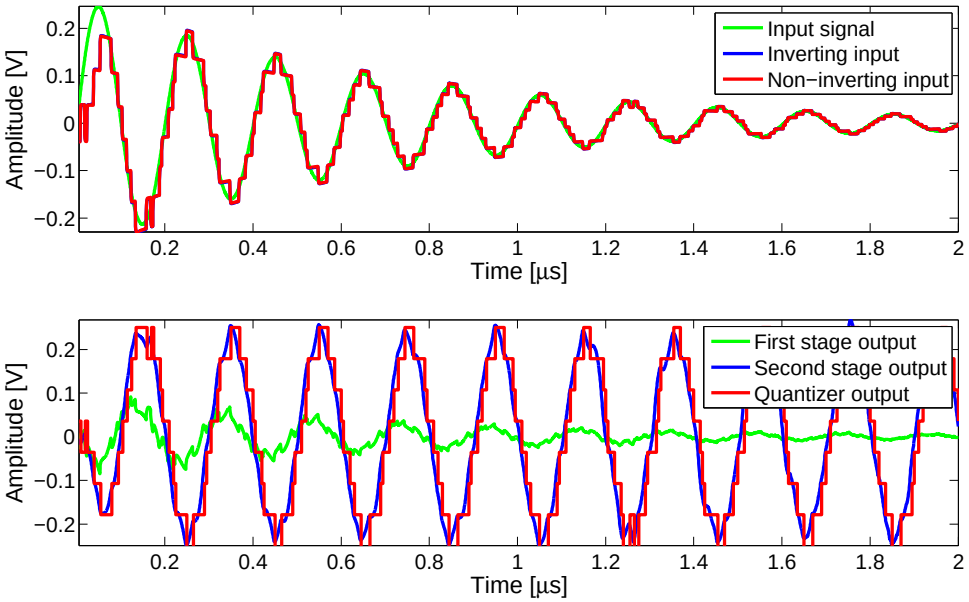


Figure 2.22: System-level Simulink simulation showing the time-domain waveforms of the modulator's critical nodes for continuous TGC.

at the end, the input signal will have to be amplified in order to take advantage of the complete full-scale range of the quantizer. This functionality has been implemented with some capacitors in the second stage feedback (visible in Figure 2.15) that can be disconnected during the receive phase, thus increasing the second stage closed-loop gain.
Figure 2.22 shows the time-domain waveforms in case the input signal, the reference voltages and the second stage gain are scaled according to the TGC. In particular, the voltages at the inputs of the first stage track the decreasing input signal, the output of the first stage gets smaller as time progresses, however, the second stage output and that of the quantizer do not seem to be influenced by the variation in the input amplitude.
Another possibility would be dynamically scaling directly the tracking ADC's full-scale, however, in this case, the comparator noise becomes increasingly significant as the signal gets smaller, which, of course, is undesirable.

2.7. Conclusion

In this Chapter, all the steps that led to the final architecture were followed. The choice converged on a continuous-time, band-pass, 3-bit $\Sigma\Delta$ modulator, clocked at $200MHz$ and featuring a tracking quantizer. The issues associated with the transducer's capacitance C_s were explained, together with a few possible solutions. Finally, an evaluation of the system's stability and robustness was performed. In Chapter 3, implementation details concerning all the blocks that were discussed in this Chapter, and more, will be provided.

3

Element-Level Design

The previous Chapter converged on a block diagram, shown in Figure 2.15, which seems to satisfy all the system-level specifications. In this Chapter, more details about the topology choices and the schematics of the implemented blocks will be provided. We will try to focus on the most critical decisions that were made, rather than extensively conveying all the simulation results.

3.1. First Stage

3.1.1. Topology Choice

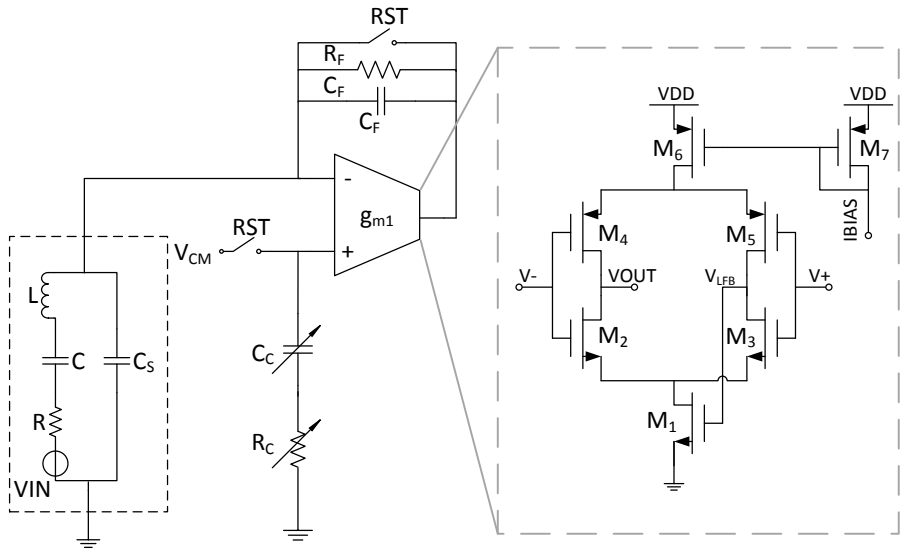


Figure 3.1: Schematic of the first stage OTA.

The first stage is the most power consuming block in the whole system, as usual in the field of analog design. The reason is that the noise introduced by the first stage ultimately defines the total SNR of the system. Thus, a careful topology choice is due, in order to obtain the desired noise performance without spending too much of the power budget.

The system is intrinsically single-ended, as only one node that interfaces with the transducer is available, however, the system-level design requires an input-differential first stage,

therefore a differential input, single-ended output topology has to be selected. Undoubtedly, one of the most efficient OTA implementations is the current-reuse amplifier. Its main advantage is the fact that it fully exploits the CMOS paradigm, as it is entirely complementary, using both an NMOS and a PMOS, sharing the same bias current, to provide a voltage to current conversion. Therefore, current-reuse amplifiers have the power to intrinsically double the achievable g_m without spending any additional current in doing so.

These kinds of OTAs have been implemented in many different ways, one that adapts well to the first stage requirements is proposed in [42]. Here, a differential to single-ended, current-reuse, self biased amplifier is proposed and referred to as CSDA. The perk of this implementation is that it does not require a biasing circuit, as both the PMOS and NMOS current sources are connected to the unused output of the amplifier, thus effectively creating a common mode feedback structure.
Additionally, this topology only features one high impedance node, thus it intrinsically has a dominant pole. However, this topology has two major problems: it has a poor power supply rejection ratio (PSRR) and a low output swing. A way to partially solve these issues is providing local feedback on only one of the current sources (V_{LFB}), in this implementation the bottom one (as the ground is thought to be more stable than V_{DD}), as shown in Figure 3.1. Here, the use of a current source biased by a local biasing block helps reduce the sensitivity of the output to transients on the power supply.

3.1.2. Theoretical Behaviour

The differential, small-signal, open-loop DC gain is easily computed and is equal to:

$$A_{OL}^{DC} = (g_{m,N} + g_{m,P})(R_{out,N} \parallel R_{out,P}) \tag{3.1}$$

The total input referred power spectral density, neglecting the contribution of the current sources, is found to be:

$$v_{n,eq}^{in} = \sqrt{\frac{8k_B T \cdot BW}{3g_m}} \tag{3.2}$$

A careful analysis of the high-frequency small-signal equivalent circuit was performed, which accounts for the load capacitance C_{load} at the drain of the input transistors, C_{gd} (the parasitic feedback caps between the drains and the gates of the input transistors) and $R_{out,N,P}$ (respectively the NMOS' and the PMOS' output resistance). M_6 was assumed ideal, while M_1 was modeled by a transconductance $g_{m,c}$. With these assumptions, the OTA has two zeroes and two poles:

$$p_{1,2} = \frac{1}{(R_{out,N} \parallel R_{out,P})(C_{gd} + C_{load})}; \quad \frac{g_{m,M_1}}{2(C_{gd} + C_{load})} \tag{3.3}$$

$$z_{1,2} = \frac{g_{m,M_1}}{C_{gd} + C_{load}}; \quad \frac{2g_m}{C_{gd}} \tag{3.4}$$

The dominant pole is the first, which determines the OTA's open-loop bandwidth. Interestingly enough, the first zero is at exactly double the frequency of the second pole. Finally, the last zero is positive, but it is normally located at very high frequencies.

3.1.3. Medium V_T NMOS Input Pair

The required input common-mode swing has to be slightly higher than the maximum input peak-to-peak amplitude expected from the transducer, which is derived in Equation 2.6. Thus,

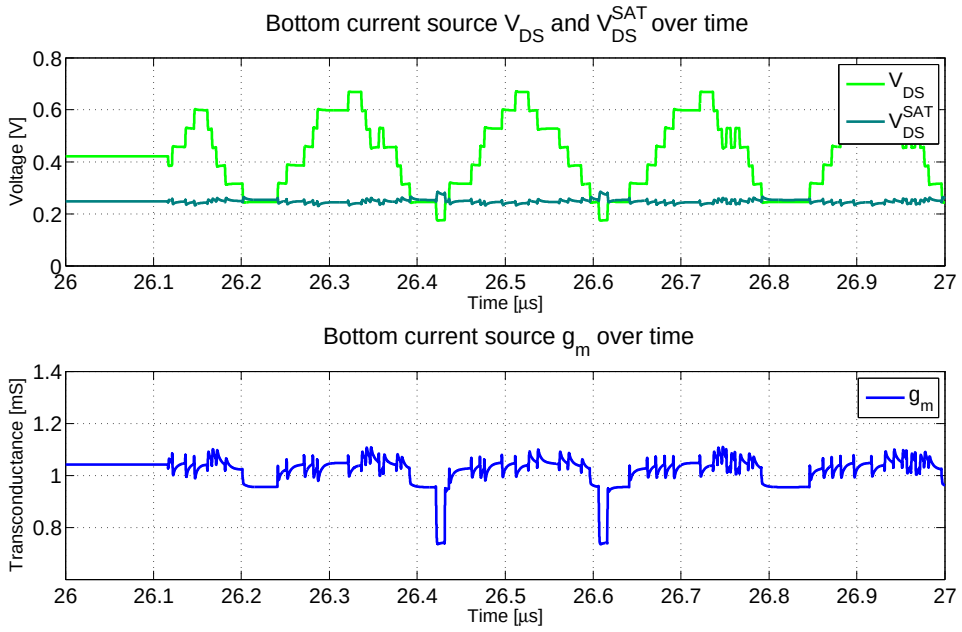


Figure 3.2: Transient simulation showing the operating point of M_1 versus time.

the first stage OTA has to be able to handle an input common-mode swing of around $500mV$. This specification is in conflict with the desired region of operation for the bottom current source M_1: theoretically, V_{LFB} should stay around the common-mode voltage of $900mV$, however, this brings the saturation voltage to $V_{DS}^{SAT} \approx V_{GS} - V_T = 900mV - 490mV = 410mV$, which is very high. Furthermore, the drain voltage has to be one V_T lower than the input one, which can reach voltages as low as $750mV$. A standard NMOS has a $V_T \approx 550mV$, depending on the body effect, thus placing the the drain of the current source at $200mV$, far away from the saturation region.
In order to solve this problem, V_{LFB} was biased at $750mV$ rather than $900mV$, thus decreasing the current source's V_{DS}^{SAT} by the same amount. Additionally, medium V_T devices were selected for the input NMOS transistors, reducing the V_T by more than $100mV$. These transistors also show a better current efficiency, but they lack in output impedance, which is only marginally important in this work. The PMOS current source also faced a similar issue, which was solved simply by connecting the bulk of the PMOS input transistors to their source, thus annulling the V_T increase due to the body effect.

The result is shown in Figure 3.2: When V_{DS} drops below V_{DS}^{SAT}, the NMOS goes out of saturation and the g_m lowers, together with the output resistance. This simulation was set up using an input voltage higher than the maximum for which the system was designed. Thus this situation should not occur in normal operation.

3.1.4. Simulation Results

The top part of Figure 3.3 shows the result of a post-layout Monte Carlo simulation of the first stage open-loop gain, accounting for process variations. The g_m of the input transistors amounts to $1.53mS$, while the total output resistance is $21k\Omega$, thus the DC gain is $A_{OL}^{DC} = 2g_mR_{OUT} = 65 = 36.3dB$. The intrinsic output capacitance is $C_{load} = 85fF$, while the feedback capacitance is $C_{gd} = 125fF$, therefore the first pole is at $p_1 = \frac{1}{2\pi R_{OUT}(C_{gd}+C_{load})} = 35.6MHz$. The simulation results follow closely the theoretical model derived above (red

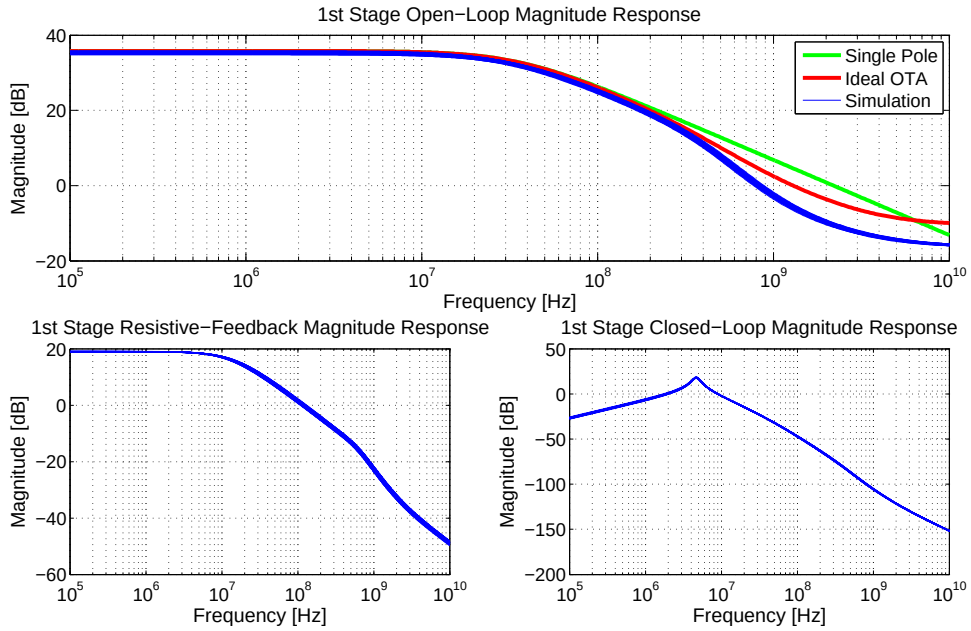


Figure 3.3: Simulations of the first stage in open loop, closed loop and resistive feedback mode.

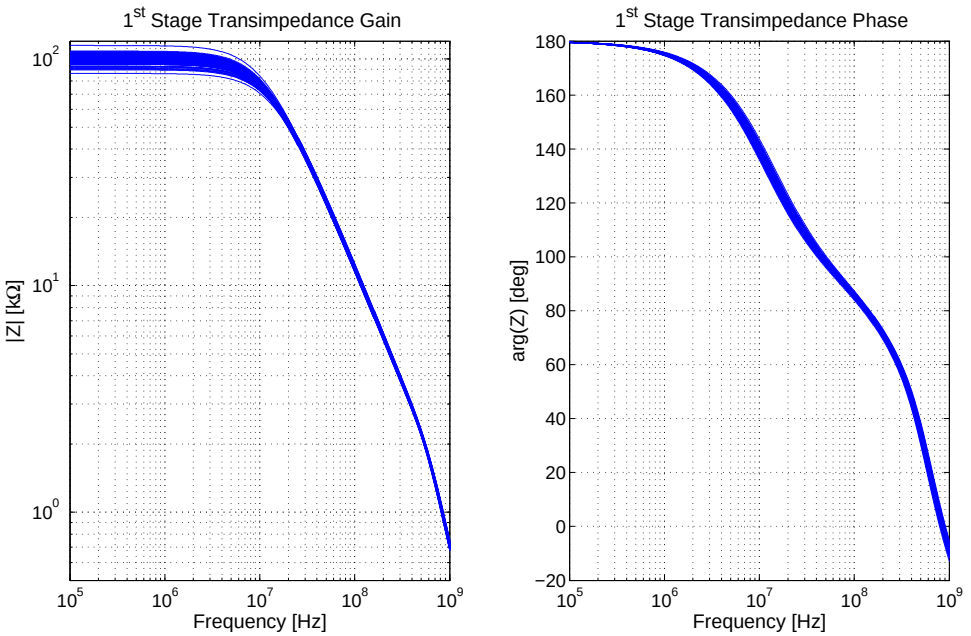


Figure 3.4: Transimpedance gain of the first stage, while injecting a current at its input.

curve), up to roughly $200MHz$. The simulated unity-gain frequency is $800MHz$.

Let us close the loop with a *resistive* feedback, adding R_F between the output and the inverting input, and a resistance equivalent to that in the transducer model between the inverting input and the AC signal source. the closed-loop magnitude response is shown in the bottom-left of Figure 3.3. The DC gain is slightly lower than the ideal: $\frac{R_F}{R} = 10.5 = 20.4dB$, because of the finite DC gain. The first pole is defined by R_F and C_F, the latter of which is composed

of an explicit capacitor of $35fF$ and a parasitic C_{gd} of $125fF$. The pole can be computed as: $\frac{1}{2\pi R_F C_F} = 10MHz$, which agrees with simulation. On the bottom right of Figure 3.3, the previous analysis was repeated, adding the reactive components in the transducer's RLC branch: the two charts reach the same value ($\approx 20dB$) at resonance ($4.6MHz$). The chart on the right has an extra zero and two additional poles that are given by the transducer resonance.

Another way to see the AC behavior of the first stage is considering it as a transimpedance amplifier, instead of a resonator. A current injected in the amplifier's virtual ground sees the transimpedance shown in Figure 3.4. The impedance is equal to the feedback resistance up to the first pole (around $10MHz$), where the impedance of C_F equals that of R_F. After that, the impedance drops at a constant rate up to very high frequencies.

The whole first stage draws $165\mu A$ from the analog supply. The input MOSFETS operate in weak inversion and achieve a transconductance of $1.53\frac{mA}{V}$. The resulting current efficiency can be calculated as $CE = \frac{g_m}{I_D} = \frac{1.5mS}{165\mu A/2} = 18$. Since the NMOS and the PMOS at the inputs share their bias current and both contribute to the effective transconductance, the current efficiency is effectively doubled, amounting to 36. This is beyond the theoretical $\frac{g_m}{I_D}$ achievable in this technology node (≈ 25) with a topology that does not employ the current-reuse technique.

Finally, the simulated post-layout input-referred integrated in-band noise is $6.97\mu V_{rms}$. The expected value from Equation 3.2 is $5.2\mu V_{rms}$. The discrepancy can be explained by the contribution of the current sources and of the flicker noise, which, although not dominant, still has a small influence on the total input-referred noise.

3.2. Second Stage

The purpose of a second stage, in this design, is merely to provide the gain needed to adapt the output amplitude of the first stage to the full-scale range of the quantizer. Thus, it is configured as a simple amplifier, which has a flat gain in the bandwidth of interest and does not provide any form of noise shaping. The gain that this amplifier should provide is not fixed, but it varies during the receive phase according to the time-gain compensation scheme.

This gain stage is not constrained by many stringent system-level requirements: its noise contribution will be attenuated by the gain of the first stage. Therefore, it does not require a high bias current, its output swing should match the quantizer's full-scale range ($500mV_{pp}$) and its maximum gain should be $33dB$.
The bandwidth of the second stage was chosen so that it filters out the high-frequency noise coming from the first stage, without adding excessive phase shift to the frequencies of interest. However, this filtering action is only important at the highest gain configuration, as at lower gains the input signal amplitude is supposed to largely overcome the thermal noise of both the transducer and the first stage, so that the most limiting factor for the modulator's SNR is its quantization noise, which scales with the input amplitude.

3.2.1. Topology Choice

Even though the second stage gain does not have to be extremely precise, a feedback configuration is still preferred over an open-loop one, because the gain switching can be easily and precisely implemented by changing the elements in the feedback loop, rather than the transconductance (or output impedance) of a g_m stage.

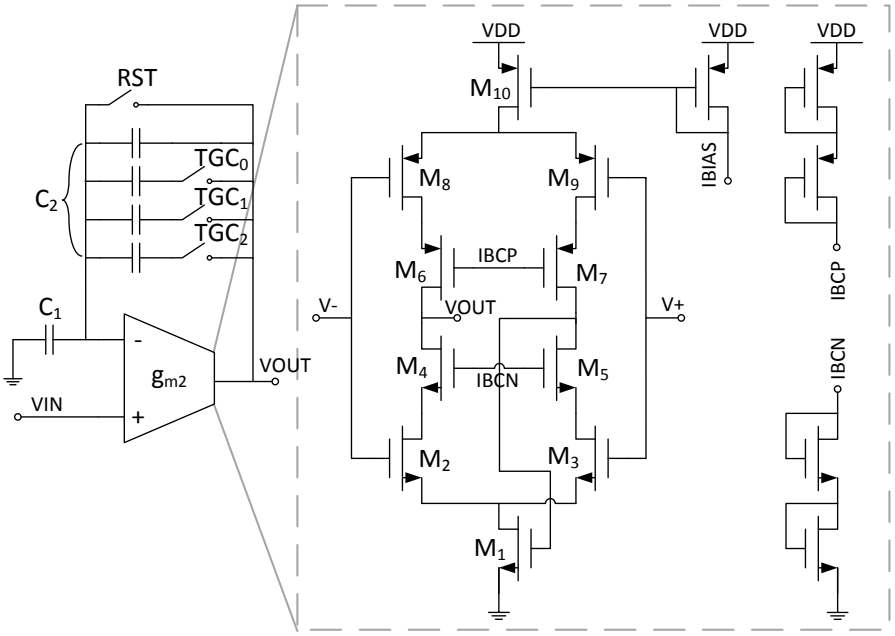


Figure 3.5: Schematic of the second stage OTA.

Capacitive feedback was preferred over a resistive one because the latter would allow any offset in the OTA to be gained up and transferred to the output, which will not consume part of the quantizer full-scale range. With capacitive feedback, since the receive phase is periodic, an auto-zero mechanism is granted “for free”, drastically reducing the effect of any offset in the loop filter. Also, resistive feedback would load the output of the OTA and reduce its output impedance, making the use of cascodes futile and yielding a lower gain precision.

The chosen feedback unit capacitance is $11fF$, slightly higher than the minimum capacitance of $9fF$ achievable with standard MOM capacitors, using metal 1 to 4 for maximum density. Implementing a gain of $33dB$, equivalent to a linear gain of 44, means having a C_1 equal to $44 \times 11fF = 484fF$. In case a single-ended, inverting configuration was chosen for the amplifier, this large capacitance would be loading the first stage and limit its bandwidth, which can lead to an unstable modulator. In order to avoid this loading effect, a non-inverting configuration was selected instead, where the first stage only drives the gate capacitance of the amplifier’s non-inverting input, rather than C_1.
The downside of a non-inverting configuration is that it can only have a differential input, thus employing twice the current (as it will have two current branches), for the same performance (bandwidth and noise). This shortcoming is tolerable, as the second stage will consume significantly less power than the first one, and using a buffer between the two stages to drive the input capacitance of the second one would burn roughly the same current as an additional branch in this amplifier.

The chosen OTA topology, shown in Figure 3.5, is very similar to that of the first stage, but the output impedance was augmented through the use of cascodes, so that the real gain of the stage is closer to the ideal one set by the components in the feedback path.

3.2.2. Gain Derivation

In this Section, the implementation of the time-gain compensation scheme will be discussed (see the dynamic range specification in Section 1.4).
The starting point is chosen to be the final, and highest gain needed to meet the specifications:

Table 3.1: Time and gain settings of the implemented TGC steps.

	Capacitor Ratio	Gain [dB]	Time [μs]	Digital Code
Initial Step	40/27	8.5	0	11
Second Step	40/9	15.5	58	10
Third Step	40/3	24	103	01
Final Step	40/1	33.2	151	00

system-level simulations suggest that the second stage should have a gain of approximately $33dB$ in order to reach the optimal SNR. The initial gain is derived by subtracting from the final one the desired TGC range (in dB). The latter is initially designed to be $27dB$, as it is a multiple of 3, which is also the selected number of TGC steps after the first one.
We are left with four gain steps: $6dB$, $15dB$, $24dB$ and $33dB$. The step size is $9dB$, which is roughly equal to a factor 3 in linear scale, enabling an easy implementation of the gain steps using capacitor ratios. However, this stage's gain is ideally $G = 1 + \frac{C_1}{C_2}$, because a non-inverting configuration was employed. In order to account for the factor $1+$, instead of changing the capacitance ratios, it was decided to adjust the time instants when the gains are switched, as illustrated in Table 3.1 and Figure 3.6.

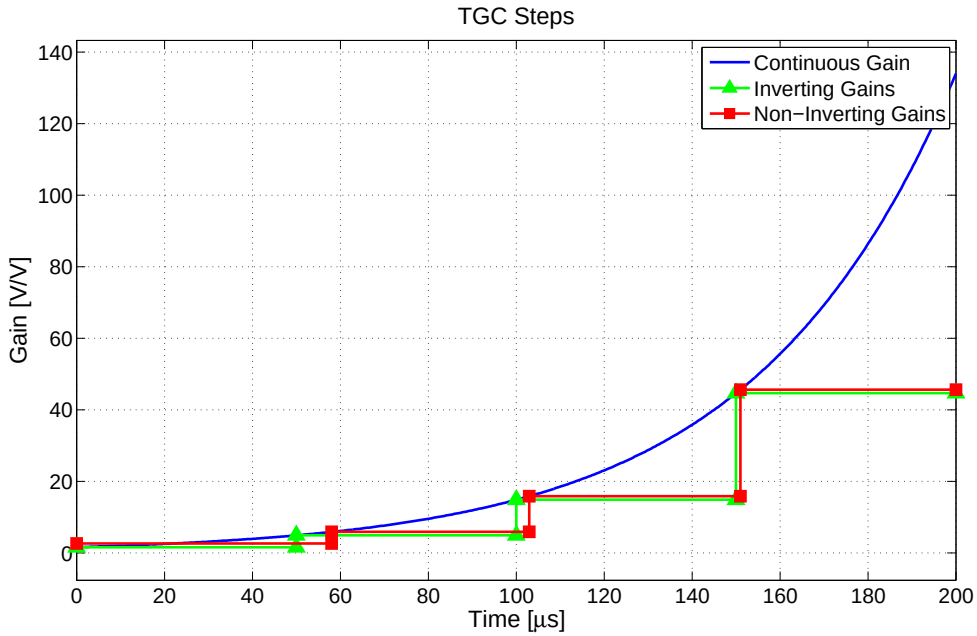


Figure 3.6: Gain settings of the TGC scheme and correspondent switching times.

3.2.3. Simulation Results

Figure 3.7 shows a post-layout Monte Carlo simulation of the second stage in the different gain settings. As expected the spread is lower when the gain is small. With the highest gain, the bandwidth is lower, this can be explained calculating the pole of the simplified system, representing all the transistors as a single g_m and assuming an infinite output impedance.

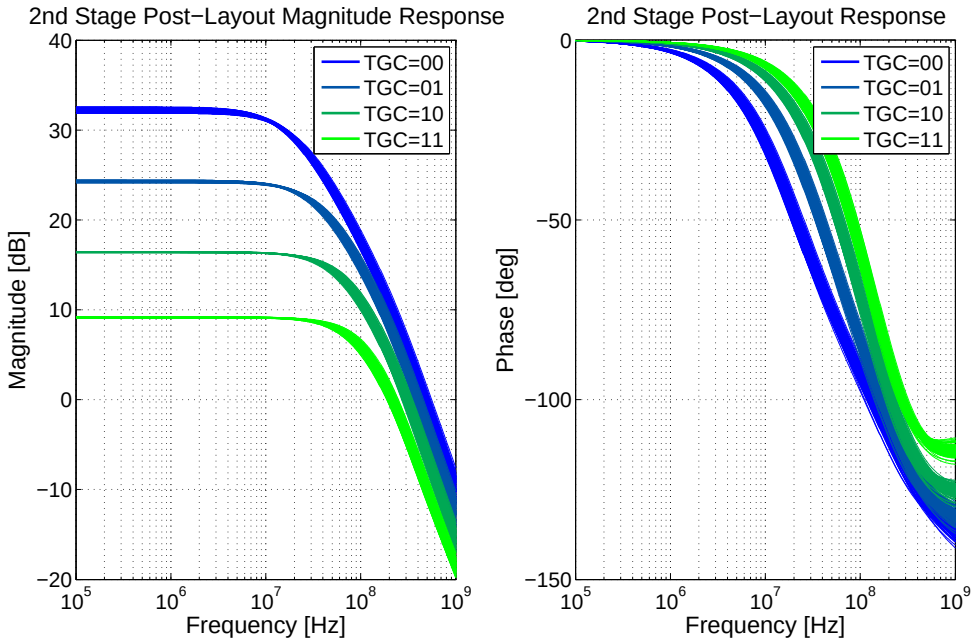


Figure 3.7: 2^{nd} stage frequency response (post-layout).

The $3dB$ angular frequency is computed as follows:

$$\omega_p = \frac{2g_m}{C_1 + C_{OUT}\left(1 + \frac{C_1}{C_2}\right)} = \frac{2g_m}{C_1 + C_{OUT}A_{CL}} \tag{3.5}$$

where g_m is the transconductance of transistors M_2, M_3, M_8, M_9, C_2 is the feedback capacitor, C_1 is the shunt capacitor connected to the OTA's inverting input and C_{OUT} is the output capacitance of the OTA. From this equation it is possible to notice how the bandwidth is lower in case the closed-loop gain is higher. This effect is exacerbated by the increase of C_{OUT} in the low gain configurations, as C_2's parasitic capacitance to ground goes in parallel with the OTA's output capacitance.

Figure 3.8 depicts the loop gain of the second stage for all the gain settings, including the reset phase, when the loop gain is equal to the open-loop gain of the OTA. In the reset phase, which is the most critical for stability, the $0dB$ axis is reached at $100MHz$, where the phase is only 75^o, thus stability is guaranteed. This stage consumes $36\mu A$ of current from the analog supply, while the transconductance of each input transistor is around $270\frac{\mu A}{V}$. They all operate in weak inversion and achieve a current efficiency of $CE = \frac{g_m}{I_D} = \frac{270\mu S}{36muA/2} = 15$, which doubles, accounting for the current-reuse structure.

3.3. Tracking Quantizer

3.3.1. Comparator

The twin comparators in each tracking quantizer are very important: they consume a significant amount of power, their delay is in the feedback's critical path and their output acts as the output of the complete system. Additionally, their noise and offset should be kept low with respect to the loop filter's output noise and to the quantizer's LSB.

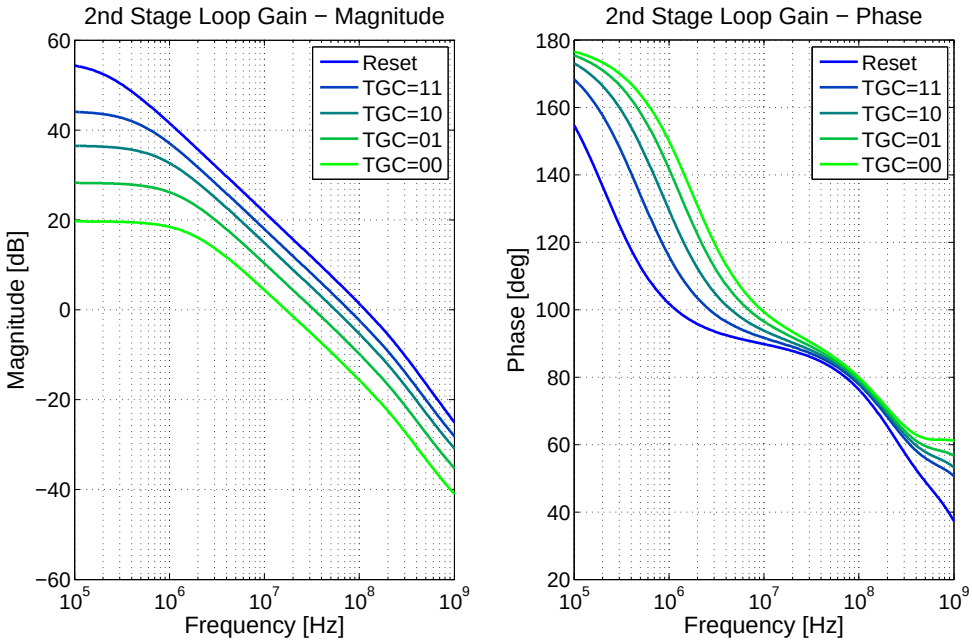


Figure 3.8: 2^{nd} stage loop gain magnitude (post-layout)

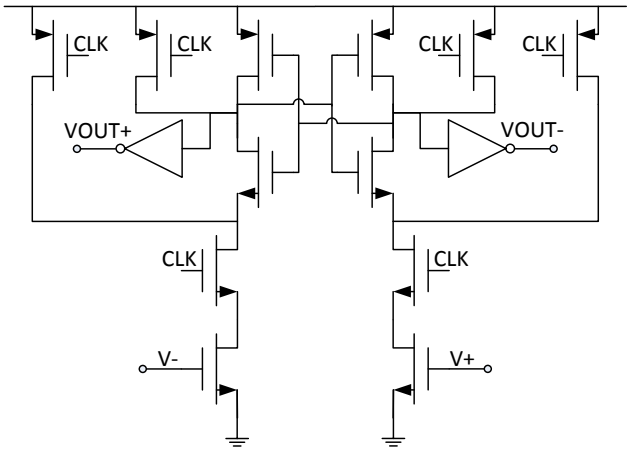


Figure 3.9: Schematic of the proposed comparator.

Topology and Kickback Noise

A dynamic comparator was chosen for this design: this type of comparator consumes no static power, while keeping the speed high. In particular, the topology discussed in [43] has been selected as a starting point for its well-known performance and simplicity. Clocking this comparator is rather easy: there is no need for an (i.e. non-overlapping) inverted clock, this helps to save some power in the local clock buffers, which, as will be explained in Section 3.3.2

Early simulations of the entire system with such a comparator showed a kick-back noise with an amplitude of several tens of millivolts. This effect can be dangerous, especially when one of the inputs of the comparator is a floating node, or is driven by a capacitive-feedback amplifier: both of these scenarios are true in this design.

In order to reduce kick-back noise, similarly to the solution proposed in [44], the NMOS switch

underneath the input pair is split in two, and moved on top of them, as shown in Figure 3.9. This technique lowers the kick-back noise because the switches disconnect the comparator inputs from the latch, when this is reset.

Noise and Offset Estimation

Because of the dynamic nature of the comparator, traditional simulation tools prove to be ineffective in the estimation of its input referred noise and offset. A common technique used for this purpose employs statistical analysis. For a noise assessment, the comparator is given a static voltage difference between its inputs and is run N times in the presence of transient noise. The voltage difference is swept within the range of interest ($[-10mV - 10mV]$) and the probability of the comparator making the expected decision is calculated. We can then plot the probabilities versus the differential input voltage, as in the top part of Figure 3.11, and fit the resulting curve with an error function. From the parameters of the obtained error function, the standard deviation of the noise σ_n can be derived, shown on the bottom part of Figure 3.11.

A similar procedure can be used to estimate the comparator's offset: instead of adding noise

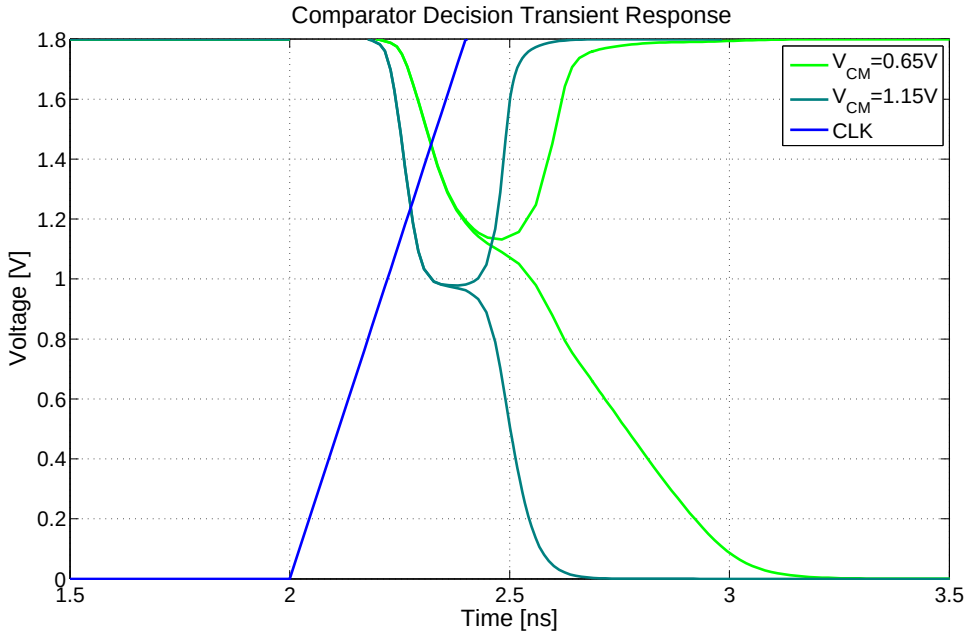


Figure 3.10: Comparator decision dynamics vs. common-mode voltage

to the simulation, one can model the transistor mismatch and run a Monte-Carlo analysis N times for every input differential voltage. The results are processed exactly as above and the result can be seen in Figure 3.12. From Figure 3.11 and 3.12 one can conclude that with a higher input common-mode voltage, the comparator generally performs worse in terms of both offset and noise. For the implemented tracking quantizer, the common-mode input voltage for the comparators is not fixed, but is data dependent and it corresponds roughly to the instantaneous level of the second stage output, assuming that the reference voltages generated by the tracking quantizer DACs are always less than one LSB away from the second stage output. Therefore, the common-mode input voltage is included in the interval $V_{CM}^{comp.} = \left[0.9 - \frac{FS_{comp.}}{2};\quad 0.9 + \frac{FS_{comp.}}{2}\right] = [0.65V;\quad 1.15V]$.

With a higher common-mode voltage at the input of the comparator, the g_m of the input differential pair becomes higher. For many analog blocks, this is desirable, however, in this case,

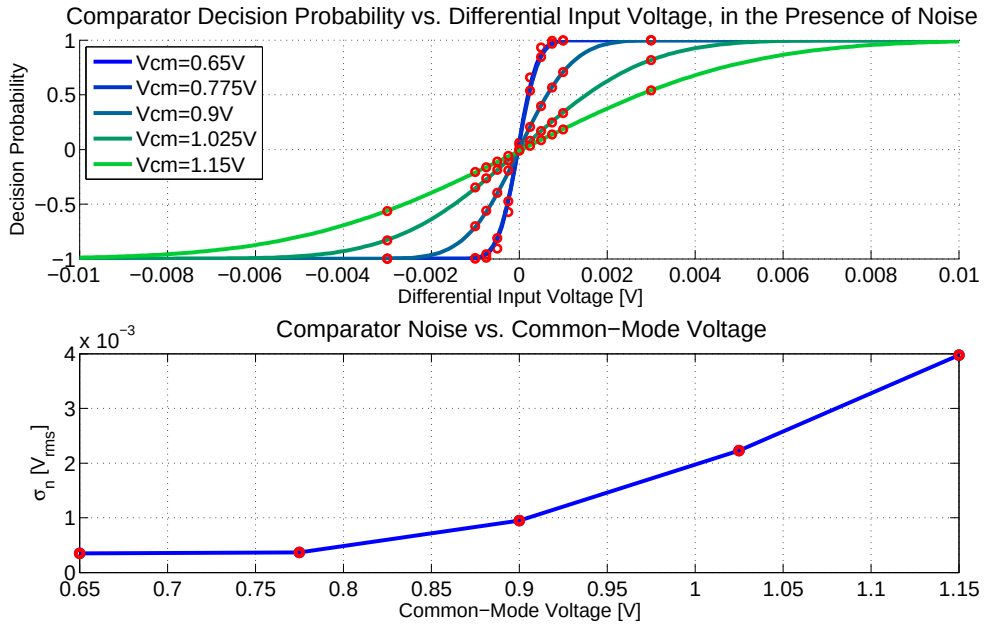


Figure 3.11: Comparator noise simulation

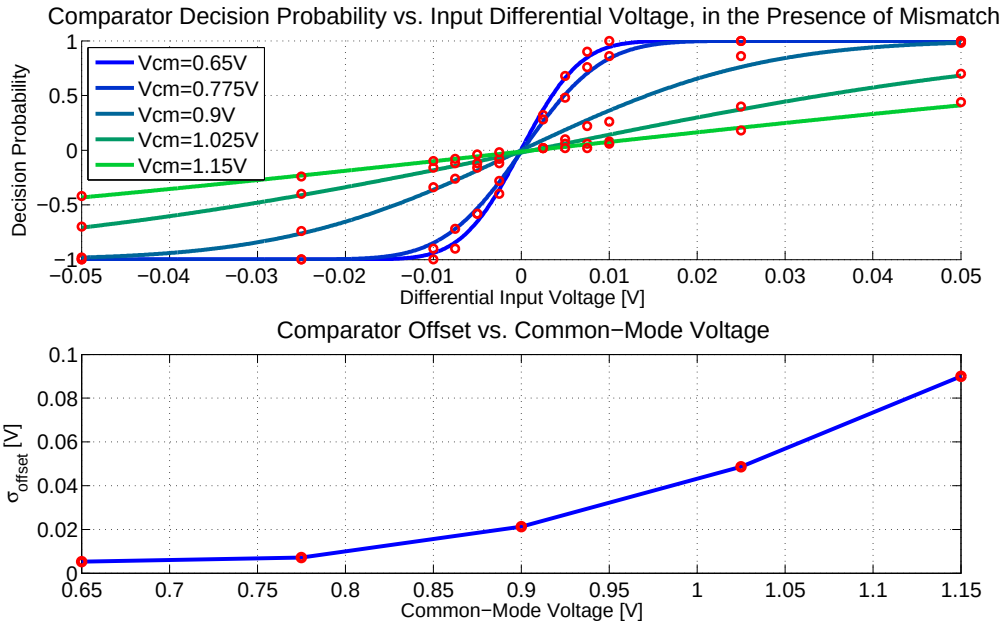


Figure 3.12: Comparator offset simulation

it leads to a higher noise and offset.
This may seem counter-intuitive, but it can be explained with the following large signal analysis. A higher g_m grants additional speed, thus the comparator will make a decision sooner, as shown in Figure 3.10. However, the voltage difference between the two terminals of the latch, at the decision moment, will be the same. This means that the situation with the highest *output* noise will be the most critical. The current noise added by the input pair to the desired output current is directly proportional to g_m: $v_{n,rms}^2 = 4k_BT \cdot BWg_m$, which, in turn is proportional to

the input common-mode. At the same time, the comparator noise bandwidth, which is also defined by its transconductance, is larger for a higher input common-mode voltage.
Intuitively, this line of reasoning can be extended to mismatch, as a higher transconductance on the input transistors will make the effect of any mismatch between them more dramatic.
It should be noted, however, that the result of the previously described procedure is the input referred integrated noise over the whole bandwidth of the comparator, which ranges from 0 to a few Gigahertz and is defined primarily by the g_m of the input pair and the parasitic capacitor at its drains. For this reason, this rms voltage cannot be directly compared with the noise introduced by the analog circuitry, as the latter was evaluated only within the system's bandwidth $[3.125MHz - 6.875MHz]$.

The data dependency of the comparators' common-mode level, and thus their noise and offset, is certainly undesirable, however, as explained in Section 2.6.4, any non-ideality introduced by the comparators does not directly yield a heavy performance loss for the modulator.

3.3.2. Digital Logic

Another block of fundamental importance that should be discussed is the logic required by the tracking quantizer. As with the previously discussed building blocks, this one should also consume a very low power and introduce a low delay in the modulator's feedback path. The complete structure is shown in Figure 3.13.

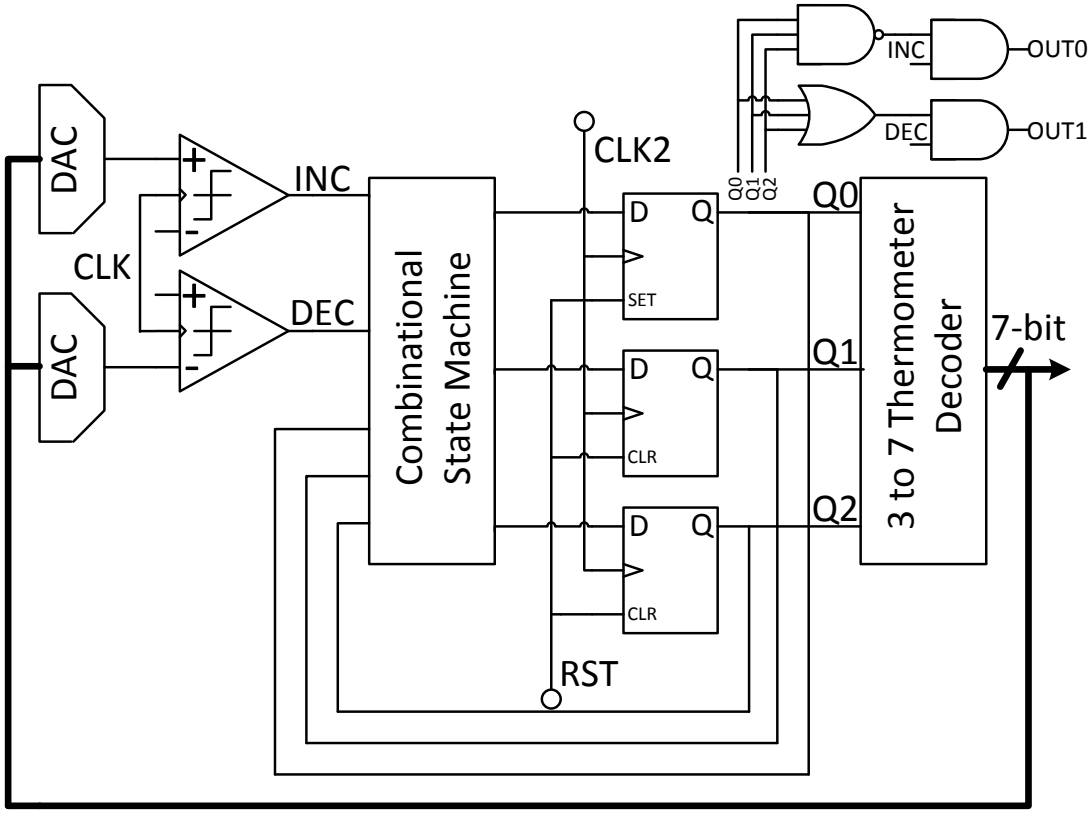


Figure 3.13: Detailed schematic of the logic block included in every element of the array.

Component Optimization

The design of the digital part of the system has as the main objective the minimization of the number of sequential logic circuits, i.e. latches and flip-flops. Such circuits are, however, necessary, since the concept of a tracking quantizer inherently requires a memory storage element to function. Furthermore, re-clocking the data between the comparators and the feedback DAC of a modulator is, in general, good practice to avoid distortion that could arise from a data-dependent delay in the comparators.
For each flip-flop in series, the data is delayed by one clock cycle, assuming the same sampling frequency for the comparators and the flip-flops. However, the feedback signal should be ready within one clock period, so more than one flip-flop in series is not acceptable. The minimum number of parallel flip-flops is three, as this is the number of bits for the quantizer, assuming the flip-flops will store the data in binary form. These three flip-flops act as the starting point for the design of the rest of the logic cells.

A state machine is introduced between the comparators and the flip-flops. It functions as a look-up table, which determines the output code for every clock cycle, based on both the comparator's decision and on the previous cycle's stored value. The state machine was written in Verilog, synthesized and mapped into strictly combinational standard cells. The layout was also generated automatically in Encounter. The synthesized schematic has a maximum of four logic gates in series for any input signal, keeping the maximum delay of the block low.

After the state machine reaches the intended value, the flip-flops are clocked by $CLK2$ (a delayed version of the comparators' clock CLK), so that the output of the state machine is latched and available for the next cycle. Note that the state machine has to be able to work within a precise time-frame: between the two clock edges of CLK and $CLK2$. In particular, the time between the two clock edges has to be between $1.5ns$ and $2.5ns$.

Next, a 3 to 7 decoder was designed and synthesized to adapt the flip-flops' binary output to the thermometer-coded feedback DACs of both the tracking quantizer and the whole modulator. Again, the delay of this block should be kept low, thus, the maximum number of gates in series is two. For this block, it's also important to keep the driving capabilities of the output gates rather constant, so that all the 7 bits of the DACs will change simultaneously. Additionally, an effort should be made to reduce glitches in the decoder as much as possible: a high glitch amplitude might trigger the feedback inverters of the modulator and lead to an unwanted feedback pulse injected at the input of the modulator.

An overload protection circuit was added (on the top right of Figure 3.13), which converts the comparators' outputs into the final outputs of the modulator: these four gates make sure that if the output code had reached its extremes (code=000 or 111) any further pulses on the DEC and INC lines respectively will be suppressed, as the state machine's output will not change as a result of them. This block also provides additional driving capabilities, as it has to drive lines up to $400\mu m$ long.

Finally, every element has a built-in clock driver, which can be enabled or disabled locally by a bit in the shift register (see Section 4.4) for testing purposes. This driver consumes a fair amount of power because its output toggles every clock cycle, even when the tracking quantizer does not change its output.

Table 3.2: Current consumption and maximum delay over corners for every block in the tracking quantizer.

	Current	Maximum Delay
Comparators	$72.5\mu A$	$390ps$
Flip-Flops	$21.7\mu A$	$250ps$ (from clock)
State Machine	$9.7\mu A$	$250ps$
Decoder	$3.3\mu A$	$150ps$
Overload Protection	$4.4\mu A$	$200ps$
Local Clock Buffers	$18.2\mu A$	$240ps$

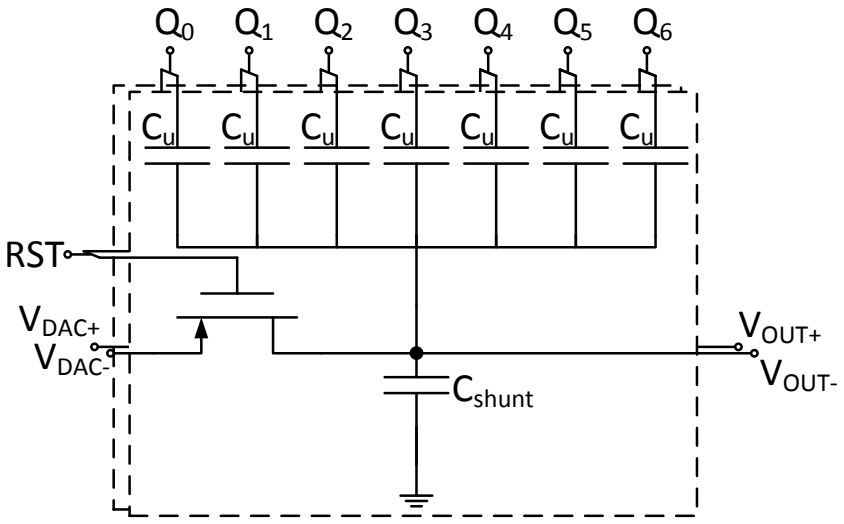


Figure 3.14: Schematic representation of the tracking quantizer's DACs

3.3.3. DAC

The tracking quantizer's DACs were implemented in a very straight-forward manner (see Figure 3.14): the thermometer outputs of the 3-to-7 decoder were directly connected to two sets of capacitive dividers, much like the global DAC's topology. The floating nodes that act as the DACs' outputs are initialized to a known value (V_{DAC+} and V_{DAC-}) during the reset phase, while the step size is defined by the digital core supply voltage and the capacitive divider ratio.
The DACs were designed so that the step size is $V_{LSB} = \frac{FS}{2^{N-1}} = \frac{500mV}{7} = 71.4mV$ when $VDDCORE = 1.8V$. The unit capacitance is the minimum allowed value for MOM capacitors ($9fF$), as the modulator has a very high tolerance to any mismatch in these DACs (see Section 2.6.4). The shunt capacitance is calculated using the following formula:

$$C_{shunt} = C_u\left(\frac{VDDCORE}{V_{LSB}} - 7\right) = 163fF \tag{3.6}$$

The actual value for C_{shunt} also accounts for the parasitic capacitance to ground of all the capacitors. The power consumption for these DACs is nearly negligible, and is included in the current drawn by the decoder, indicated in Table 3.2.

3.3.4. Simulation Results

In Figure 3.15, a post-layout simulation of the tracking ADC performance across corners is shown. Accounting for the expected delay, the full-scale, $5MHz$ input signal is kept between the two references by the quantizer's internal feedback. A mismatch in the comparators could

lead to a rather non-linear characteristic (not shown in Figure 3.15), however, as explained in Section 2.6.4, any non-linearity in the tracking quantizer does not influence the converter's performance significantly.

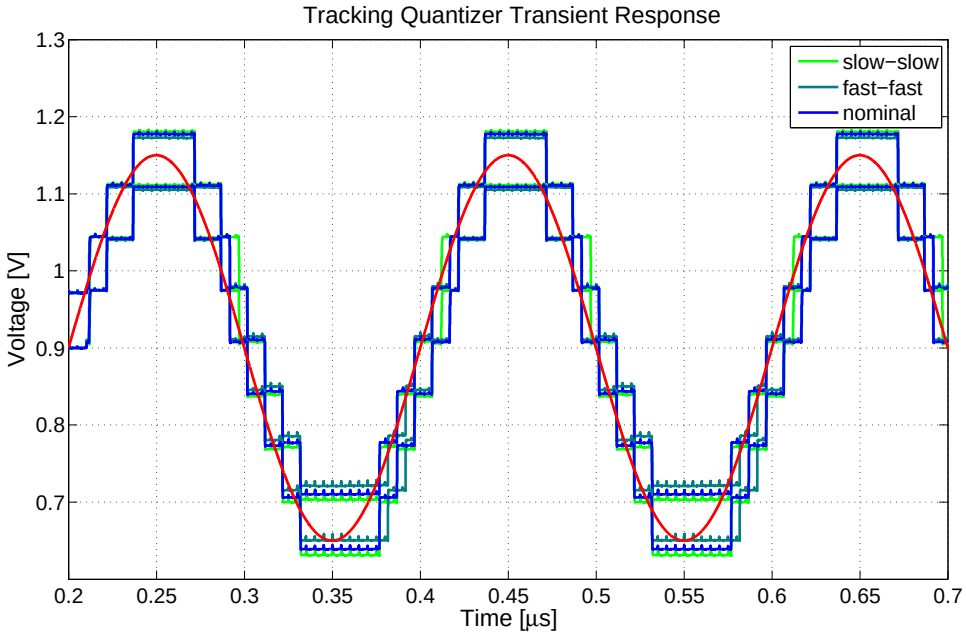


Figure 3.15: Tracking quantizer transient response (post-layout, over corners)

3.4. Other Blocks

Here, the few blocks left that complete the converter architecture will be discussed, before moving to the design and simulation of the pixel as a whole.

3.4.1. Biasing Block

In this design, each pixel was provided with its dedicated bias block, even though its area is quite large (around 12% of the pixel). This is inevitable, as wiring each pixel to a bias block placed outside of the array would lead to an unmanageable number of connections, while distributing voltages throughout the array leads to a very high sensitivity to any mismatch in the local current sources. Furthermore, any local disturbance on the supply line will be common to the amplifiers and the bias block, thus improving the PSRR of the system.
The implementation of the bias circuitry, shown in Figure 3.16, uses a known constant g_m biasing technique [45] as the core bias current generation block. This topology was chosen because it is theoretically insensitive to changes in the supply voltage and in the transistor parameters. In fact, the transconductance of the current sources is only dependent on the ratio between the aspect ratios of the two NMOS current sources (k) and the resistor value R:

$$g_m = \frac{2}{R}\left(1 - \frac{1}{\sqrt{k}}\right) \tag{3.7}$$

In order to reach the desired mode of operation, this block requires a start-up circuit, which lifts the NMOS gate voltage above ground, while lowering the PMOS gate potential.

The main current ($1\mu A$) generated by the constant g_m block is then mirrored, with the due ratios, and connected to the diode-connected MOSFETs within the amplifiers. Three bias

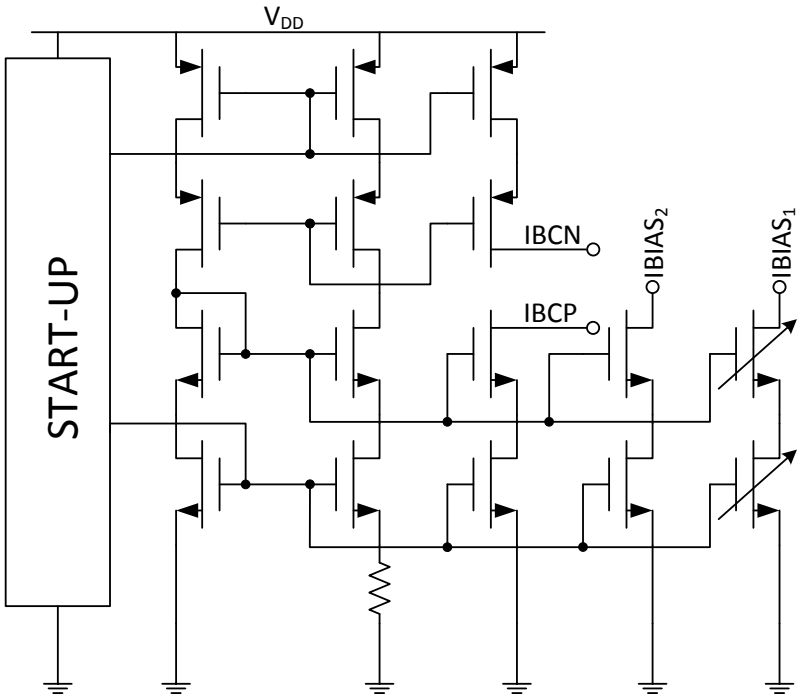


Figure 3.16: In-pixel constant-Gm biasing block.

currents are needed for the second stage, and only one for the first one. In particular, the latter can be tuned, so that it is possible to control the noise introduced by the amplifier. All current sources are cascoded, in order to boost their output impedance. This block consumes $21\mu A$ from the analog supply, including the current mirrors. A static supply voltage change will result in a variation in the main current, with a sensitivity of $13\frac{nA}{V}$.

3.4.2. Tunable Compensation RC

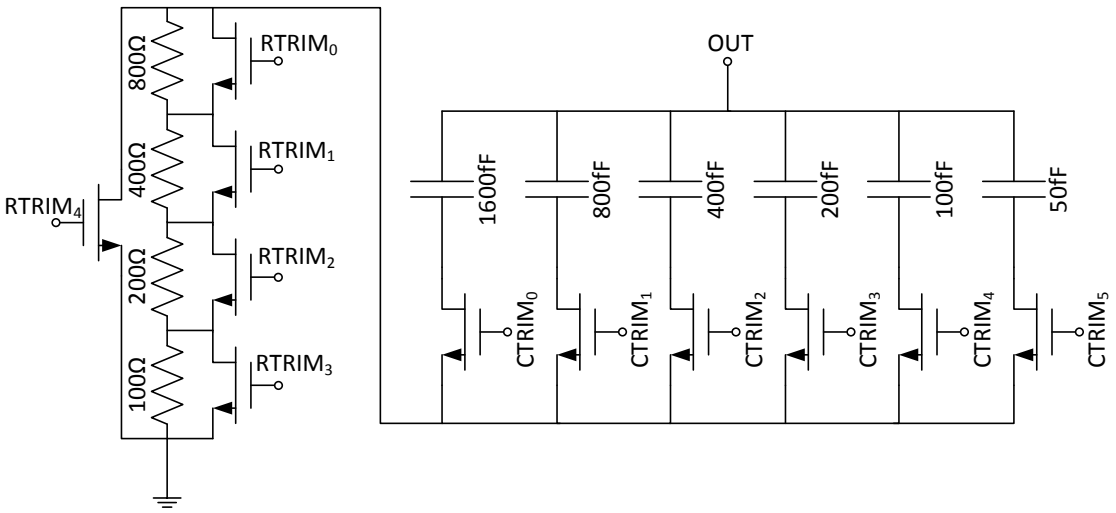


Figure 3.17: Schematic of the tunable capacitor C_c and resistor R_c.

The tunable C_c and R_c (see Figure 2.15) have been implemented as in Figure 3.17. In the in-pixel shift register (see Section 4.1) 6 bits were reserved for C_s, so that the full-scale is

$3.15pF$, which limits the maximum allowable parasitic capacitance C_s to this value. The LSB is $50fF$, which is roughly half of the width of the stable interval in Figure 2.20. Such a resolution should allow us to always find an optimum configuration where C_s is almost perfectly compensated for. The on resistance of the NMOS switches is kept low by using a large aspect ratio, additionally, it is likely that more than one capacitor will be selected, thus further reducing the equivalent resistance of the switches by having several in parallel. A binary weighted scale was chosen for its ease of control as linearity is not a concern in this case.

For the tunable resistor R_c, 5 bits were chosen. The first four bits ($RTRIM_{0-3}$) implement a binary weighted RDAC with an LSB of 100Ω and a full-scale of 1500Ω. The LSB was chosen after measuring the effect of a mismatch between R_s and R_c on the SNR. The full-scale is kept in the order of the $k\Omega$ as the influence of R_s was not visible in the available impedance measurements, which had a maximum frequency of less than $10MHz$. This suggests a cut-off frequency of at least a decade further, and a resistor value of $R_s < \frac{1}{2\pi f_{cut-off} C_s} = \frac{1}{2\pi 100MHz \cdot 2pF} \approx 800\Omega$.
$RTRIM_4$ is used to drastically lower the on resistance of the other NMOS switches, in case they are all selected, since these resistances, even if designed to be very low, can become relevant if there are four in series.

3.5. Pixel-Level Design

In this Section, the overall simulation results of the pixel as a whole will be presented.
The estimated current consumption is illustrated in Figure 3.18. The total post-layout current drawn by the supplies is $458\mu A$, which yields a total power of $824\mu W$. In the post-layout pie chart, the analog current includes that of the two stages and the biasing, while the digital current comprises that of the core digital blocks as well as the local clock buffers.
Figure 3.19 shows the spectra estimated by taking the FFT of the time-domain modulator's

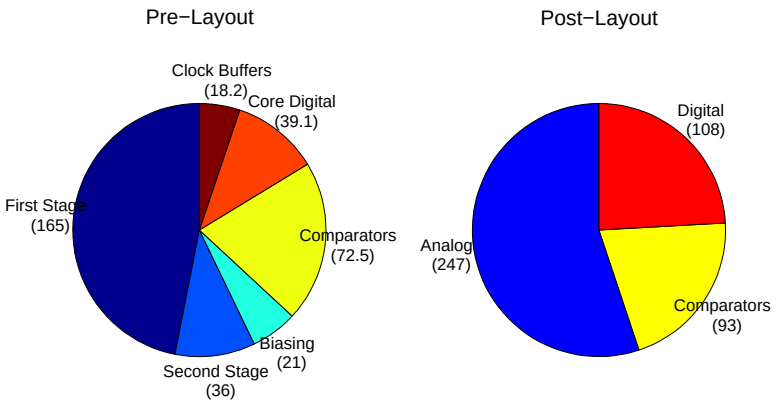


Figure 3.18: Pie chart representing the current consumption in μA for every block in the pixel.

digital output (post-layout), with and without transient noise. The spectra are very similar to the one obtained from system-level simulations (Figure 2.16), however the SQNR and consequently the SNR are slightly lower. Some harmonic distortion components are also visible, because of the first and second stage non-linearities. Figure 3.20 shows the AC behaviour of the loop filter; in particular, the transfer function from the voltage source internal to the trans-

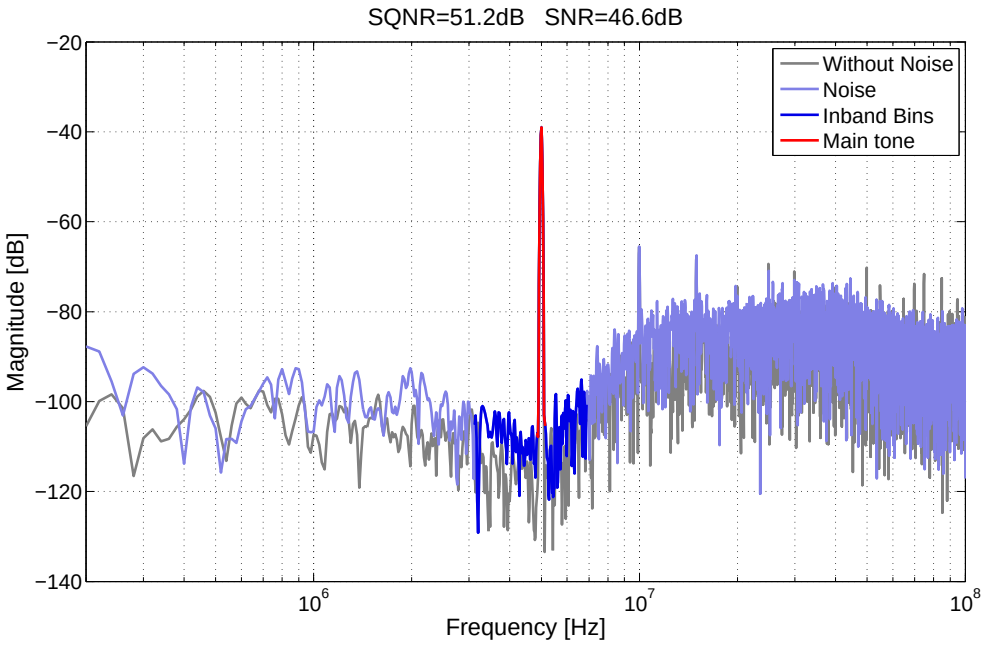


Figure 3.19: Spectra of the post-layout pixel-level simulations, without and with transient noise.

ducer model to the outputs of each stage are plotted when the gain is maximum (TGC=00). As expected, the first stage gain at resonance is roughly $\frac{R_F}{R} = 20dB$, while the output of the second stage lies above that of the first by $\approx 33dB$, which is the maximum gain for the second stage. Above $15MHz$, the pole of the second stage starts closing the gap between the two curves. The phase response for both curves starts from -90^o, as the stage is inverting and there is a zero in the origin, and it passes through -180^o at resonance, where the transducer does not contribute with any phase-shift. The phase then drops to -270^o because of the two resonant poles provided by the transducer. After that, because of the poles of the two stages, the phase drops further.

In Table 3.3, the contribution of the main sources of noise is provided. In particular, the noise at the output of the loop-filter, integrated in the bandwidth $[3.125MHz - 6.875MHz]$ has been computed. The resistor R within the transducer model contributes for almost half of the in-band noise, therefore fulfilling the requirements expressed in Section 2.5.5. The noise figure is therefore close to $3dB$. The first stage, as expected, has the largest noise contribution, amounting to a total of 30%, while the first stage's feedback resistor R_F adds 14% of the noise. The second stage and the biasing both contribute with 3% of the noise.
Since the lower boundary of the bandwidth of interest is in the MHz range, the contribution of flicker noise does not dominate, which helps keeping the area of the analog stages low.
It should be noted that this table only shows an estimation of the noise contributions, as it does not account for the modulator's feedback. In other words it is an "open loop" measurement. In order to get the complete picture, one would have to account for the entire thermal noise transfer function of each source to the loop filter output. For example, as described in 2.3.1, the in-band noise added by the first stage is only part of its contribution, as the noise at high frequencies can fold back to the bandwidth of interest if not properly filtered.

3.6. Element-Level Layout

One of the most important objectives for this project is being able to fit the modulator underneath each transducer element. In order for this to be possible, a careful layout procedure is

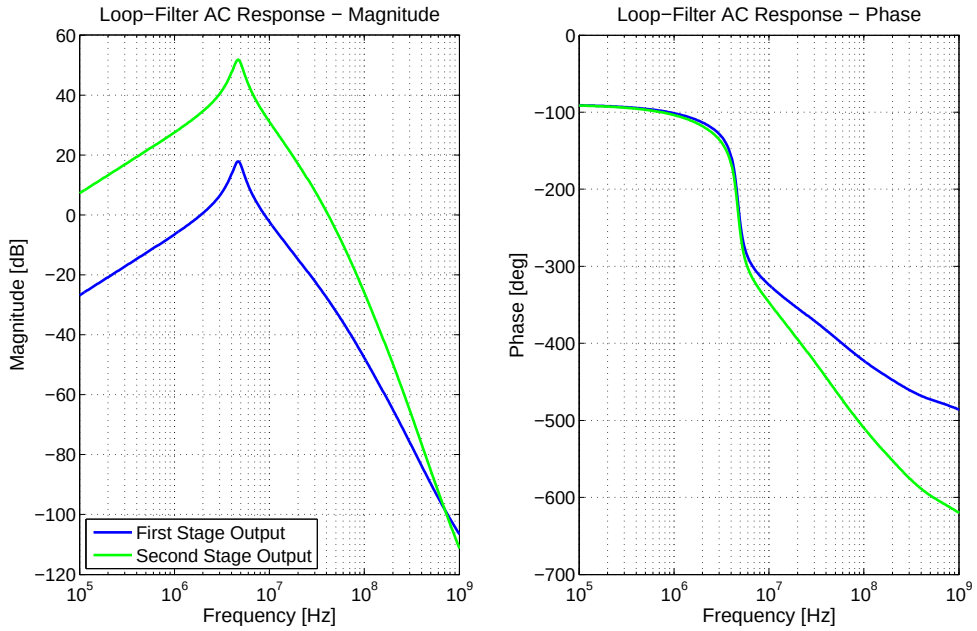


Figure 3.20: AC response of the loop filter.

Table 3.3: Noise contributions of the critical transistors and resistors within the pixel.

Device	Noise Type	Output Noise	Percentage
R_{TRANSD}	Thermal	$4.66mV_{rms}$	46.58%
R_F	Thermal	$2.56mV_{rms}$	14.04%
$PM_{1^{st},1}$	Thermal	$1.84mV_{rms}$	7.37%
$PM_{1^{st},2}$	Thermal	$1.74mV_{rms}$	7.37%
$NM_{1^{st},1}$	Thermal	$1.74mV_{rms}$	6.5%
$NM_{1^{st},2}$	Thermal	$1.74mV_{rms}$	6.5%
Biasing	Thermal	$1.14mV_{rms}$	3%
Second Stage	Flicker	$1.12mV_{rms}$	3%
$NM_{1^{st},1-2}$	Flicker	$1mV_{rms}$	2.3%
Total	Thermal/Flicker	$6.83mV_{rms}$	96.7%

just as important as a mindful system-level design. A render of the pixel layout is shown in Figure 3.21, where the main functional blocks have been highlighted.
The layout is quite tight (all the density checks are met automatically) and there is almost no free space, as you can see in Figure 3.22.

The majority of the area (44%) is occupied by capacitors: the large test capacitor on the top right, the second stage feedback, the quantizer DACs and the feedback capacitors on the very top. Both Metal-Insulator-Metal (MIM) and Metal-Oxide-Metal (MOM) capacitors were used: MIM caps (pink) were preferred when a high density is required, for example in the feedback capacitors and for the tunable capacitor, while MOM were selected when a low unit capacitance size was needed (second stage feedback and quantizer DACs). The MIM capacitors employ Metals 5 and 6, thus their polarity was chosen such that the quieter node is implemented in Metal 6, to avoid any cross-talk with the transducer.
The two OTAs, the comparators and the second stage feedback capacitors were laid out using the common centroid technique, so that the mismatch effects are minimized, while the layout of the logic blocks was synthesized using Encounter.

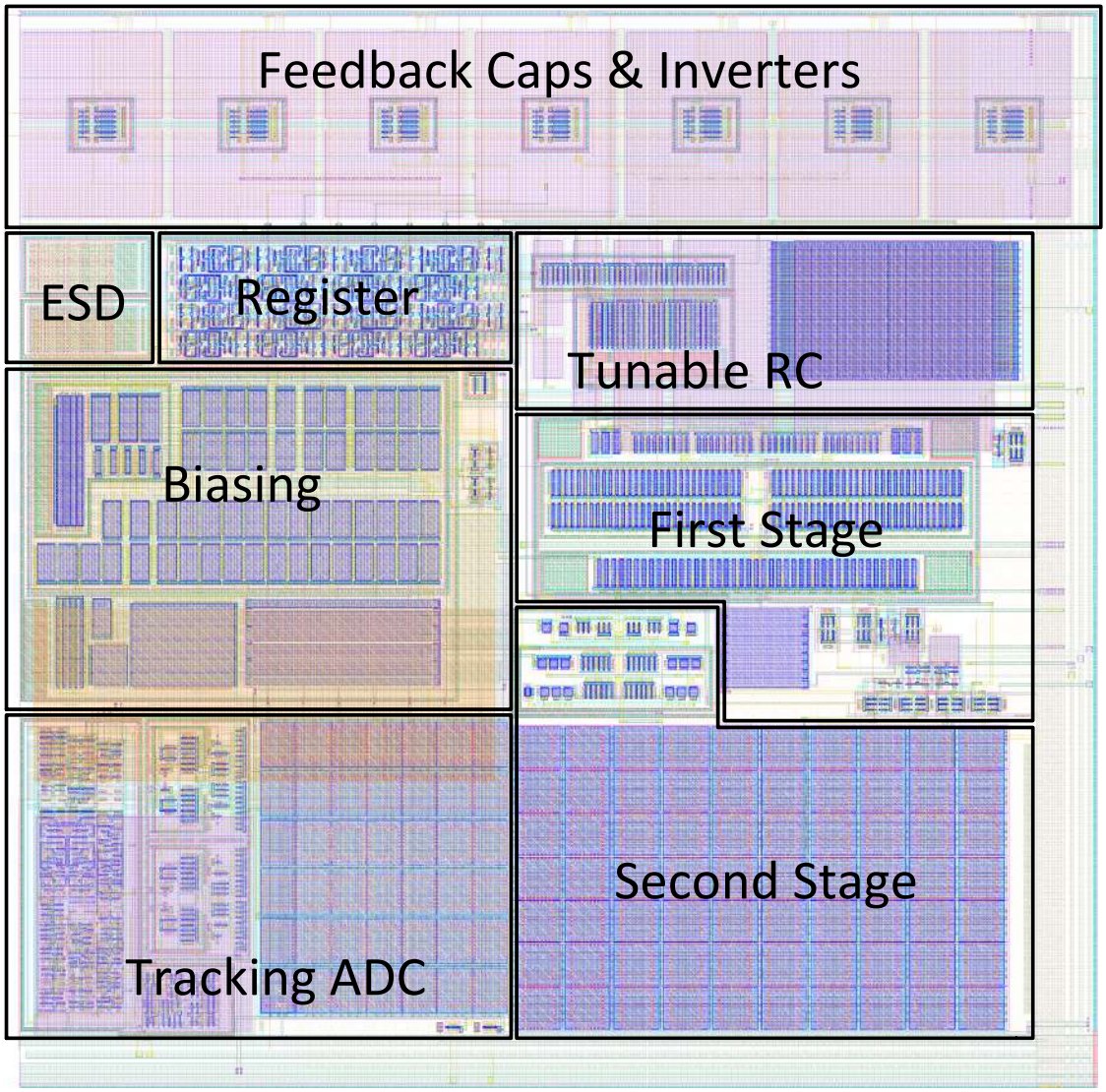


Figure 3.21: Pixel Layout

The passivation opening needed to connect each element to the relative transducer is visible as an orange shadow on the center-left of the layout. This pad is not horizontally centered, however, the transducers are built so that the pad is in their center, thus there will be a slight horizontal offset between the ADC area and that of the relative transducer, as they will not be perfectly overlapping. This could lead to some cross-talk between adjacent pixels, in case they are not properly shielded.

This pad, built with 20kÅ thick Metal 6, is the most sensitive node in the whole system, thus it should be kept far from digital signals or high-swing nodes, or shielded from them. For this reason, static circuits were placed directly underneath the pad: the biasing block and the register, which only changes its outputs during the configuration phase. Additionally, the entire left side of the chip is covered with a Metal 5 shield, to further protect the pad. The feedback capacitors are also protected by a full layer of Metal 4 below them.

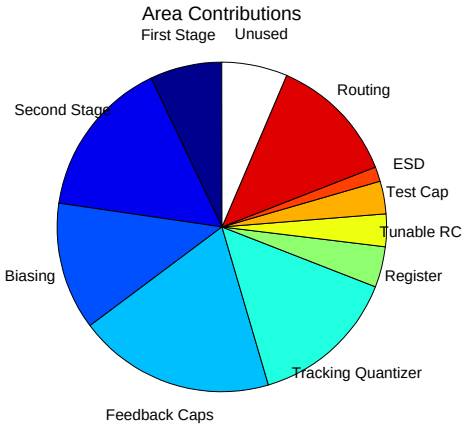


Figure 3.22: Area contributions of every block.

The signal path follows a clockwise cycle, starting from the previously described pad and ending with the feedback capacitors. The tracking quantizer's 7-bit output, which should reach the feedback inverters, is routed in Metal 4 and passes over the biasing block, which is protected fully by a Metal 3 shield.
The pixel layout was designed to be modular, so that if an array of these elements is created, the shared inputs and outputs connect automatically. An easy way to reach the required modularity is having a grid-like voltage distribution. In this case, the East and South edges of the pixel were reserved for this purpose: the East edge distributes voltages through Metals 2 and 4 (overlapping), while the South edge uses Metals 1 and 3, so that the vertical and horizontal lines are interleaved.

Finally, Table 3.4 shows a list of all the inputs and outputs of each pixel. All of these I/Os are shared between pixels, except for those dedicated to the configuration loading (RCLK_IN, RDATA_IN, RCLK_OUT and RDATA_OUT), which are routed from pixel to pixel in a daisy chain (see Section 4.1), and the two digital outputs of the tracking quantizer (OUT0 and OUT1).

Table 3.4: I/O connections for each pixel.

Power Pins	VDDA, VDDC, VDDCORE	3
Ground Pins	GNDA, GNDC, GNDCORE, GROUND_FOIL	4
Analog References and Signals	VREF+, VREF-, VDAC+, VDAC-, V_CM, RF_IN	5
Digital Inputs	CLK, CLK2, RST, TGC0, TGC1, RCLK_IN, RDATA_IN	5
Digital Outputs	OUT0, OUT1, RCLK_OUT, RDATA_OUT	3
Total		20

3.7. Conclusion

In this Chapter, a description was provided of the implementation details for each of the key blocks that compose the element-level $\Sigma\Delta$ ADC. A heavy use of current-reuse topologies help to keep the analog power low, while the digital design was optimized to minimize the delay of the key components, as well as the power. An effort was made to provide the reader with all the simulation results needed to judge the performance, as well as the robustness of the system. The following Chapter 4 will describe the path that lead to the full chip implementation.

4

Chip-Level Design

After having discussed the main circuits included in every element, in this Chapter, the blocks which concern the chip as a whole will be described. This Chapter also includes some chip-level considerations such as details on the layout and positioning of the transducers.
The array size was decided to be 5×4 elements. A smaller size would not allow for a proper acoustic evaluation, such as the measurement of the delay between two elements, which would be too low to measure if the maximum distance between elements is reduced. Furthermore, a smaller array would make the transducer fabrication process significantly more challenging. On the other hand, a larger array would require a greater chip area and cost. It would also make the routing of the modulator outputs more challenging and in general the complexity of the array-level design would increase.

4.1. Shift-Register

In order to load the configuration data for the entire chip, a shift register in a daisy chain configuration has been employed, as shown in Figure 4.1.
The big squares represent the element circuits, while the small one stands for the register used to select the wanted signal as an input of the on-chip analog buffer (see Section 4.4). Finally, the last flip-flop is used to select which sub-array should be connected to the LVDS transmitters (see Section 4.2). The output of this last flip-flop is buffered and connected to a pad, in order to be able to establish if the shift-register is working properly. Each section of the register, highlighted in the top part of the figure, is composed of several D-flip-flops in series, and some local buffering inverters to ensure a correct set-up and hold timing. The chosen protocol only requires the clock and the data lines and omits the enalble line found in a more traditional SPI protocol. A 3-wire SPI would require an additional set of registers, the area of which would take a toll in the element's constrained area budget.
The data and clock lines are routed in parallel through the array, in a snake-like fashion, but in opposite directions, so that the last register is always the one which gets clocked first.
The in-pixel shift-register sections are composed of 16 flip-flops, the function of which is summarized in Table 4.1.

4.2. LVDS Outputs

In order to convey the modulator's output bitstreams out of the chip and into the data acquisition system, 20 LVDS transmitters were implemented on chip, as the available standard I/O pads were not fast enough. This communication protocol was selected for its high speed and its differential nature, which makes the signal less sensitive to digital cross-talk or electro-

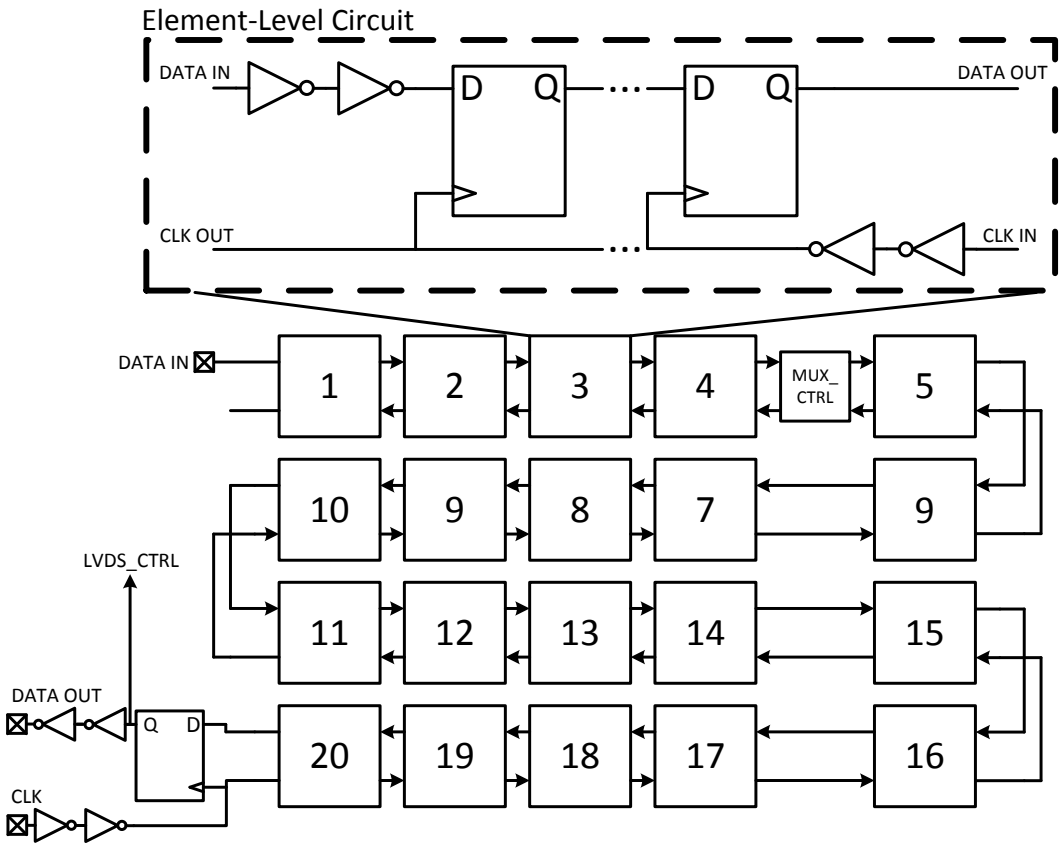


Figure 4.1: Block diagram of the shift register daisy chain.

Table 4.1: Shift-register bits.

Name	Function	#bits	Per pixel/Global
$CTRIM_{0-5}$	Trims the compensation capacitor C_c	6	Per pixel
$RTRIM_{0-4}$	Trims the compensation resistor R_c	5	Per pixel
$ITRIM_{0-1}$	Trims the first stage's bias current	2	Per pixel
$FCAP$	Adds a 50fF cap in the first stage's feedback	1	Per pixel
$DISABLE$	Grounds the transducer and disables the clocks	1	Per pixel
$TEST$	Connects the test capacitor	1	Per pixel
MUX_CTRL_{0-4}	Selects the input for the PMOS analog buffer	5	Global
$LVDS\_CTRL$	Selects the array section to be connected to the LVDS	1	Global
Total		16	

magnetic coupling. Every LVDS transmitter drives two pads, while the array is composed of 20 elements, with two digital outputs per element. This would bring the pad count for the LVDS outputs to 80, thus increasing the chip area and cost.

In order to avoid such a high number of output pads and to reduce the substantial power drawn by the transmitters, the whole array was divided in two 10-element sections horizontally. Either one of the two sections is connected to the 20 LVDS transmitters, depending on the LVDS_CTRL bit in the shift-register.

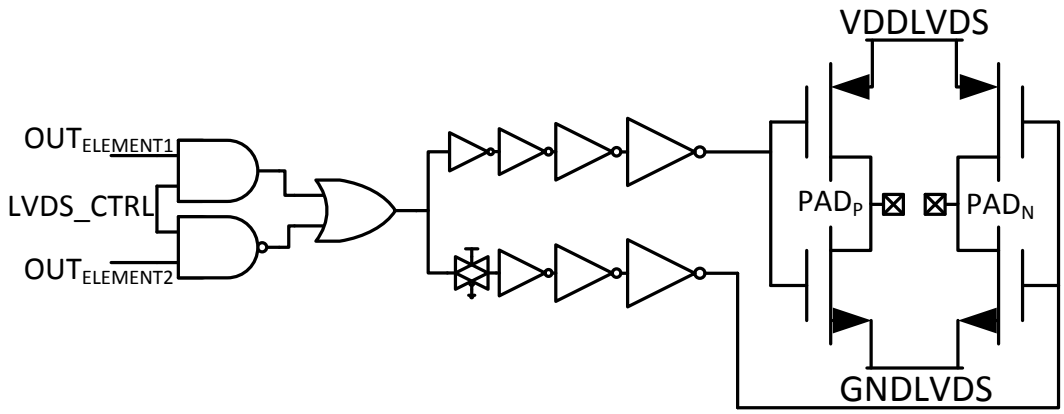


Figure 4.2: Schematic of the implemented LVDS driver and pre-driver.

The schematic diagram of each LVDS transmitter is shown in Figure 4.2: the two outputs from the top and the bottom parts of the array are multiplexed through the LVDS_CTRL bit. The result is fed into a pre-driver with a high driving capability, which divides the signal in two parts with different polarities, but with the same delay. Finally, the pre-driver's outputs control the LVDS circuitry, which in turn drives the pads.

The voltages for VDDLVDS and GNDLVDS were chosen to be 1.2V and 0.6V respectively, so that the NMOS and PMOS get the same overdrive voltage. The average current drawn by the pre-driver is $160\mu A$ per channel, while the LVDS driver consumes $1.8mA$ per channel, assuming a load capacitance of $20pF$ and accounting for the fact that the output does not necessarily toggle for each clock cycle.

The power consumption for this circuit was not optimized, as a mature product based on this test chip would have on-chip or in-pixel decimation filters to reduce the sampling rate and possibly digital beamformers, which limit the amount of information that has to be transmitted to the data acquisition system.

4.3. Clock Distribution

Since the sampling frequency for the modulators is rather high, it is important to carefully distribute the two clock signals, so that they reach every element at the same time. These clocks are needed for the comparators and for the flip-flops within the tracking quantizer. Since the two clocks need to have a precise phase difference, they must be routed in parallel, so that the total delay is the same for both. The implemented clock distribution strategy is shown in Figure 4.3, where the grey squares represent the elements.

A tree-like routing strategy brings the clock signals from the pads to the vertical edges of the array, then these signals are injected inside the array without further repeaters. This is only feasible as the array is quite small and the propagation delay along the clock lines between two adjacent pixels is simulated to be less than 0.1% of the clock period. For bigger arrays or higher sampling frequencies, an H-like distribution, with repeaters inside the main array should be adopted.
The clock distribution inverters consume an average of $2mA$ from a 1.8V supply, however, this current can be drastically reduced with a careful sizing of the inverters.

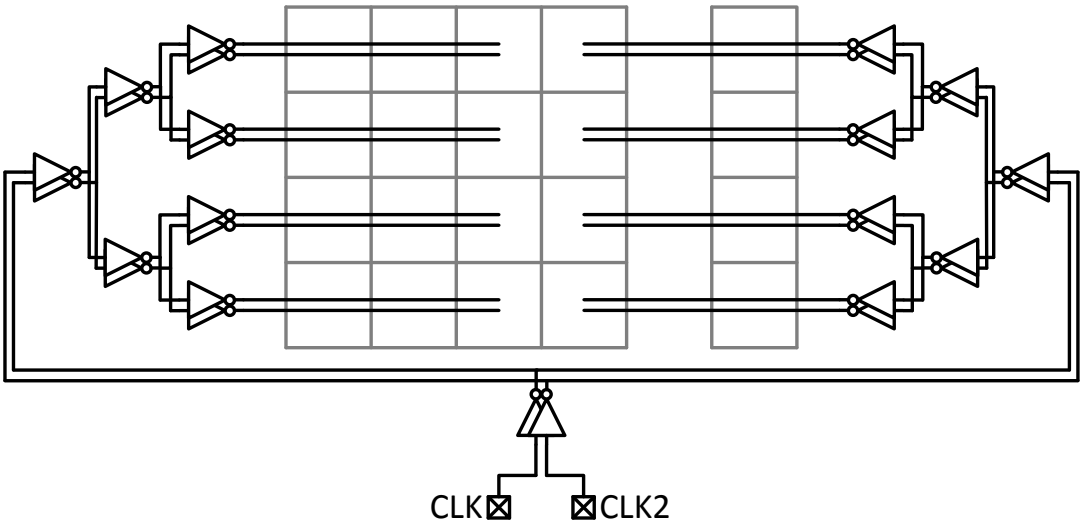


Figure 4.3: Block diagram of the array's clock distribution strategy.

4.4. Design for Test

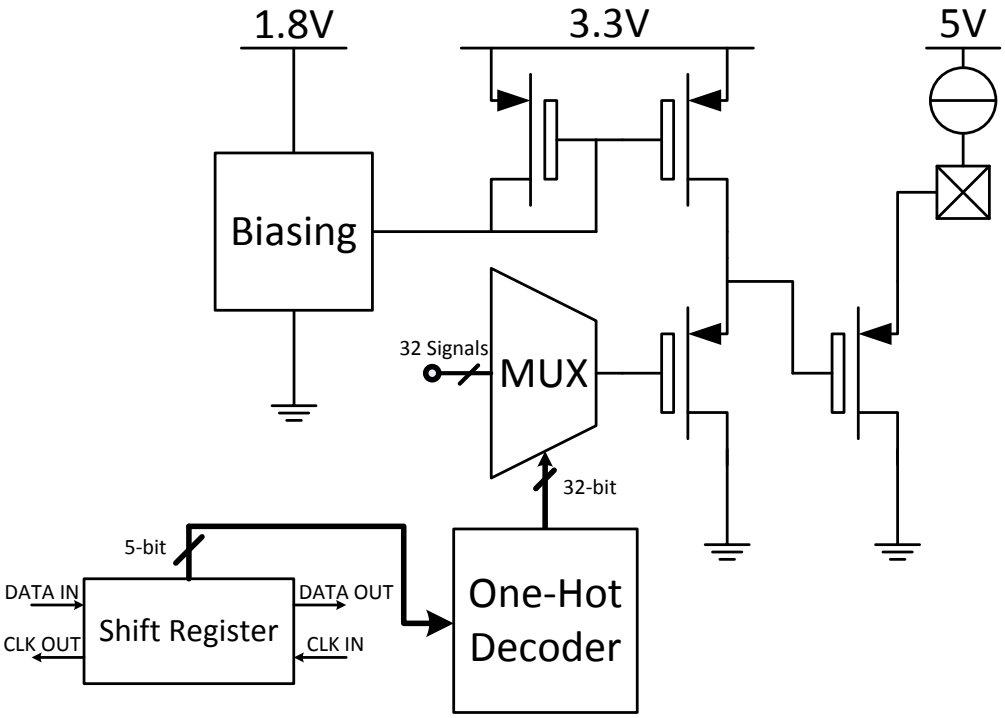


Figure 4.4: Block diagram of the PMOS buffer designed for test purposes.

The main testing tool that was included in the chip is shown in Figure 4.4. Its main purpose is to access the most critical nodes of the modulator and buffer them, so that they can be probed through an oscilloscope.
The desired channel is sent by the measurement equipment together with the configuration data for the ADCs, and stored in a 5-bit shift register between elements 4 and 5 (the num-

bering refers to Figure 4.1). A 5-to-32 one-hot decoder turns the binary-weighted bits in the register into a 32-bit code which selects the signal to be buffered, through a 32-to-1 analog multiplexer.
Only the four right-most "test" elements (number 5, 6, 15, 16) are connected to the multiplexer, so that for each of these pixels it is possible to monitor the two inputs of the first stage, its output, the output of the second stage and the two digital outputs. These points are highlighted in blue in Figure 4.5. Additionally, all the analog references can be selected by the multiplexer, including the analog ground.

The buffer is designed to provide a very high bandwidth of 200MHz, while offering a low input capacitance, as a large one could load excessively one of the critical nodes of the element, thus disturbing its normal operation. At the same time, the buffer should be able to drive a large capacitance, estimated to be in the order of $20pF$, formed by the pad, the PCB traces and the oscilloscope probe. For the sake of simplicity, a PMOS buffer topology was chosen, however, a single stage is insufficient to provide the desired specifications. A transconductance of roughly $g_m = 2\pi C_{out} f_{cut-off} = 25mS$ is needed to meet the bandwidth and loading requirements, but this g_m can only be provided by a rather large PMOS, which in turn will offer a high input impedance (in this case $1.7pF$). Thus, a pre-buffering stage is added using the same topology, which drives the large PMOS buffer while keeping the input capacitance to only $50fF$.

In order to accommodate for the entire voltage range (0 to 1.8V), these followers were implemented using thick-oxide $3.3V$ transistors and were provided with progressively higher supply voltages, so that a large voltage on the input will not turn off the PMOS. The first stage is biased by a simplified version of the circuit described in 3.4.1, and is supplied by a dedicated 3.3V supply. On the other hand, the second stage is biased by an external off-the-shelf current source of around $7mA$ on the PCB, through a pad stripped of any ESD protection. The output offset due to the two V_T drops can be calibrated out subtracting to every measurement the output voltage obtained while selecting the analog ground as an input.

Another test feature that has been introduced is an analog input signal (RF_IN) coming from an external function generator, distributed to every element. This signal reaches one terminal of a $300fF$ capacitor, present in every pixel. The test capacitor is connected on the other side to the first stage's inverting input through a switch, controlled by a dedicated bit in the shift register (the $TEST$ bit shown in Table 4.1). This technique allows for the injection of an electrical input signal for the modulator, which will keep its NTF unchanged (the test capacitor will go in parallel with C_s and needs additional compensation), so it is possible to evaluate the noise shaping behaviour of the converter even without an acoustic input signal.

4.5. Power Domains

Many different power domains were used for the implementation of this chip. In particular, the comparators were given a dedicated supply, so that they will not disturb the analog core, while also being protected from the switching transients in the digital domain. The 1.8V digital supplies were separated in four parts: the element-level digital circuitry has a separate supply, so that the current drawn by it can be measured. The pad-ring supply has no connection to the core, so that it can work at a different voltage level. The supply for the clock buffers and data repeaters and that of the LVDS pre-drivers were separated, so that, if needed, they can be tuned independently. A summary of the employed voltage domains is provided below.

- **VDDA, GNDA:** Core supply for the element's analog section and shift registers. Nominal voltage: 1.8V.
- **VDDC, GNDC:** Supply voltage dedicated to the tracking quantizer's comparators. Nominal voltage: 1.8V.
- **VDDCORE, GNDCORE:** Supply voltage for the in-pixel digital circuitry. Nominal voltage: 1.8V.
- **VDD3V3, GND3V3:** Supply voltage for the PMOS buffer's first stage. Nominal voltage: 3.3V.
- **VDDD, GNDD:** (double pads) Provides voltage for the global clock buffers and data repeaters. Nominal voltage: 1.8V.
- **VDDDESD, GNDESD:** Pad-ring post-driver digital supply. Nominal voltage: 2.5V in accordance with the FPGA specifications.
- **VDDLVDS, GNDLVDS:** (double pads) Supply voltage for the LVDS drivers. Nominal voltage: 1.2V and 0.6 respectively.
- **VDDLVDSESD, GNDLVDSESD:** (double pads) Used for the LVDS pre-drivers. Nominal voltage: 1.8V.

The whole designed chip has been schematized in Figure 4.5

4.6. Chip Layout & Fabrication

Figure 4.6 shows the full chip layout which has been taped out. The main blocks are highlighted by black outlines. The chip measures $2.38mm \times 2.1mm$ for a total chip area of $5mm^2$. The active area, however, is much smaller, and the chip size is mainly defined by the number of pads (82). The main 4×4 array is placed exactly in the center of the chip. The four test ADCs are placed on the right of the main array, such that they will not have any transducer above them. This will allow to diagnose any problems related to unwanted cross-talk between the transducers and the ADCs underneath them. The test circuits (see Section 4.4) are located between the main array and the test ADCs.

The main array and the test ADCs are surrounded by many metal rings: each of them is connected to a pad in the pad-ring and it distributes a voltage across the whole array. The two outputs of every modulator reach the periphery of the main array, where they are buffered by data buffers, which in turn drive the lines that finally reach the LVDS transmitters (see Section 4.2), located close to the pad-ring. Each of the transmitters drives two pads, for a total of 40, located on the top and left edges of the chip. The bottom edge is dedicated mostly to the digital signals and supplies, while the right one hosts the analog pads.

Immediately on the left of the main array, four pads (highlighted in green in Figure 4.9) were wired directly to four corresponding pads in the pad-ring, without any connection to any circuit. Both the pads in the pad-ring and those beside the main array had their ESD protection removed, so they can be used as transmit elements, which requires a large voltage across the transducers. These elements can also be used to measure their electrical impedance, simply by connecting them to an impedance analyzer. The results of this operation will be provided in Chapter 5

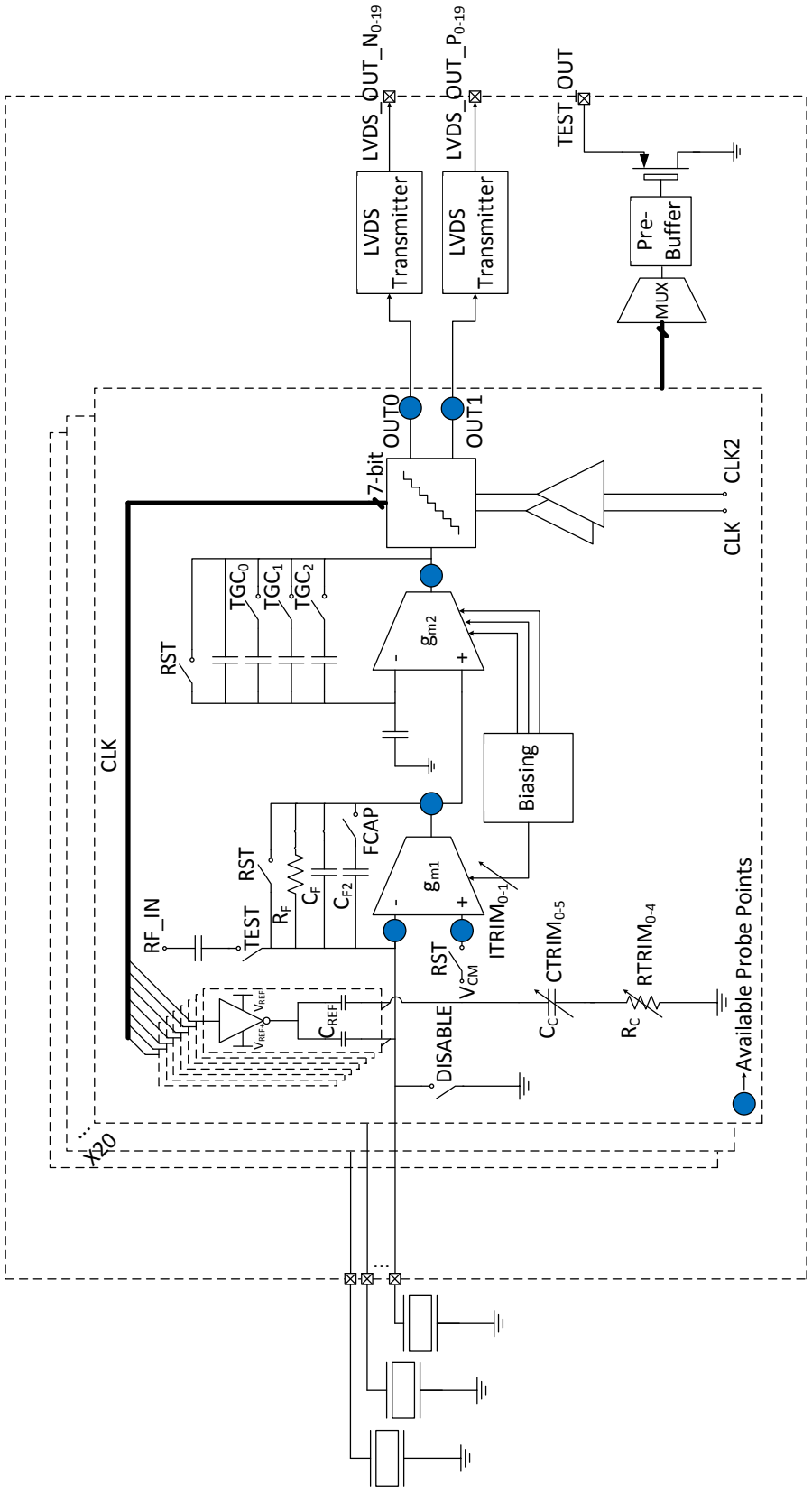


Figure 4.5: Block diagram of the complete chip

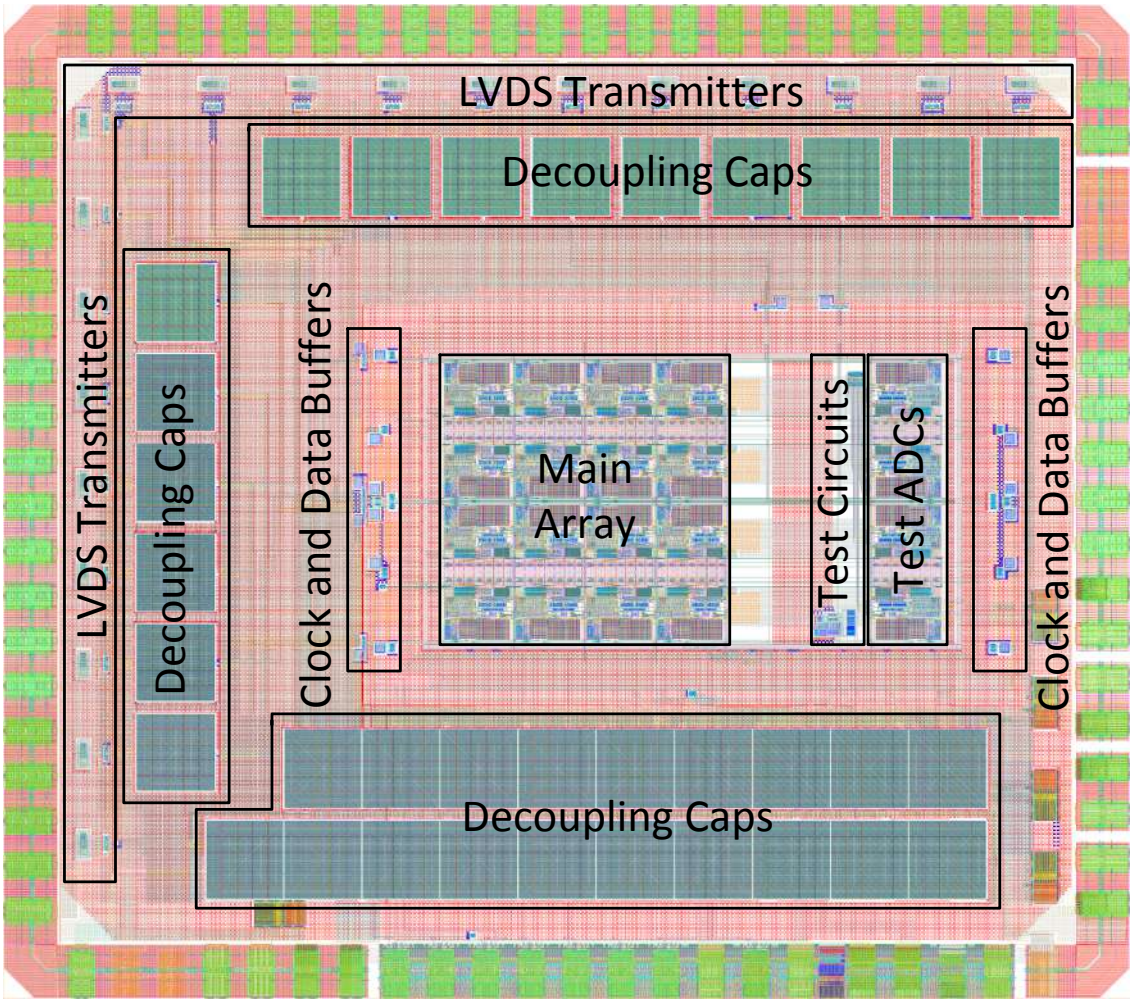


Figure 4.6: Full Chip Layout

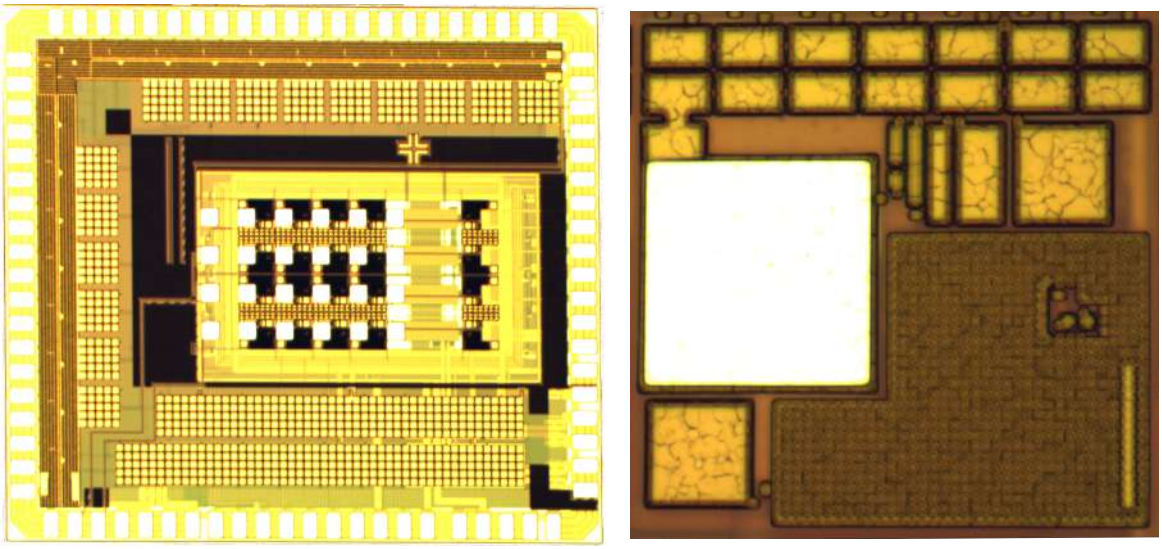
Finally, the empty space between the pad-ring and the core circuits was filled with high density decoupling capacitors (MIM and MOM overlapping): the ones on the bottom are connected in triplets to each supply (except for the $3.3V$ one), while the capacitors on the top and left are dedicated specifically to the decoupling of the LVDS supplies, which draw the most instantaneous current.

The chip has been fabricated in TSMC $0.18\mu m$ MS-RF technology. The micrograph of the ASIC is shown in Figure 4.7a, together with a close-up of the pixel (Figure 4.7b), which can be easily compared to Figure 3.21. Here, a few structures can be distinguished, implemented in Metal 6: the pad, the feedback capacitors on the top, the tunable capacitor C_c on the top-left and the shield which covers most of the analog and digital circuitry.

4.7. Transducer Integration

In this Section, more details will be provided on the transducer fabrication process and its integration on the ASIC [14]. A section view of the complete acoustic stack is shown in Figure 4.8.

- As described in Section 3.6, during the standard IC fabrication, an opening in the passivation layer is created for every ADC, so that the designated input pads are exposed for further post processing.



(a) Chip micrograph.

(b) Zoom on a pixel.

Figure 4.7: Micrograph of the fabricated ASIC.

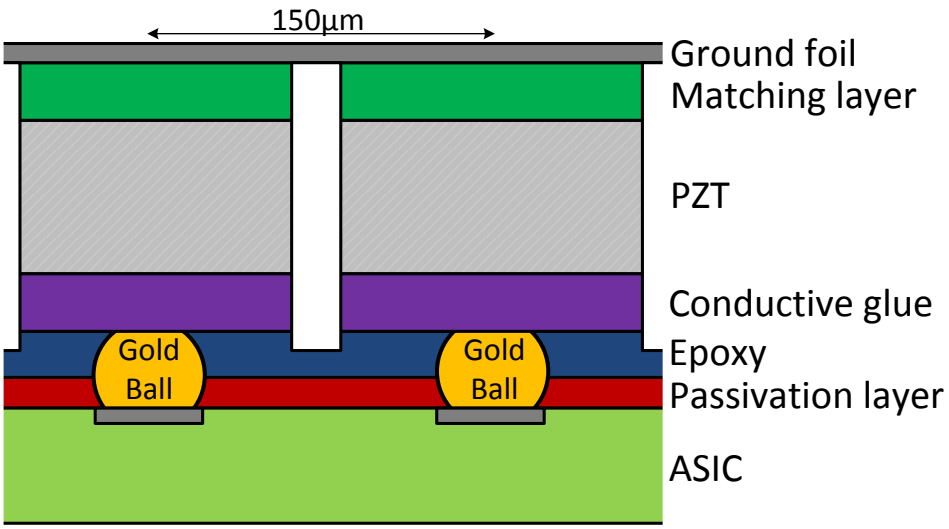


Figure 4.8: Section view of the acoustic stack, divided into layers.

- On each pad, a gold ball is bonded in place.
- The array area is covered with an epoxy layer, acting as a buffer for the dicing procedure.
- The gold balls and the epoxy layer are grinded down, in order to leave a flat surface for the following steps.
- The conductive glue, PZT and matching layer are sequentially applied, forming the acoustic stack.
- The acoustic stack is diced, leaving behind stand-alone transducers.
- Finally, an aluminum ground foil is applied on the elements.

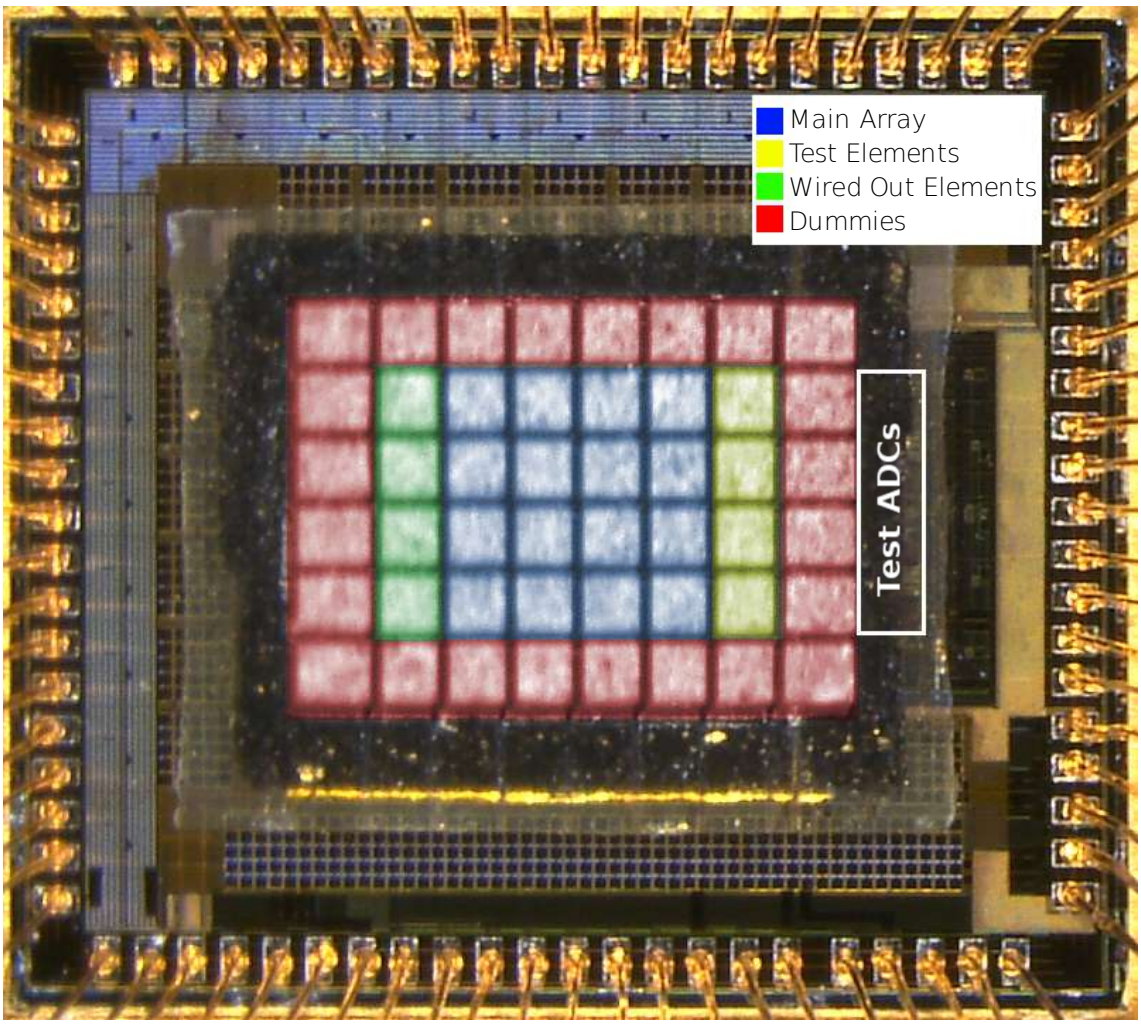


Figure 4.9: Micrograph showing the transducers on top of the ASIC.

Figure 4.9 shows a picture of the ASIC with PZT transducers integrated on top of it, before the application of the ground foil (hence the single elements are still visible).
The transducers connected to the main array are highlighted in blue, while the yellow ones are wired to the relative test ADCs on the right side of the acoustic stack.
The green elements are directly connected to four pads in the pad-ring. Finally, the active transducers are surrounded by a ring of dummy elements, in order to provide a homogeneous environment also for the transducers on the edges of the array.

4.8. Conclusion

In this Chapter, the procedure that leads from a single modulator to a full chip has been explained. In particular, the process of building the array has been desribed, together with the configuration technique, the strategy that allows each modulator's output to be transmitted outside of the chip and the test features that were added for troubleshooting. Finally, the chip-level layout and the additional fabrication steps needed to build the piezoelectric transducers on top of the ASIC have been presented.
Chapter 5 will focus on the test set-up and the measurements results.

5 Measurements

In this final Chapter, the test setup, custom built to measure the previously discussed ASIC, is briefly presented. The measurement results will follow, which will be compared with the simulations presented in Chapter 3.
Before diving into the measurements, a quick word on the employed samples is due. A total of seven chips were measured, four of them have transducers built on top, but none of them can be considered completely "healthy":

- **Samples 1EL, 2EL, 1TR, 2TR:** These chips suffered a catastrophic failure: they were connected to a bench-top voltage reference, which, while powering down, caused the chips to fail. The measured impedance between $VDDA$ and ground is about 2Ω, making these samples completely useless. The clear mechanism that caused this failure is still unknown, but all the samples were healthy before this occurred.
- **Sample 3EL:** Working chip without transducers, used for electrical measurements described in Section 5.3.
- **Sample 3TR:** This chip is electrically functional, but an error in the dicing procedure made the top half of the array unusable, and many elements on the south-east corner have been damaged during the bonding process, leaving only 4 measurable elements. Furthermore, the acoustic stack has been exposed to water, which degrades the isolation between channels. Despite these limitations, this sample was used for the measurements discussed in Section 5.5.
- **Sample 4TR:** Finally, in this chip the $VDDLVDSESD$ pin (see Section 4.5), which powers the LVDS pre-drivers (see Section 4.2), draws a high current and causes a malfunction. However, the chip operates well if this pin is left unconnected. The digital LVDS are then not functional, however, the performance of the test ADCs can be evaluated through the on-chip buffer. In particular, the first stage inputs can be probed, and will show the modulator's DAC voltage, which will have the same shape and spectrum as the reconstructed digital output. This sample was used for the measurements in Section 5.4.

5.1. Measurement Setup and PCB Design

Let us start with a description of the measurement set-up that was designed and built. A schematic representation of the full test set-up is presented in Figure 5.1, while the layout of the PCBs is provided in the appendix A. There are a number of functionalities to be implemented

in order for the chip to work properly and to conveniently communicate with a computer, where further processing takes place:

1. Provide all the supply and reference voltages (see Section 4.5).
2. Drive the on-chip buffer described in Section 4.4 with a constant current source.
3. Create the reference voltages V_{REF+} and V_{REF-}, visible in Figure 2.15
4. Buffer the chip's high speed LVDS outputs.
5. Generate and send the configuration to the chip, together with the rest of the digital signals: the two clocks, the reset and the TGC setting.
6. Acquire, save and send to the computer the buffered LVDS outputs.
7. Generate the test input analog voltage, presented at the end of Section 4.4.
8. Acquire and present the analog output voltage of the on-chip buffer.

Many of these functions can be implemented using a custom built PCB which directly connects to the chip. Since the ASIC needs a post-processing step for the transducers to be built on top of it (see Section 4.7), a conventional packaging cannot be employed. Instead, the chip, after the transducer integration step, is directly glued and bonded to a carrier PCB, which is referred to as daughter board. The design of this PCB is kept very simple, as any component it hosts should be soldered manually for each chip. Thus, the daughter board only contains 30 decoupling capacitors, that need to be as close to the chip as possible, and two high-speed connectors, used as both an electrical and mechanical link to the mother board. The top part of the daughter board, which hosts the ASIC, is kept flat (all the components are placed on the bottom side); this is important as during acoustic measurements, a water bag will be placed on top of the acoustic stack (see Figure 5.11), so there should be nothing blocking its way.

Another PCB, the mother board, powered through a $6V$ bench-top power supply unit, has been designed to provide the first four of the functionalities listed above. This board hosts 20 voltage regulators, with a wide range of specifications, to implement function 1. Function 2 is solved by a current reference IC on the mother-board. For the third functionality, a pair of DACs were used, followed by two OP-AMPs in buffer configuration. This implementation is necessary as the reference voltages should follow an exponential curve during the receive phase, as explained in Section 2.6.5. Finally, function 4 is implemented using 20 high-speed LVDS buffers.

The digital inputs and outputs (functions 5 and 6) are managed by an FPGA (ALTERA Cyclone IV GX Development Board), which is linked to the mother board through two high-speed connectors. A PLL in the FPGA provides the clocks for the modulators, the on-chip shift register and the DACs on the mother board. In order to keep the communication between the FPGA and the computer as simple as possible, an RS-232 protocol was selected. The configuration data is sent by the computer through an USB cable, using Matlab's Serial Communication Toolbox. The data is then converted in RS-232 protocol by an FTDI chip on the mother board, which in turn sends it to the FPGA. The modulators' output data, after having been acquired by the FPGA LVDS receivers and stored in a FIFO, follows the opposite flow from the FPGA to the computer.
The DACs on the mother board are programmed directly by the FPGA, through an SPI interface.

The development board only has 17 LVDS input channels, which is not sufficient to acquire data from the whole array simultaneously, thus some multiplexing was implemented on the mother-board.

Functions 7 and 8 are implemented by an arbitrary function generator and a digital oscilloscope respectively. Both of these instruments are connected to the computer and controlled via Matlab. The function generator is also connected to the FPGA, which provides it with an external clock reference and a trigger signal, in order to synchronize the input signal generation with the start of the receive phase. Thus, the entire set-up is fully automated.

5.2. Transducer Impedance Characterization

5.2.1. Measurements in Air

Before moving on with the description of the modulator's measurement results, an impedance characterization of the fabricated transducers is presented, as it is a key piece of information to understand the measurements that will follow.
A total of four samples went through the transducer integration steps described in 4.7. Each sample has four transducers directly connected to the pad-ring, thus the total number of measurable transducers is 16. However, only 11 elements show a resonance, while the others display a capacitive behavior.
Figure 5.2 shows the magnitude and phase response of the transducers, measured in air by an impedance analyzer. These plots show two evident resonant modes, one around $3.75MHz$ and the other at $5.65MHz$, both with a relatively high Q factor. Because there are two resonances within the bandwidth of interest, the single-resonance model shown in Figure 1.4 used so far is no longer viable. Thus, the second resonance is modeled by an additional RLC branch, in parallel with the first. The best fitting model parameters, obtained using the ZView software, are listed in Table 5.1, together with their standard deviation.
The transducer's parallel capacitance C_s is a fundamental parameter for this design, thus it

Table 5.1: Statistical data of the electrical transducer model parameters, measured in air (11 measured transducers).

	C_s	C_1	L_1	R_1	C_2	L_2	R_2
Mean	$1.463pF$	$472fF$	$1.8mH$	$9.8k\Omega$	$319fF$	$5.8mH$	$11k\Omega$
Standard Deviation	$182fF$	$147fF$	$0.378mH$	$4.9k\Omega$	$60.8fF$	$1.1mH$	$2.97k\Omega$

has to be estimated carefully. In order to achieve a precise evaluation, the capacitance measured on a chip without the transducers was subtracted from the one found with the fitting process (Table 5.1 already shows the de-embedded capacitance). This way, the capacitance of the PCB traces and the on-chip routing and pads is not accounted for.
In addition to the two dominant resonance modes, there are a few other peaks, clearly visible in the phase plot, at $8.2MHz$, $15MHz$, $18.4MHz$, $21.3MHz$ and $24.6MHz$. These high frequency resonance modes were anticipated, yet they were not modeled in the design process, as no information about them was available.

5.2.2. Measurements in Gel

As discussed in Section 1.2, the electrical impedance of the ultrasound transducer also depends on its acoustic environment. Thus, the measurements proposed in the previous Section were repeated after placing a drop of gel on top of the acoustic stack. The measured impedance is shown in Figure 5.3.
The first resonance is attenuated to a point where it is hardly visible, while the second

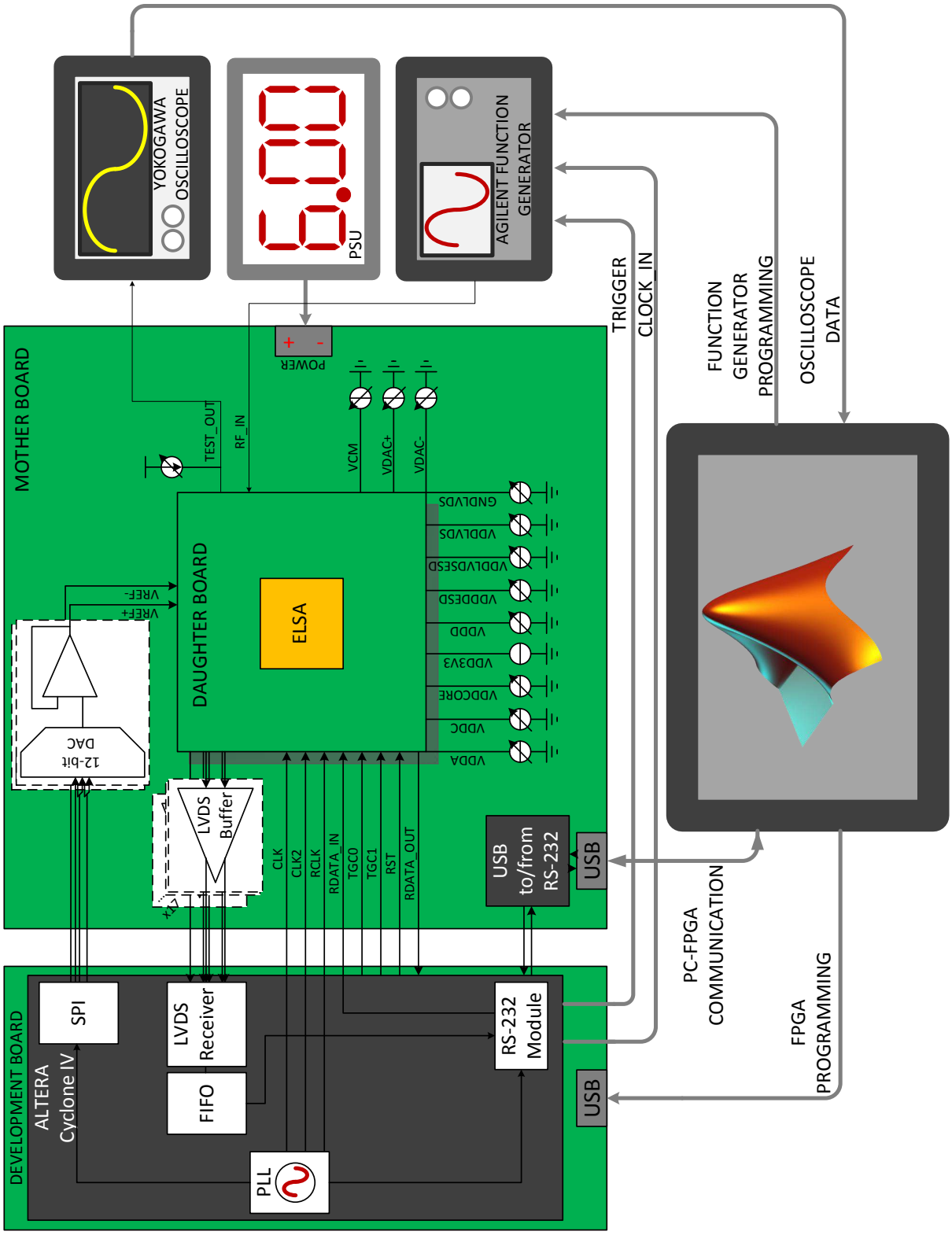


Figure 5.1: Block diagram of the implemented measurement set-up.

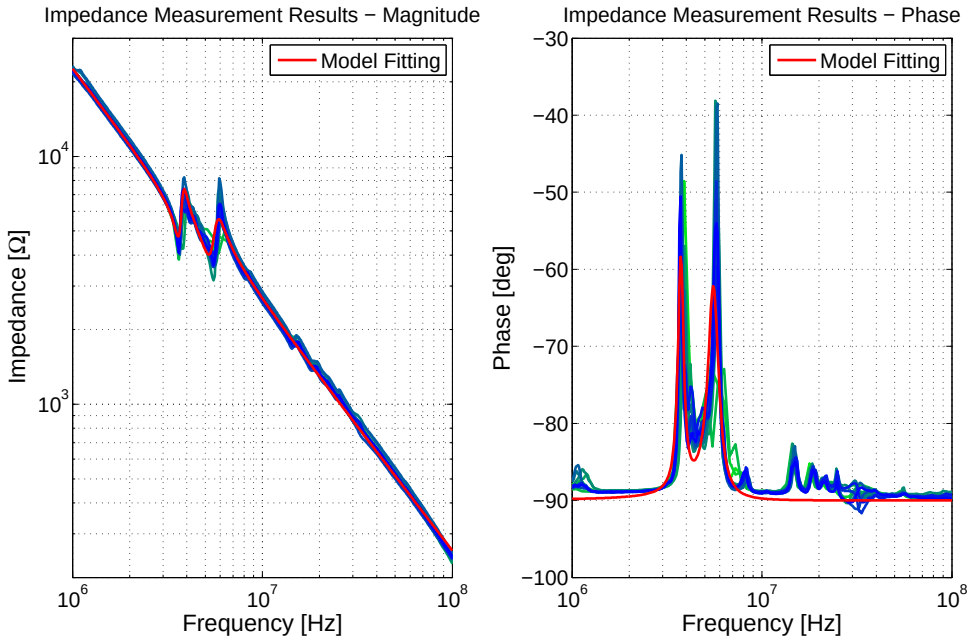


Figure 5.2: Measurement of transducer impedance through an impedance analyzer.

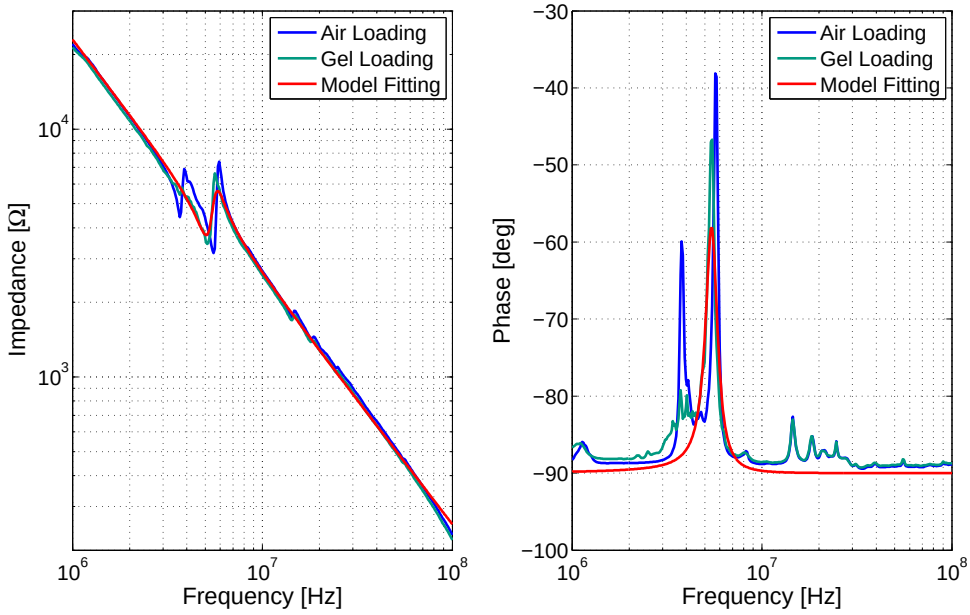


Figure 5.3: Measurement of transducer impedance with and without applying gel.

one slightly shifts to a lower frequency ($5.35MHz$) and increases its bandwidth. Thus, the first resonance is thought to be caused by the acoustic impedance mismatch between the transducer and air. In this case, the single-resonance Butterworth-Van Dyke model (shown in Figure 1.4) can still be used. The impedance parameters obtained by the fitting are listed in Table 5.2, where C_s has already been de-embedded. Luckily, these values closely resemble the ones for which the system was designed, and are actually slightly better (lower average R and C_s).

Table 5.2: Extracted electrical transducer model parameters, measured in acoustic gel (2 measured transducers).

	C_s	C	L	R
Transd. 1	$1.525pF$	$718fF$	$1.25mH$	$10.4k\Omega$
Transd. 2	$1.525pF$	$563fF$	$1.56mH$	$5.9k\Omega$
Expected Parameters	$2pF$	$731fF$	$1.6mH$	$9.44k\Omega$

5.3. Electrical Measurements without Transducers

In this Section, the measurement results obtained from the electrical samples (those that did not go through the transducer integration process) will be discussed. Of course, without transducers, the modulator's loop filter is missing, thus no noise shaping will be visible. However, some information about the chip functionalities can still be gathered. These measurements were performed by disabling the modulator's feedback, by forcing the voltages V_{REF+} and V_{REF-} to the common-mode voltage of $0.9V$. In this "open-loop" configuration, the system will act like a more conventional ultrasound front-end, composed of a TIA, a VGA and a 3-bit ADC.

5.3.1. Open-Loop Spectra

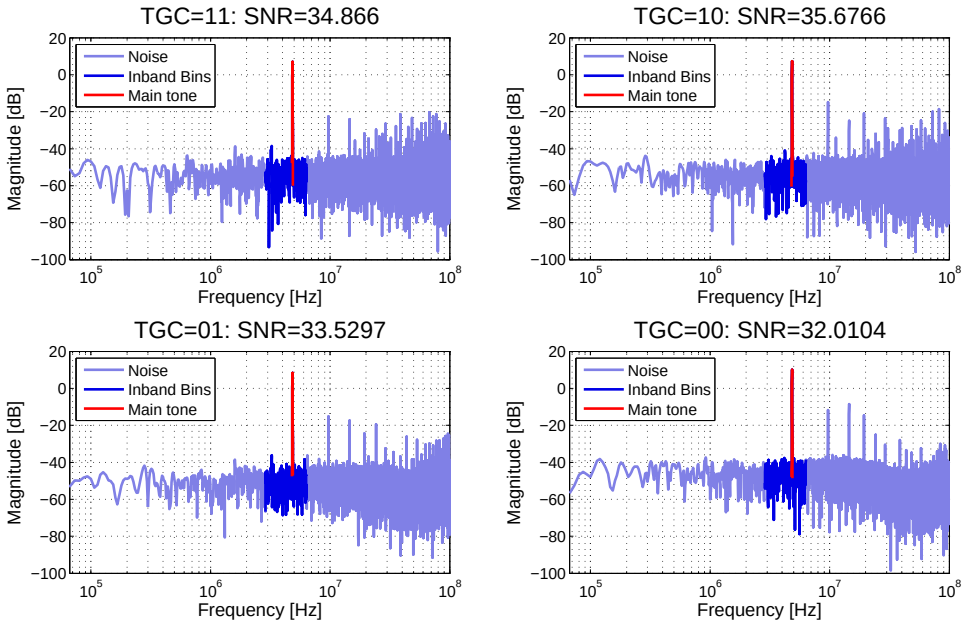


Figure 5.4: Spectra of the open-loop structure's output for all TGC configurations.

The expected SNR of such a system is roughly $34dB$ (see Equation 2.8). Indeed, measurements at all TGC configurations display a similar performance, as shown in Figure 5.4. As expected, the spectra are mostly flat, except for the distortion peaks. This behavior can be easily explained: the linearity is now defined by the quantizer DAC (see Section 3.3.3), which has a unit capacitor of only $9fF$. Furthermore, the comparators' noise and offset are dependent on the input voltage (see Section 3.3.1), adding further non-linearity to the system.

5.3.2. Transfer Functions & TGC Measurement

Figure 5.5 shows the measured transimpedance of the first stage. In this measurement, the frequency response of the on-chip buffer and of the function generator were de-embedded. Theoretically, the transimpedance should be equal to R_F at low frequencies and have a pole

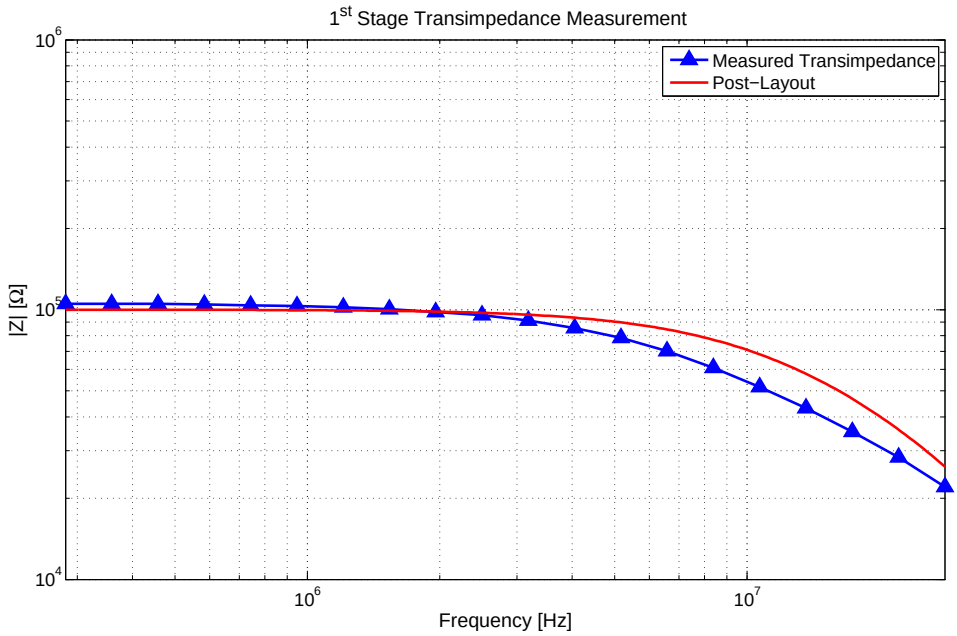


Figure 5.5: First stage transimpedance measurement, compared to post-layout simulations.

at $\frac{1}{2\pi R_F C_F}$ and another after f_s (see Figure 3.4). However, the additional loading of the buffer brings the second pole to lower frequencies, such that its influence is already visible at 20MHz, the maximum frequency that the function generator can provide. The first pole is also at lower frequencies, with respect to post-layout simulations, possibly due to a parasitic feedback capacitance larger than expected.

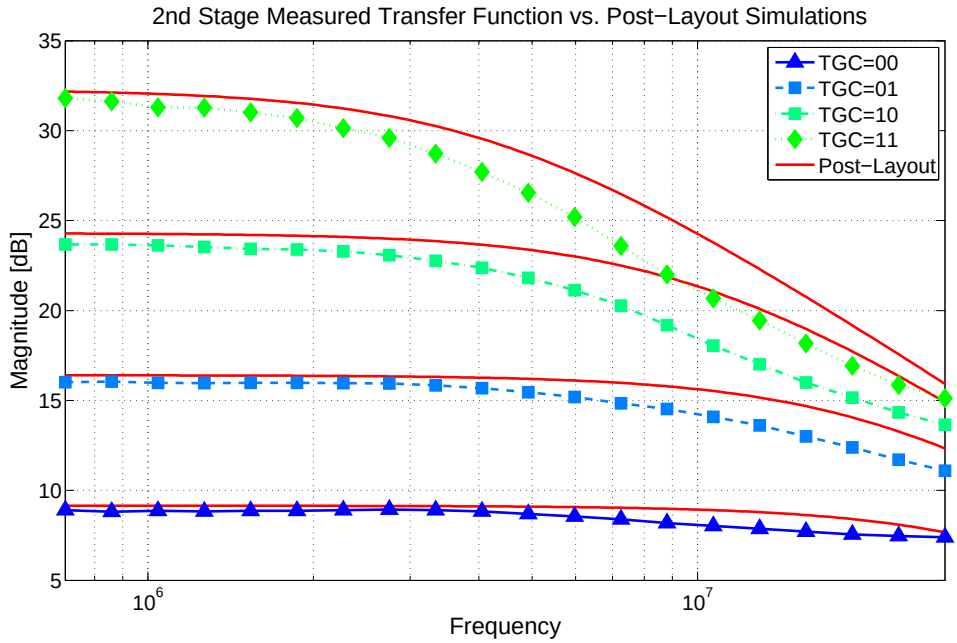


Figure 5.6: Second stage measured transfer function, compared to post-layout simulations.

The ability of the system to adapt to different input signal levels was also tested. Figure 5.6

shows the measured transfer functions for the second stage for all the TGC configurations. The proposed measurements were achieved by using the on-chip buffer and evaluating the amplitude of the second stage output divided by that of the first stage. The post-layout simulations presented were performed while loading the second stage with the estimated input capacitance of the on-chip buffer and the associated routing parasitics. Even accounting for the extra loading, the measured bandwidth is slightly lower than expected, however, the TGC low-frequency gains that were decided in the design phase (see Section 3.2.2) are very close to the measured ones. The measured transfer functions appear to have a slope of less than the expected $20\frac{dB}{dec}$ after $10MHz$. This could be caused by the fact that while measuring the first stage, its second pole is shifted towards low frequencies by the buffer loading, while, when the second stage output is measured, the first stage is only loaded by the second one.

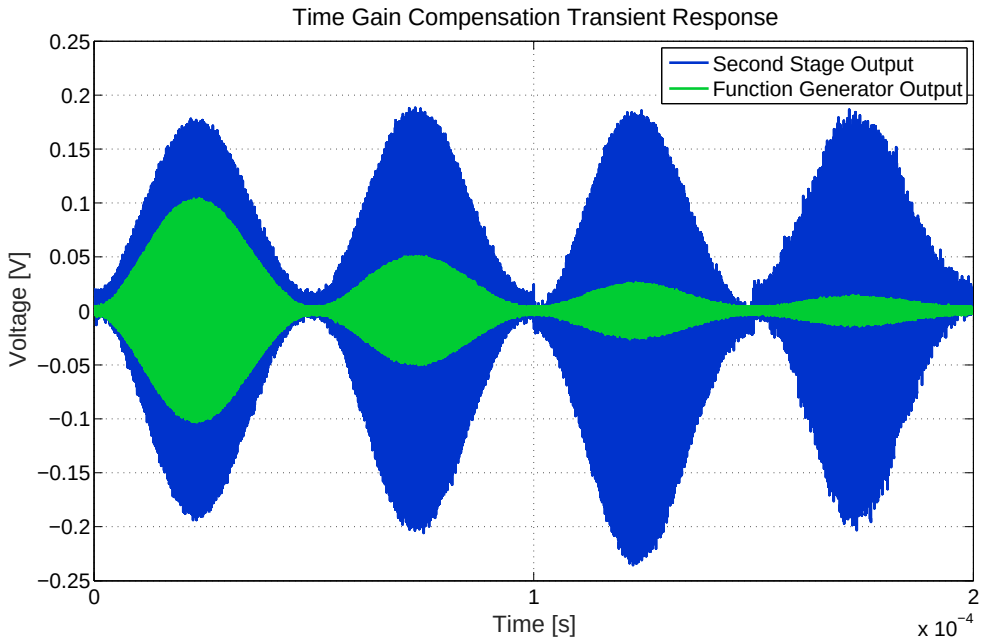


Figure 5.7: Measured time-domain trace showing the TGC functionality.

Figure 5.7 shows a transient simulation of the full receive phase ($200\mu s$). The input applied to the pixel-level test capacitor, shown in green, is a $5MHz$ sine wave with a decaying amplitude and a sinusoidal envelope. The envelope was added to make sure that at the transition points between different TGC configurations, no input is applied. As a matter of fact, if this condition is violated, the biasing point of the second stage could be jeopardized. In order for this condition to be valid, it was assumed that the ultrasound echoes are sparse enough in the time domain, so that the probability of receiving a reflection during the TGC switching point is low. If this assumption is observed to be invalid, a brief reset phase during the TGC transition can be added, keeping the second stage common-mode voltage under control.

The expected second stage output amplitude is $V_{out} = V_{in} 2\pi f_{in} C_{test} \cdot A_{2^{nd}stage}$. The impedance of the test capacitor and of R_F were designed to be almost identical at $5MHz$, so the first stage acts as a voltage buffer with respect to the test input. The blue curve shows how the amplitude at the output of the second stage is kept roughly constant by the TGC functionality.

5.4. Electrical Measurements with Transducers

In this Section, the converter performance is evaluated using the electrical test input, as in the previous section. However, this time, the employed sample has transducer integrated on top, and the modulator feedback is applied.

5.4.1. Closed-Loop Spectra

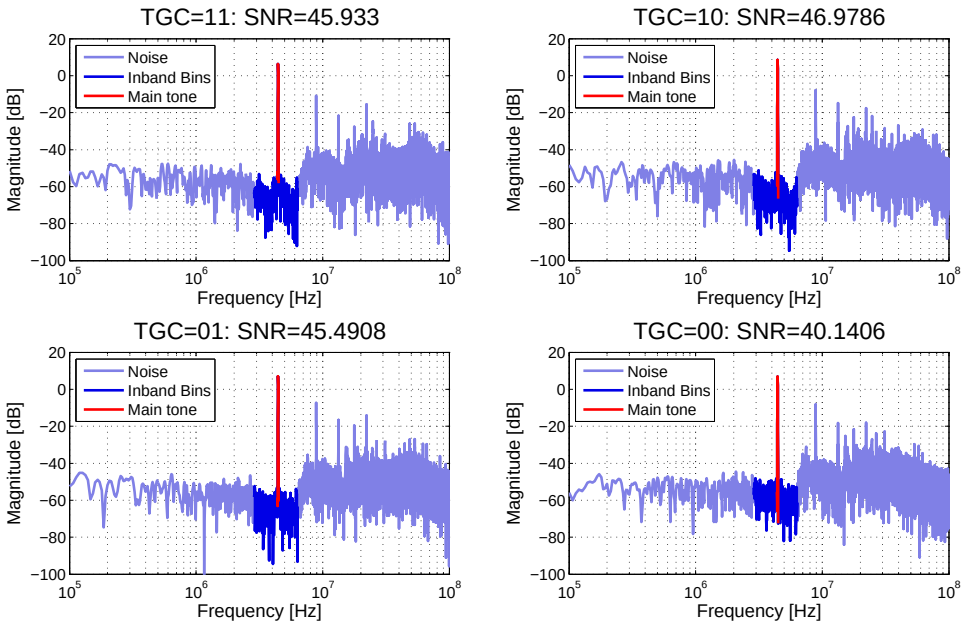


Figure 5.8: Spectra of the reconstructed modulator's output for all TGC configurations.

Figure 5.8 shows the measured modulator spectra for all TGC configurations, with a full-scale input. The desired noise shaping behavior can be clearly observed in all cases. The second and third order distortion peaks are quite high, at around $-20dBFS$ and $-30dBFS$ respectively, however this proves only slightly worse with respect the simulation shown in Figure 3.19, where mismatch is not accounted for.
The SNR is very close to the simulated one, or even exceeding the expectations, except in the highest gain configuration (TGC=00). Here, a SNR of only $40dB - 41dB$ is reached. A possible reason for this performance degradation is the noise coming from the reference voltages V_{REF+} and V_{REF-}, which enters the modulator at a very critical spot. This noise is the combination of that of the DAC that generates the reference voltages and the OPAMP that buffers them, for a total noise of $234\mu V_{rms}$, integrated in the bandwidth of interest. A very rough estimation of the equivalent integrated noise at the first stage's input yields a 25 times higher noise with respect to that of first stage, accounting for the different transfer functions. This could be an overestimation due to the limited information that can be found in the DAC data-sheet.

Let us take a closer look at the measured spectrum shape: in Figure 5.9 the modulator output's FFT is shown together with the measured phase of the transducer impedance (see Section 5.2). The correlation between the two plots should be evident. In particular, the bandwidth of interest presents two peaks instead of the single resonance that was assumed during the design phase. This slightly increases the usable bandwidth, but it could harm the SNR, as the central part between the two peaks will show a high quantization noise.

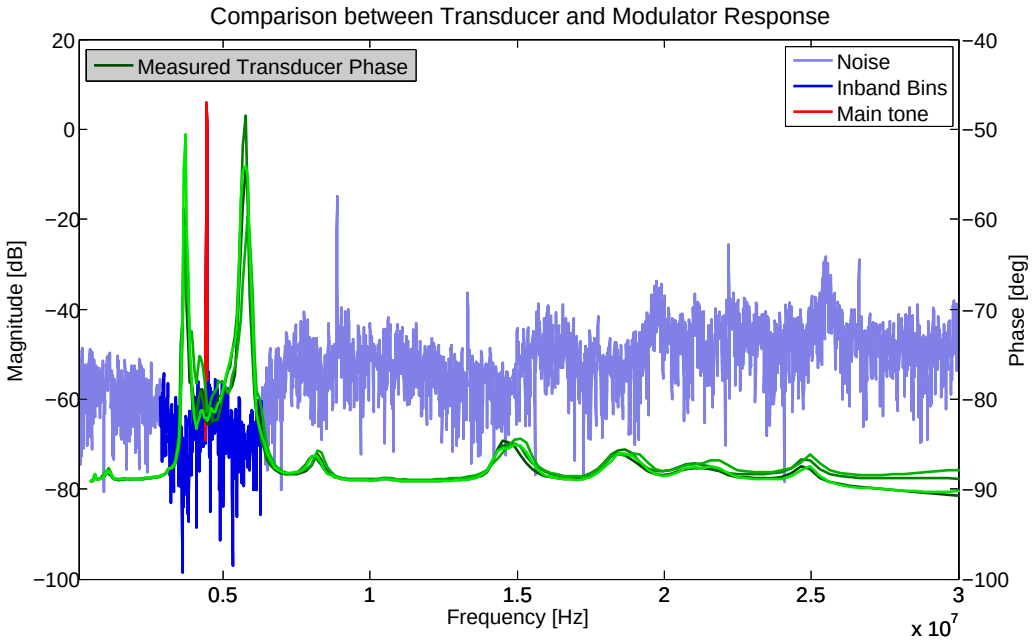


Figure 5.9: Spectrum of the modulator output bitstream, combined with a phase plot of that sample's measured transducers.

Moreover, the transducer's high frequency peaks are also visible in the modulator output. Finally, the downward slope of the spectrum between $10MHz$ and $15MHz$ could be caused by a feedback capacitance larger than expected in the first stage.

5.4.2. Dynamic Range Measurement

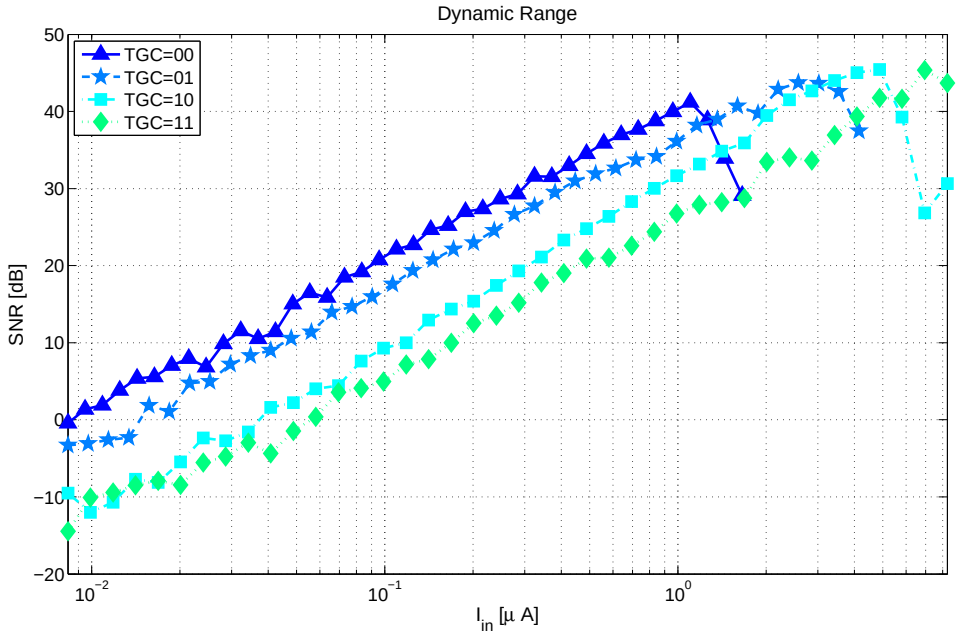


Figure 5.10: Measured Dynamic Range for different TGC configurations, $f_{in} = 4.4MHz$.

Figure 5.10 shows the measured dynamic range for all TGC configurations. The quantity

on the x axis is the first stage's input current, calculated as $I_{in} = 2\pi f_{in} C_{test} V_{in}$, where V_{in} is the voltage provided by the function generator. In high gain mode ($TGC = 00$) the highest point is reached at $I_{in}^{max} = 1.2\mu A$, which is similar to the predicted value: $I_{in}^{max} = \frac{A_{in}^{max}}{R} \approx 1.5\mu A$, where A_{in}^{max} is the maximum amplitude of the transducer model's voltage source, derived in Section 2.5.5 and R is the resistor. Because of the limited range of the function generator, the point where the DR plot flattens cannot be measured for the high gain configuration. However, one can conclude that the dynamic range, accounting for the time gain compensation, exceeds 3 decades ($60dB$).

5.4.3. Power Consumption

Table 5.3: Comparison between simulated current consumption and measured one

	VDDA				VDDC	VDDCORE	VDDLVDS	VDDLVDSESD
ITRIM	00	01	10	11	-	-	-	-
Simulations (μA)	247	275	303	331	93	108	1800	160
Measurements (μA)	190	212	233	255	92	138	921	140

Table 5.3 shows a comparison between the expected current consumption and the measured one. The most important difference is the analog current consumption, which shows a reduction of almost 30% with respect to simulation. The function of the mentioned power domains is explained in Section 4.5. The reason for this behavior is still unclear, and it could be the cause of the bandwidth reduction observed for both the stages. Trimming the first stage current leads to a variation of $22\mu A$ in the analog current consumption, as opposed to a simulated value of $28\mu A$. This points towards a problem in the biasing (unit current lower than anticipated). Moreover, this difference could be also blamed on the indirect measurement that was performed (total analog current divided by the number of elements).
The core digital current (from VDDCORE) is larger than expected, possibly due to an increase in the modulator's activity factor, which was observed during the measurements.
The LVDS pre-drivers show a similar current, drawn by the VDDLVDSESD pin, with respect to the simulated one. Finally, the 50% difference in current consumption for the LVDS drivers can be attributed to a load capacitance lower than the expected one.
The total measured per-channel power consumption goes from $756\mu W$ to $873\mu W$ depending on the selected first stage current, without including the LVDS transmitters.

5.5. Acoustic Measurements

In this Section, the results obtained with an acoustic input will be presented. A few additions (see Figure 5.11) to the previously described experimental set-up are required in order to estimate the modulator's response to an acoustic signal. A "water bag" is laid over the acoustic stack, which mimics the acoustic impedance of the human body. Additionally, an unfocused (with no intrinsic directionality) transmit transducer with a center frequency of $5MHz$ is placed inside the water bag. This transducer can either be connected to the function generator, which has a maximum amplitude of $10V_{pp}$, or to a dedicated high-voltage bench-top pulser. The position of the transmitter relative to the ASIC can be precisely adjusted through an X-Y-Z stage; furthermore the transmitter can be tilted along one of the directions.

5.5.1. Closed-Loop Spectra with an Acoustic Input

Figure 5.12 shows the FFTs of the modulator output for each TGC configuration. This time, the input is an acoustic wave sent by the transmitter, directly above the chip. In this case, the transmitter is connected to the function generator and is driven by a sine wave. These spectra

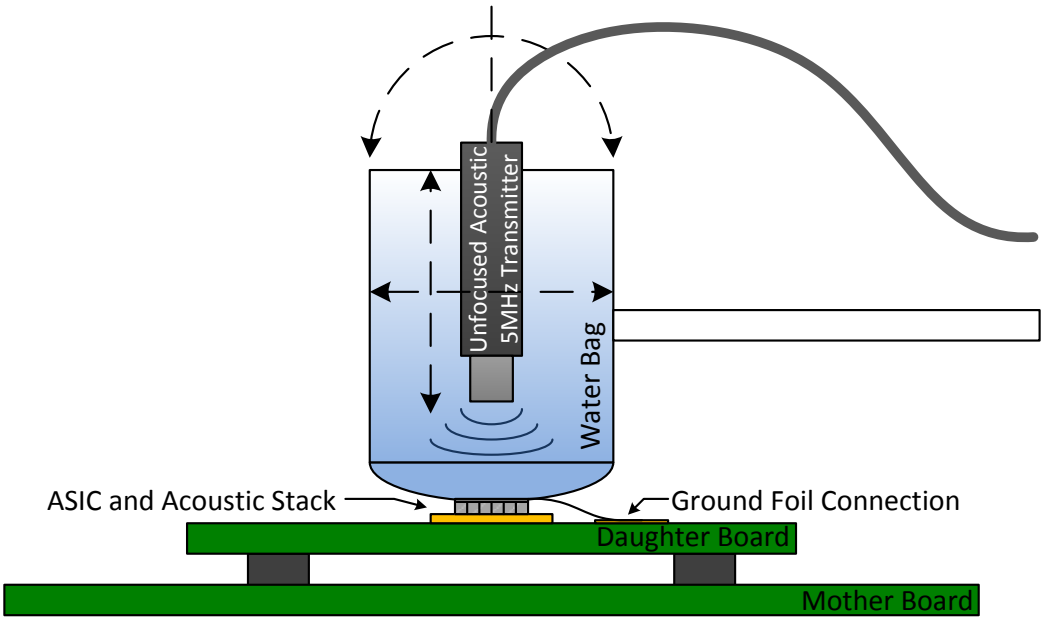


Figure 5.11: Block diagram of the set-up that allows for acoustic measurements.

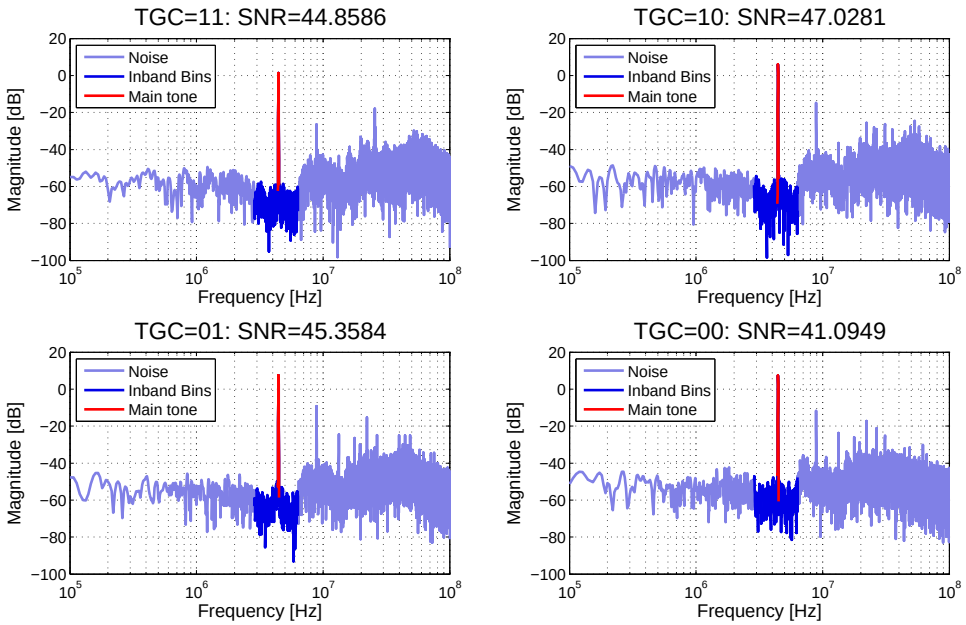


Figure 5.12: Spectra obtained at different TGC configurations with an acoustic input.

are practically identical to those shown in Figure 5.8, and so are the maximum signal to noise ratios. This might sound counter-intuitive in view of the discussion introduced in Section 5.2 about the different impedance that should be expected from the transducer in case of water loading. However, the water bag solution used for these measurements might induce a larger acoustic impedance mismatch between the transducer matching layer and the surrounding environment, with respect to the measurements in gel shown in Figure 5.3.

5.5.2. Transducer and Modulator Transfer Functions

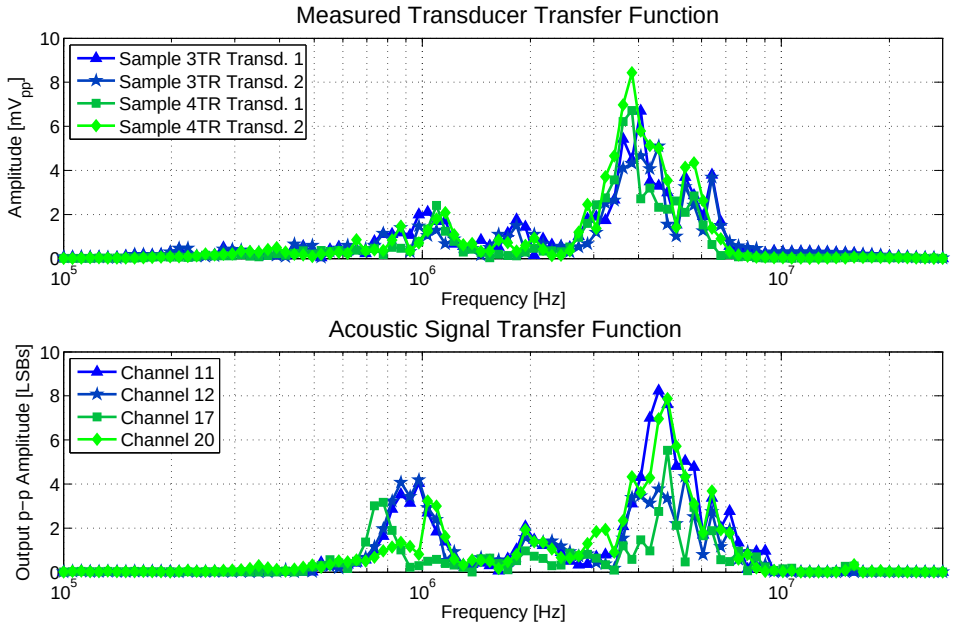


Figure 5.13: Measured modulator and transducer frequency response with an acoustic input.

Figure 5.13 shows a comparison between the frequency response of the modulator's digital output and that of the transducer. The measured elements are not the same in the two subplots: for the upper plot, the wired-out transducers were employed (see Figure 4.9). This measurement has been performed by sweeping the transmit frequency of the function generator, which was connected to the acoustic transmitter, while the amplitude was fixed at $10V_{pp}$. The transmitter was designed to resonate at around $5MHz$, thus, the shown results also include the bandwidth limitations introduced by it. This measurement can be seen as an evaluation of the acoustic STF of the modulator, which should not limit the intrinsic acoustic signal bandwidth. The general shape of the measured frequency response is similar: the main peak appears around $4MHz$ for the transducers and $4.5MHz$ for the modulators. This discrepancy could be caused by the different transducer loading condition in the two measurements, in particular, the wired out transducers will have a much higher load capacitance to drive. There seems to be a second peak at $1MHz$, which appears to be more relevant in the bottom sub-plot.

5.5.3. Time-Domain Acoustic Measurements

Now that the spectral characteristics of the modulator in the presence of an acoustic input were discussed, some time-domain measurements will be provided.
First of all, it is important to evaluate if the modulator output is a good approximation of the actual signal that the transducer provides. In order to verify that, the response of the modulator to an acoustic pulse has been compared in Figure 5.14 to the same pixel's TIA output, after having disabled the modulator feedback, in the same way as in Section 5.3. In this case, the TIA provides an amplified version of the transducer's output voltage. The two responses show a clear correlation, however they are not perfectly equivalent. A possible explanation for this effect has been provided in Section 2.6.2: when the modulator's feedback is applied, the termination condition for the transducer is changed.
In the measurements, the increase in the transducer's bandwidth that was anticipated in the

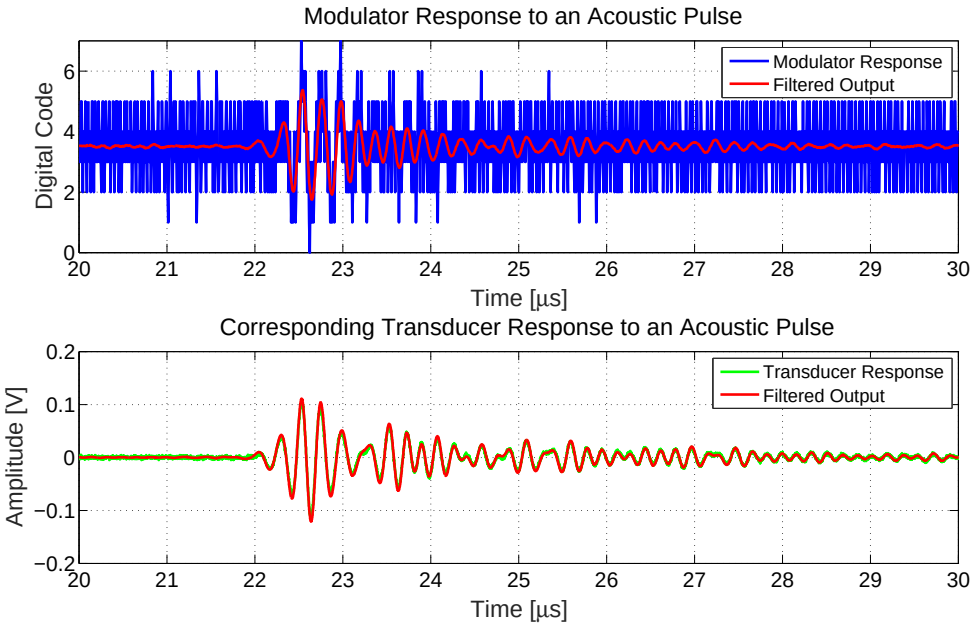


Figure 5.14: Comparison between the modulator output and the transducer response, for the same pixel.

system-level simulations was not observed, probably because of the limitations of the adopted Butterworth-Van Dyke model concerning the electro-mechanical conversion performed by the transducer.

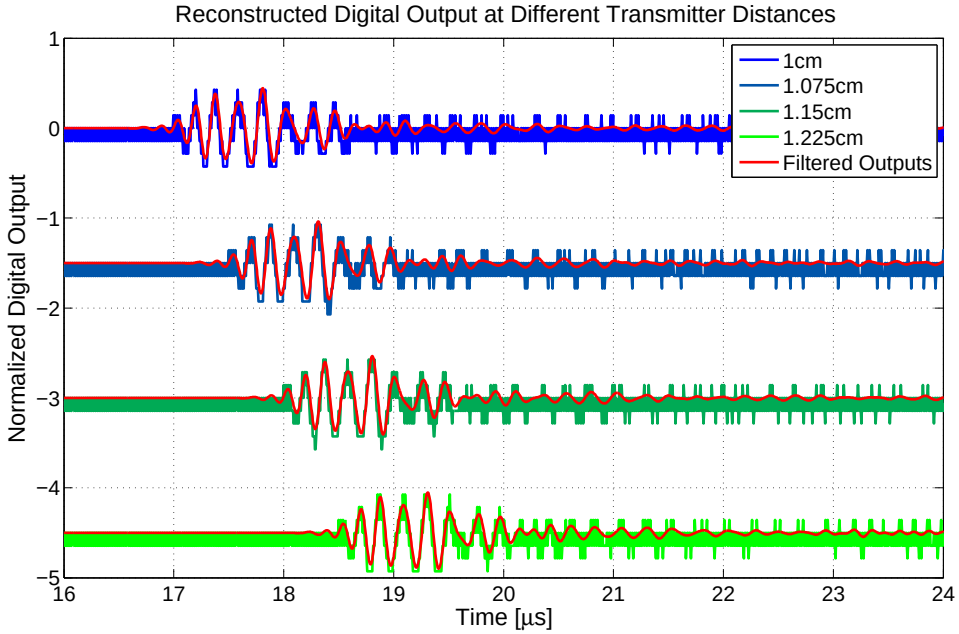


Figure 5.15: Modulator output with an acoustic pulse as an input, at different distances between the acoustic transmitter and the ASIC.

Figure 5.15 shows the response of the modulator to an acoustic pulse, while the transmitter position is varied. The starting distance d between the acoustic stack and the transmitter is

around $1cm$, which is increased by $0.75mm$ at each step. The observed delay (t_d) between two successive measurements is $500ns$, according to expectations: $t_d = \frac{d}{v} = \frac{0.75mm}{1500\frac{m}{s}} = 500ns$, where v is the speed of sound in water.
Another important result conveyed by Figure 5.15 is that the measured waveform keeps its shape constant, even with repeated measurements. This effect is more evident looking at the filtered outputs: yet another confirmation that, regardless of the modulator's chaotic behavior, the in-band quantization noise is low.

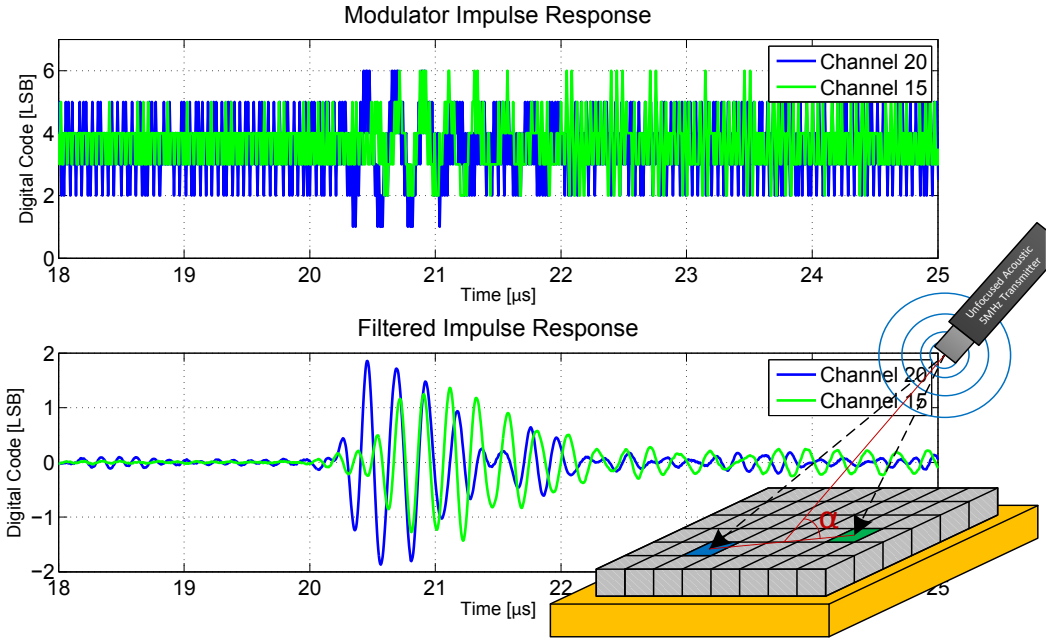


Figure 5.16: Arrival delay between two elements, applying an acoustic pulse from a tilted position

Finally, Figure 5.16 shows the response of two different elements to an acoustic pulse provided by the transmitter, which, this time, is placed in a tilted position. Assuming that the acoustic wave propagates through spherical wave-fronts, the element highlighted in green will be reached sooner by the wave with respect to the blue one. Indeed, the measurements show a different time of arrival for the two pixels. The time difference is approximately equal to one period ($200ns$), which corresponds to a mean path length difference of $300\mu m$, while the physical distance between the two elements along the tilting axis is $600\mu m$. Of course, the estimated and the physical distance can be different, and their difference is a function of the tilting angle α, which in this case was approximately 60^o, and the distance between the center of the array and the transmitter ($2cm$). Only with an angle of 0^o with respect to the chip surface one would measure the actual physical distance between the two elements, but this proves to be impractical.

5.6. Performance Comparison

This work uses the ultrasound transducer as the loop filter for a $BP\Sigma\Delta$ converter, thus, the area typically reserved for the reactive components needed to implement standard loop filters is saved. This allows the designed converter to be very competitive in terms of area.
Figure 5.17 shows a performance comparison between this work and other $\Sigma\Delta$ converters published in two of the most important circuit conferences. The green markers show designs that

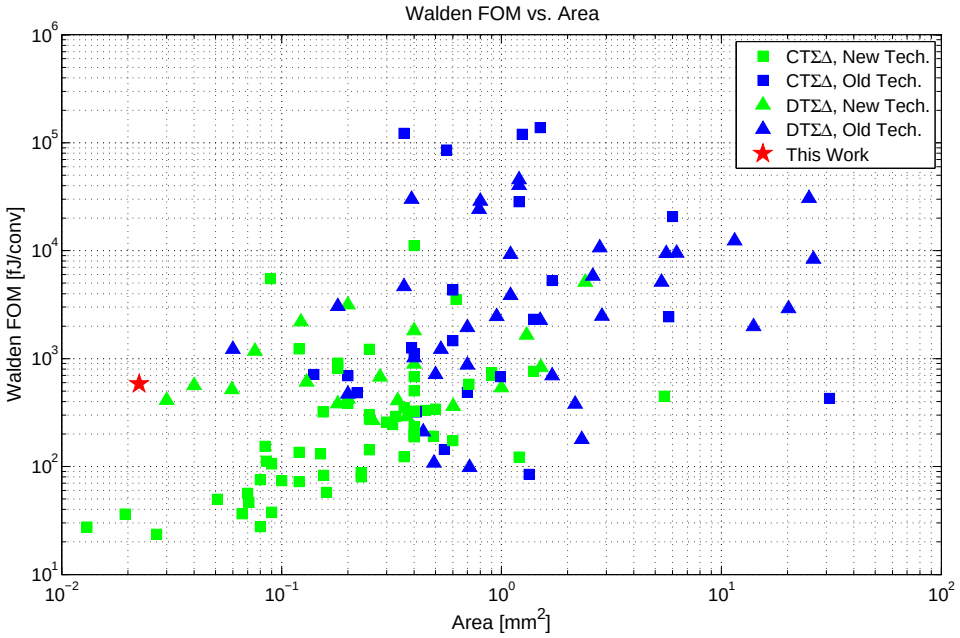


Figure 5.17: Walden Figure of Merit vs. area of modulators published in ISSCC and VLSI [4].

use a feature size smaller than $0.18\mu m$, while the blue ones represent modulators employing a $0.18\mu m$ technology or an older one. The x-axis represents the area, while on the y-axis, the Walden figure of merit is shown, defined as:

$$FOM_W = \frac{Power}{\frac{f_s}{OSR}2^{\frac{SNR-1.76}{6.02}}} = \frac{800\mu W}{\frac{200MHz}{26.7}2^{\frac{47dB-1.76}{6.02}}} = 0.583\frac{pJ}{conv} \tag{5.1}$$

A general trend can be identified, which associates a lower active area with a smaller FOM. This work achieves by far the lowest area, compared to designs using a similar technology. Furthermore, only two of the converters employing a smaller feature size achieve a lower area. The Walden FOM, while being more than one decade above the state of the art, can still be considered as a competitive result. This is because, in the current implementation, 50% of the power goes to the digital core and the comparators, which would benefit greatly from a smaller feature size.

Finally, Figure 5.18 shows a direct comparison between area and power consumption, for modulators with similar SNR and bandwidth, using any technology node. This work features the lowest area, as well as the smallest power consumption among all the designs.

5.7. Conclusion

In this Chapter, the implemented test set-up was described and the most relevant measurements were presented and discussed. The electrical and acoustic results are, for the most part, in line with the expectations, and demonstrate not only that the concept is viable, but also that the predicted performance can be achieved without significant roadblocks.
Finally, this chip's measured performance was compared to that of similar works.

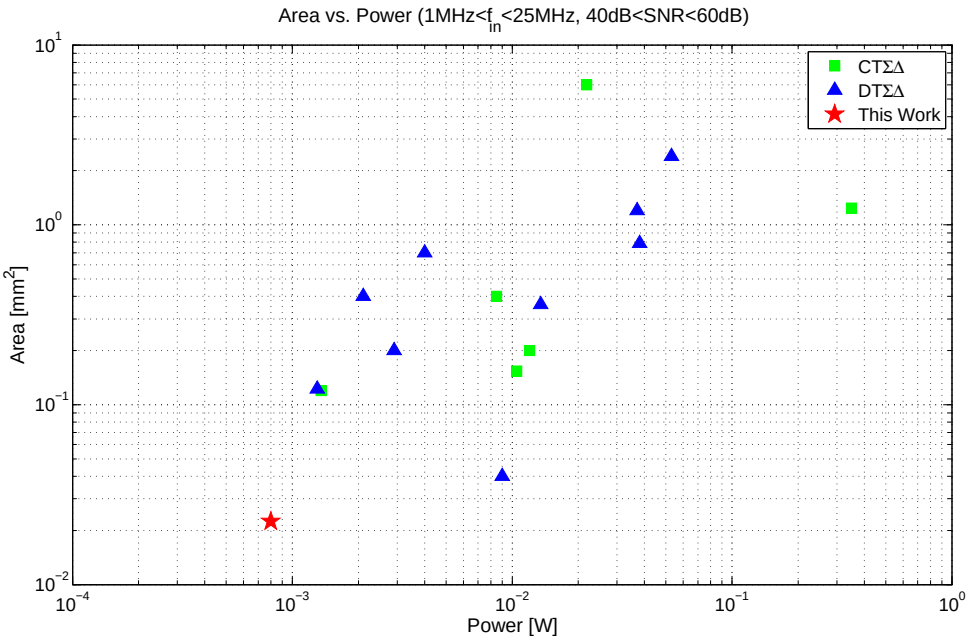


Figure 5.18: Comparison of this work with respect to others in the same SNR and BW range [4].

6 Conclusions

6.1. Thesis Contribution

The most important result for this thesis is that using the frequency response of the sensor as the loop filter of a $\Sigma\Delta$ modulator can inherently save complexity, area an power. This concept can be theoretically extended to any "flavor" of ultrasound imaging, and is not limited to TEE probes.
Moreover, this idea, although not entirely new (see Section 1.5.1), can be applied to any field that requires a sensor with a known and exploitable transfer function. As a matter of fact, sensors showing a capacitive behavior in the bandwidth of interest (see Section 2.2) would be the preferred candidates for the application of this concept, since the capacitance compensation techniques used in this design would not be required.

Another interesting aspect of this work is that it does not introduce much additional circuitry with respect to a more traditional analog front-end. The TIA and the second stage (which acts as a VGA) are used in the majority of ultrasound readout systems, while the only extra blocks needed to convert an AFE into a complete ADC are the tunable RC branch, the feedback DACs and the tracking quantizer. These circuits occupy only 37% of the area and consume roughly 50% of the power.

The implemented tracking ADC, aside from the problems with the comparators (see Section 3.3.1), represents a power- and area-efficient solution for multi-bit quantizers to be used in a $\Sigma\Delta$ loop. Its linearity does not directly influence that of the modulator. The number of comparators is reduced to two, regardless of the required number of bits, furthermore, the full conversion is completed within one clock cycle, as opposed to SAR converters. The associated digital logic can be kept rather simple, allowing for the use of such a quantizer even for relatively high speed applications.

Finally, the Butterworth-Van Dyke model for ultrasound transducers was validated once more, proving that the purely electrical response of such a transducer can be described by an RLC branch in parallel with a capacitance. Of course, more sophisticated models would have been able to predict other features of the transducer, such as higher order resonances and the precise conversion coefficients between the mechanical and the electrical domain. These models could have been used to predict the acoustic STF of the modulator, which is not well represented by the model that was adopted, as highlighted by the difference between the expected STF and the measured one.

6.2. Future Work

The main concepts that were put to a test through the design of this chip have been proven rather successful. Nonetheless, there are a few aspects of this work that need further attention, before it can be proposed as a commercial product.

First of all, a strategy to decimate and perform beam-forming, possibly combining these two functions, must be found. The need for decimation and beamforming is highlighted by the high power consumption required in the current implementation in order to stream out the 20 differential digital outputs at $200MHz$. Solving this issue is fundamental, especially if, like in most cases, a bigger array is required.

In order to reach a mature prototype, with a higher number of elements, a better strategy for the routing of the modulator outputs is required. A possible approach would be combining the pixels in sub-arrays and introduce in every sub-array a decimation filter, which can be partially shared between elements, and a multiplexer, in order to reduce the amount of outputs to be routed. Furthermore, the current implementation features two digital outputs per pixel. However, these two signals could be serialized or combined using tri-state circuits or trivalent logic.

In a second iteration, the additional circuitry needed to implement the acoustic transmitters could be implemented. This would eliminate the need for an external transmitter and increase the range of acoustic experiments that can be performed.

The selected capacitance compensation method (see Section 2.4), although effective, is not very elegant, as it requires a calibration step and an additional bank of capacitors. Further research on this topic could uncover new, interesting techniques.

A better way to generate and distribute the reference voltages V_{REF+} and V_{REF-} should be investigated. As a matter of fact, the SNR reached in high gain configuration, in the current implementation, is limited by the noise on these lines. An on-chip reference generation, with an eye on noise, could solve the issue.

The implemented two-clock arrangement for the tracking quantizer is somewhat delicate and unreliable. An in-pixel generation of the second clock, in a process-independent manner, could benefit the system robustness.

The designed comparator clearly suffers from a data-dependent noise and offset, introducing many unpredictable non-idealities in the tracking quantizer. The use of a topology which includes a continuous pre-amplifier could help mitigate this effect.

Finally, a careful evaluation of the electrical and acoustic cross-talk between different elements should be performed.

Bibliography

[1] M. Heron, “National Vital Statistics Reports Volume 65, Number 5 June 30, 2016 Deaths: Leading Causes for 2014,” Tech. Rep. 5, NVSS, 2016.

[2] “Echocardiogram - CHOC Children’s,” in http://www.choc.org/heart/cardiodiagnostics/echocardiogram/.

[3] Y. Zhang, C.-H. Chen, T. He, and G. C. Temes, “A Continuous-Time Delta-Sigma Modulator for Biomedical Ultrasound Beamformer Using Digital ELD Compensation and FIR Feedback,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 62, no. 7, pp. 1689–1698, 2015.

[4] B. Murmann, “ADC Performance Survey 1997-2016,” 2016.

[5] M. O. Culjat, D. Goldenberg, P. Tewari, R. S. Singh, and L. Anthony, “A review of tissue substitutes for ultrasound imaging,” Ultrasound in medicine & biology, vol. 36, no. 6, pp. 861–73, 2010.

[6] Food and Drugs Administration, “Information for Manufacturers Seeking Marketing Clearance of Diagnostic Ultrasound Systems and Transducers,” 2008.

[7] W. P. Mason, “Physical Acoustics and the Properties of Solids,” The Journal of the Acoustical Society of America, vol. 28, no. 6, p. 1197, 1956.

[8] R. Krimholtz, D. Leedom, and G. Matthaei, “New equivalent circuits for elementary piezoelectric transducers,” Electronics Letters, vol. 6, no. 13, p. 398, 1970.

[9] S. Sherrit, S. Leary, B. Dolgin, and Y. Bar-Cohen, “Comparison of the Mason and KLM equivalent circuits for piezoelectric resonators in the thickness mode,” 1999 IEEE Ultrasonics Symposium. Proceedings. International Symposium (Cat. No.99CH37027), vol. 2, no. July, pp. 921–926, 1999.

[10] D. A. Berlincourt, D. R. Curran, and H. Jaffe, “Piezoelectric and Piezomagnetic Materials and Their Function in Transducers,” in Physical Acoustics, pp. 169–270, Academic Press, 1964.

[11] M. Guan and W.-H. Liao, “Studies on the circuit models of piezoelectric ceramics,” in International Conference on Information Acquisition, pp. 26–31, IEEE, 2004.

[12] Q. Liu, C. Chen, Z.-Y. Chang, C. Prins, and M. A. P. Pertijs, “A mixed-signal multiplexing system for cable-count reduction in ultrasound probes,” in 2015 IEEE International Ultrasonics Symposium (IUS), pp. 1–4, IEEE, 2015.

[13] P. L. Muntal, K. Faerch, I. H. Jorgensen, and E. Bruun, “System level design of a continuous-time $\Delta\Sigma$ modulator for portable ultrasound scanners,” 2015 Nordic Circuits and Systems Conference (NORCAS): NORCHIP & International Symposium on System-on-Chip (SoC), pp. 1–4, 2015.

[14] C. Chen, S. B. Raghunathan, Z. Yu, M. Shabanimotlagh, Z. Chen, Z.-y. Chang, S. Blaak, C. Prins, J. Ponte, E. Noothout, H. J. Vos, J. G. Bosch, M. D. Verweij, N. de Jong, and M. A. P. Pertijs, "A Prototype PZT Matrix Transducer With Low-Power Integrated Receive ASIC for 3-D Transesophageal Echocardiography," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 63, no. 1, pp. 47–59, 2016.

[15] S. Freeman, M. Quick, M. Morin, R. Anderson, C. Desilets, T. Linnenbrink, and M. O'Donnell, "Delta-sigma oversampled ultrasound beamformer with dynamic delays," IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 46, no. 2, pp. 320–332, 1999.

[16] O. Norman, "A Band-Pass Sigma-Delta Modulator for Ultrasound Imaging at 160MHz Clock Rate," IEEE Solid-State Circuits Conference, vol. 28, no. 5, pp. 15–17, 1993.

[17] P. Song, K. T. Tiew, Y. Lam, and L. M. Koh, "A CMOS 3.4 mW 200 MHz continuous-time delta-sigma modulator with 61.5 dB dynamic range and 5 MHz bandwidth for ultrasound application," Midwest Symposium on Circuits and Systems, pp. 152–155, 2007.

[18] T.-C. C. Cheng and T.-H. Tsai, "CMOS Ultrasonic Receiver With On-Chip Analog-to-Digital Front End for," IEEE Sensors Journal, vol. 16, no. 20, pp. 7454–7463, 2016.

[19] Y. S. Shu, J. Y. Tsai, P. Chen, T. Y. Lo, and P. C. Chiu, "A 28fJ/conv-step CT $\Delta\Sigma$ modulator with 78dB DR and 18MHz BW in 28nm CMOS using a highly digital multibit quantizer," Digest of Technical Papers - IEEE International Solid-State Circuits Conference, vol. 56, pp. 268–269, 2013.

[20] M. K. Chirala, P. Huynh, J. Ryu, and Y. H. Kim, "A 128-ch $\Sigma\Delta$ ADC based mixed signal IC for full digital beamforming Wireless handheld Ultrasound imaging system," Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, vol. 2015-Novem, no. Cic, pp. 1339–1342, 2015.

[21] J. Jeong, N. Collins, and M. P. Flynn, "A 260 MHz IF Sampling Bit-Stream Processing Digital Beamformer With an Integrated Array of Continuous-Time Band-Pass $\Delta\Sigma$ Modulators," IEEE Journal of Solid-State Circuits, pp. 1–9, 2016.

[22] P. Llimos Muntal, "A 10 MHz Bandwidth Continuous-Time Delta-Sigma Modulator for Portable Ultrasound Scanners," Proceedings of Norcas 2016, 2016.

[23] Y.-J. Kim, S.-E. Cho, J.-Y. Um, M.-K. Chae, J. Bang, J. Song, T. Jeon, B. Kim, J.-Y. Sim, and H.-J. Park, "A Single-Chip 64-Channel Ultrasound RX-Beamformer Including Analog Front-End and an LUT for Non-Uniform ADC-sample-clock Generation," IEEE Transactions on Biomedical Circuits and Systems, vol. 11, no. 1, pp. 87–97, 2017.

[24] M.-C. Chen, S.-R. Kothapalli, P. Cathelin, A. Cathelin, S. S. Gambhir, and B. Murmann, "A Pixel-Pitch-Matched Ultrasound Receiver for 3D Photoacoustic Imaging with Integrated Delta-Sigma Beamformer in 28nm UTBB FDSOI," in IEEE Solid-State Circuits Conference, (San Franciso, USA), pp. 456–458, 2017.

[25] A. Spineanu, P. Bénabès, and R. Kielbasa, "A digital piezoelectric accelerometer with sigma-delta servo technique," Sensors Actuators A Phys., vol. 60, no. 1-3, pp. 127–133, 1997.

[26] S. Sutula, J. Pallares Cuxart, J. Gonzalo-Ruiz, F. X. Munoz-Pascual, L. Teres, and F. Serra-Graells, “A 25-µW All-MOS potentiostatic delta-sigma ADC for smart electrochemical sensors,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 61, no. 3, pp. 671–679, 2014.

[27] K. A. A. Makinwa and J. H. Huijsing, “A Smart Wind Sensor using Time-Multiplexed Thermal ΣΔ Modulators Thermal ΣΔ modulation,” European Solid-State Circuits Conference, pp. 452–455, 2001.

[28] Y. P. Xu and R. Yu, “Electromechanical resonator based bandpass sigma-delta modulator for wireless transceivers,” Proceedings - 2005 IEEE International Workshop on Radio-Frequency Integration Technology: Integrated Circuits for Wideband Communication and Wireless Sensor Networks, RFIT 2005, pp. 101–104, 2005.

[29] R. Yu and Y. P. Xu, “A 47.3-MHz SAW resonator based CMOS second-order bandpass sigma-delta modulator with 54-dB peak SNDR,” Proceedings of the Custom Integrated Circuits Conference, no. 1, pp. 204–207, 2005.

[30] B. K. Thandri and J. Silva-Martinez, “A 63 dB SNR, 75-mW bandpass RF ΣΔ at 950 MHz using 3.8-GHz clock in 0.25-μm SiGe BiCMOS technology,” IEEE Journal of Solid-State Circuits, vol. 42, no. 2, pp. 269–279, 2007.

[31] R. Sobot, S. Stapleton, and M. Syrzycki, “Continuous time bandpass ΣΔ modulators for a software defined radio,” Canadian Conference on Electrical and Computer Engineering 2004 (IEEE Cat. No.04CH37513), vol. 3, pp. 1253–1256, 2004.

[32] R. Schreier and M. Snelgrove, “Bandpass Sigma Delta Modulation,” Electonics Letters, vol. 25, no. 23, pp. 1560–1561, 1989.

[33] F. Ortmanns, M. , Gerfers, Continuous-Time Sigma-Delta A/D Conversion. Springer, 2005.

[34] G. Wei, P. Shettigar, F. Su, X. Yu, and T. Kwan, “A 13-ENOB, 5 MHz BW, 3.16 mW multi-bit continuous-time ΔΣ ADC in 28 nm CMOS with excess-loop-delay compensation embedded in SAR quantizer,” in IEEE Symposium on VLSI Circuits, Digest of Technical Papers, pp. C292–C293, IEEE, 2015.

[35] F. Colodro and A. Torralba, “Continuous-time sigma - Delta modulator with a fast tracking quantizer and reduced number of comparators,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 57, no. 9, pp. 2413–2425, 2010.

[36] H. Park, K. Nam, D. K. Su, K. Vleugels, and B. A. Wooley, “A 0.7-V 100-dB 870μW digital audio ΣΔ modulator,” IEEE Symposium on VLSI Circuits, pp. 178–179, 2008.

[37] Y. H. Leow and L. Siek, “A high speed tracking quantizer for Continuous-Time multi-bit sigma delta modulators,” IEEE Asia-Pacific Conference on Circuits and Systems, Proceedings, APCCAS, pp. 776–779, 2010.

[38] L. Dörrer, F. Kuttner, P. Greco, P. Torta, and T. Hartig, “A 3-mW 74-dB SNR 2-MHz continuous-time delta-sigma ADC with a tracking ADC quantizer in 0.13-μm CMOS,” IEEE Journal of Solid-State Circuits, vol. 40, no. 12, pp. 2416–2426, 2005.

[39] R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters. Wiley Interscience, 2005.

[40] K. Matsukawa, Y. Mitani, M. Takayama, K. Obata, S. Dosho, and A. Matsuzawa, “A fifth-order continuous-time delta-sigma modulator with single-opamp resonator,” IEEE Journal of Solid-State Circuits, vol. 45, no. 4, pp. 697–706, 2010.

[41] M. Bolatkale, High Speed and Wide Bandwidth Delta-Sigma ADCs. PhD thesis, TU Delft, Delft, 2013.

[42] M. Bazes, “Two novel fully complementary self-biased CMOS differential amplifiers,” IEEE Journal of Solid-State Circuits, vol. 26, no. 2, pp. 165–168, 1991.

[43] B. Razavi, “The StrongARM latch [A Circuit for All Seasons],” IEEE Solid-State Circuits Magazine, vol. 7, no. 2, pp. 12–17, 2015.

[44] D. V. Duong, “A Capacitive Dynamic Comparator With Low Kickback Noise For Pipelined ADC,” IEEE International Conference on Electronics, Computing and Communication Technologies, vol. 2, no. 3, pp. 1–4, 2013.

[45] B. Razavi, Design of analog CMOS integrated circuits. McGraw-Hill, 2001.

A

Appendix

A.1. PCB Design

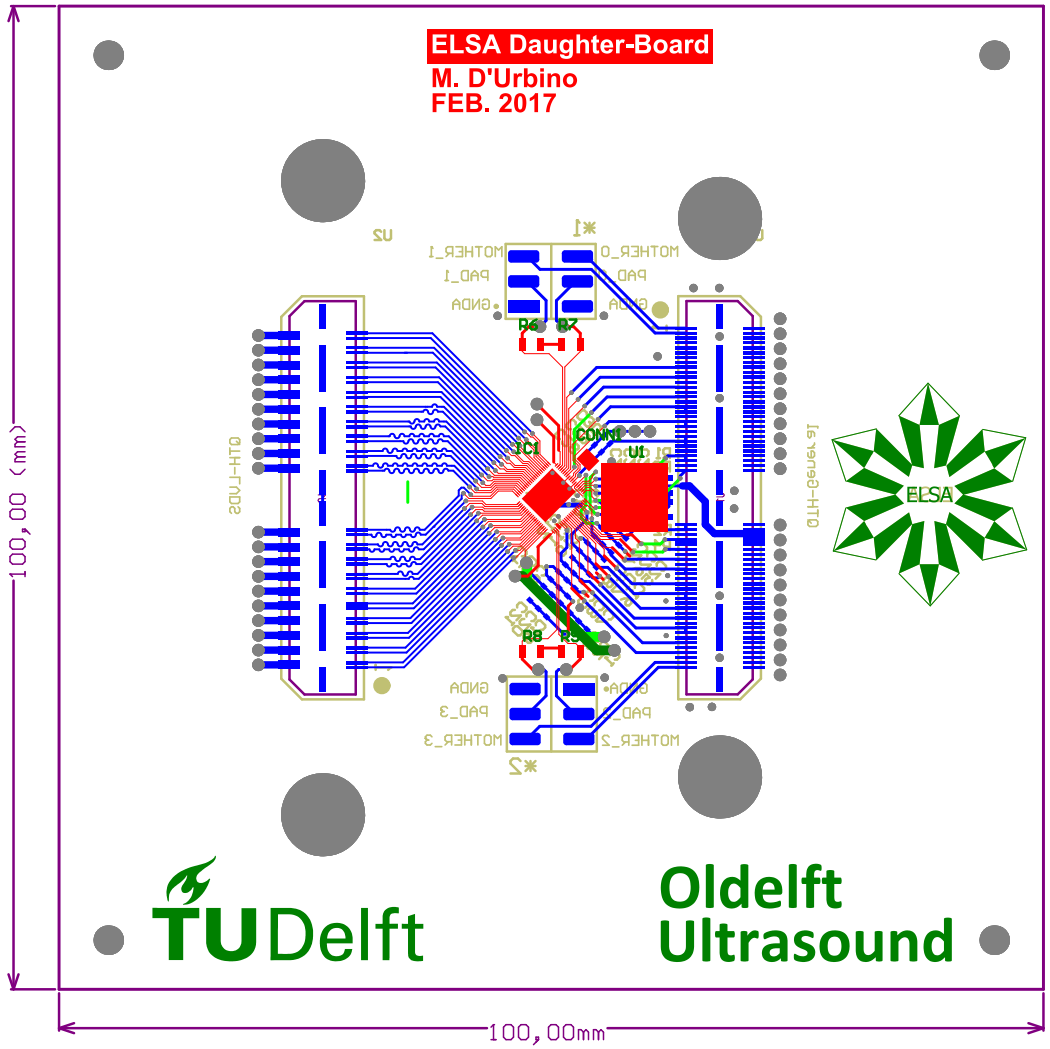


Figure A.1: Daughter-board PCB layout.

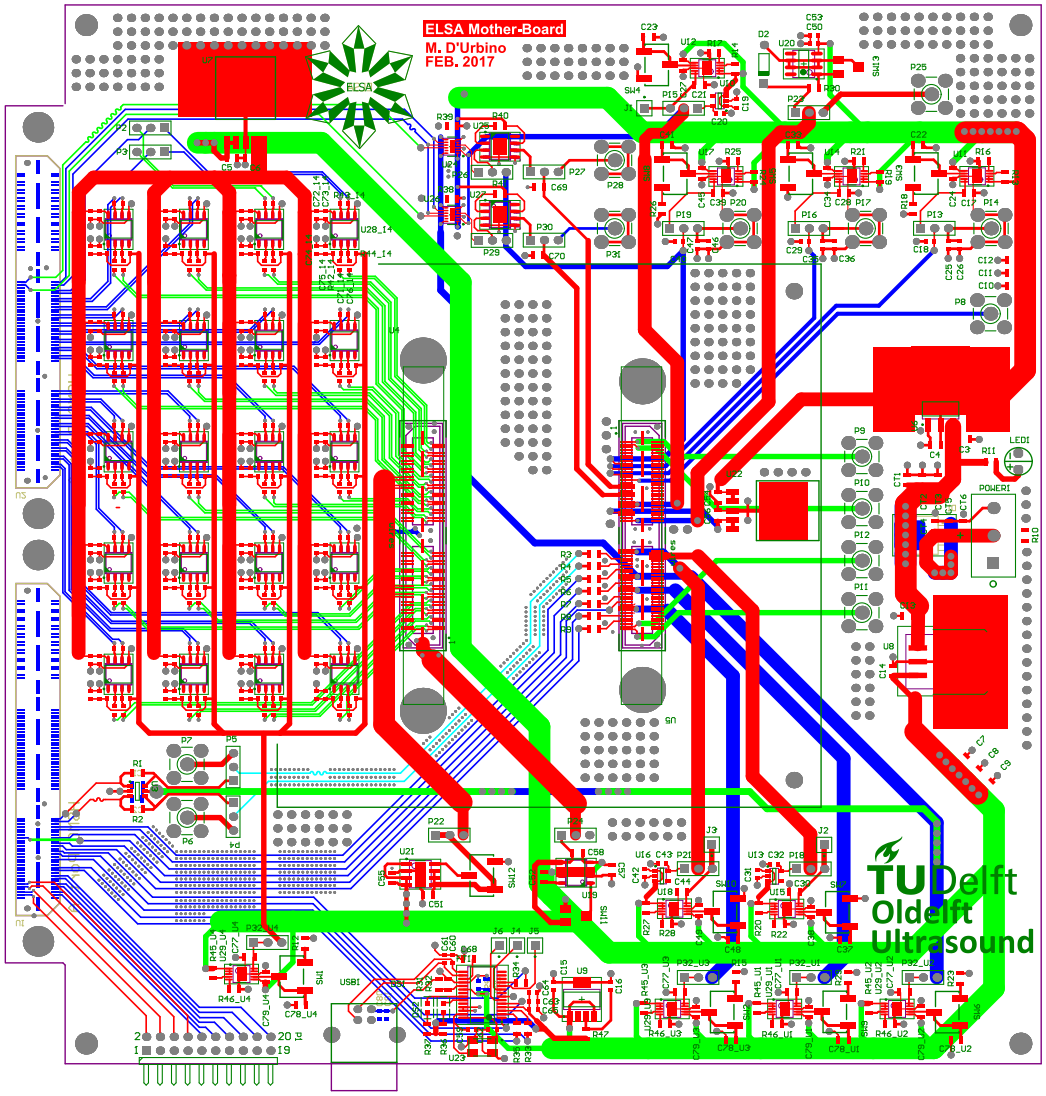


Figure A.2: Mother-board PCB layout.