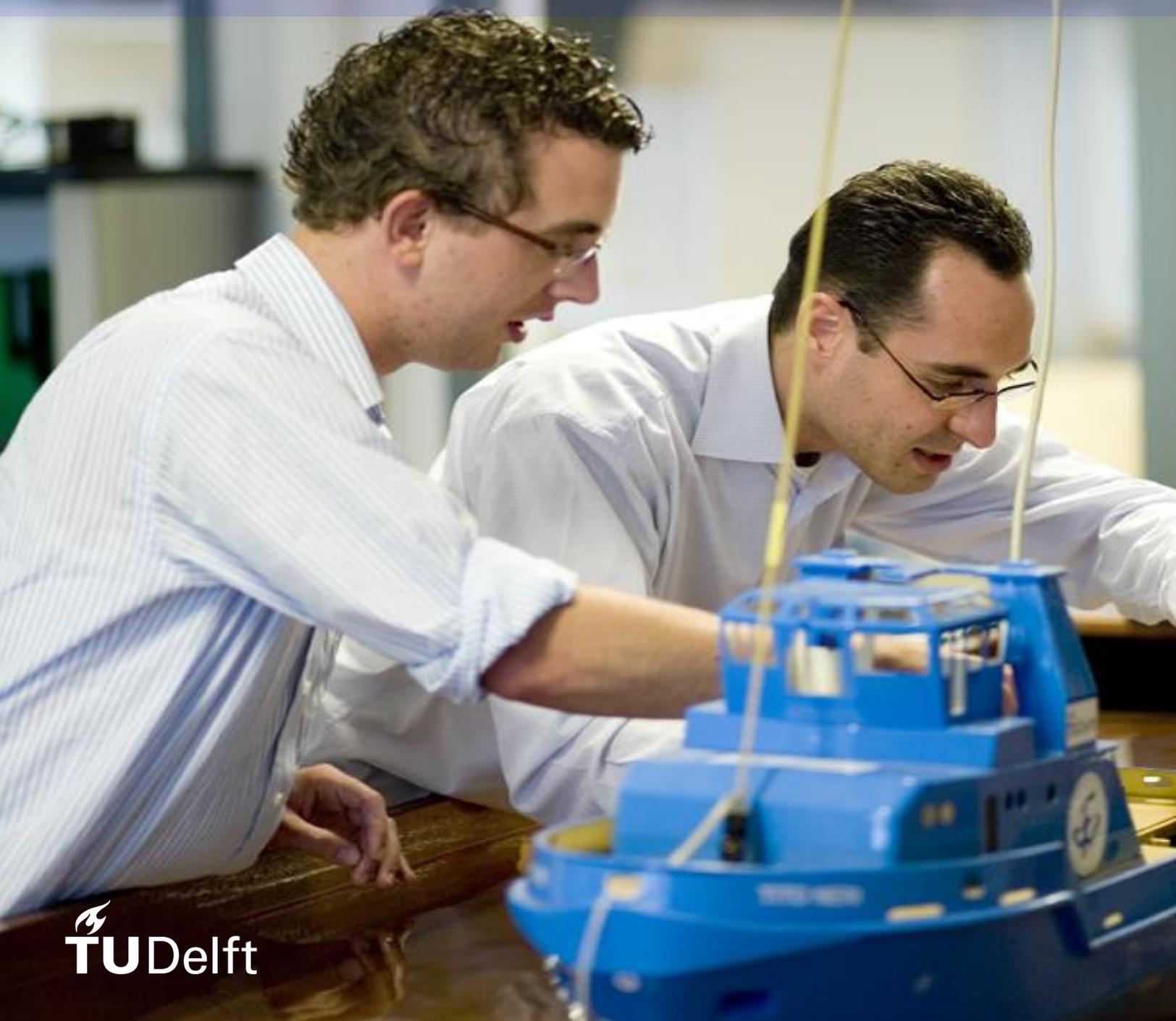


Analysis and Design of a 2.5GS/s 6-bit SAR ADC with a 3-bit/cycle Resolving Scheme

Magda Ursulean



Copyright © 2016 by the authors(s)

All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Analysis and Design of a 2.5GS/s 6-bit SAR ADC with a 3-bit/cycle Resolving Scheme

Master of Science Thesis

A thesis submitted in partial fulfillment for the degree of Master of Science

Electronic Instrumentation Laboratory

Faculty of Electrical Engineering, Mathematics and Computer Science

Microelectronics Department

Delft University of Technology

by

Magda URSULEAN

August, 2016





SECURE CONNECTIONS
FOR A SMARTER WORLD

Thesis committee:

Associate Professor Dr. Ir. Michiel Pertijs Technische Universiteit Delft, supervisor

Dr. Ir. Athon Zanicopoulos NXP Semiconductors, supervisor

Dr. Ir. Marcello Ganzerli NXP Semiconductors, supervisor

Associate Professor Dr. Ir. Marco Spirito Technische Universiteit Delft, reviewer

Copyright © 2016 by the author(s).
All rights reserved.

This thesis can be made freely available for research purposes only after the disclosure stamp has been put on this document by NXP Semiconductors.

Acknowledgements

Formally, this thesis began in August 2015, but the true start of this research dates back more than two years ago, when the members of the Microelectronics department of TU Delft awarded me a scholarship that allowed me to attend this prestigious university in the first place. Without their generous financial support, you would not be reading these pages today, which is why I would like to thank them first and foremost for giving me the opportunity to turn from a neophyte into an engineer under the guidance of the best lecturers in the Netherlands.

I am lucky to have been advised by Prof. Dr. Ir. Michiel Pertijs, who helped me continuously improve my research through many fruitful discussions, and whose careful and constant review of my thesis has led to the refined version you are reading today.

I am also very grateful to NXP Semiconductors who, through Dr. Ir. Athon Zanicopoulos and Dr. Ir. Kostas Doris, have welcomed me in a team of exemplary circuit designers who lent me their insight and experience throughout the writing of this thesis and who gave me a level of excellence to aspire to.

Last, but not least, I would like to express my sincerest gratitude to Dr. Ir. Marcello Ganzerli, who was my mentor throughout the project and who always made time to answer my questions and share his design experience with me. I am sure that the insight I gained under his supervision will give me a unique vantage point at the beginning of my career.

Magda Ursulean

Eindhoven, August 2016

Table of Contents

| | |
|---|----|
| Abstract..... | 13 |
| 1. Introduction | 14 |
| 1.1. Motivation..... | 14 |
| 1.2. Research Objectives | 14 |
| 2. Considerations on the design of high-speed ADCs in CMOS technology | 16 |
| 2.1. The impact of technology scaling on ADC design | 16 |
| 2.2. The effect of scaling on supply voltages..... | 17 |
| 2.3. The effect of scaling on sampling rate | 18 |
| 2.4. The effect of scaling on area and power dissipation..... | 19 |
| 2.5. The effect of scaling on the figure of merit (FoM) | 20 |
| 2.6. Turning to prior art for inspiration | 21 |
| 2.6.1. The Time Interleaving Technique..... | 21 |
| 2.6.2. Choosing the best sub-ADC..... | 23 |
| 3. Architecture Study | 27 |
| 3.1. Assumptions | 27 |
| 3.2. Architectures without inter-stage amplifier | 30 |
| 3.2.1. 6-bit synchronous SAR ADC (1 bit/cycle)..... | 30 |
| 3.2.2. 6-bit asynchronous SAR ADC (1bit/cycle) | 33 |
| 3.2.3. 2-bit asynchronous SAR ADC (1bit/cycle) | 35 |
| 3.2.4. 6-bit asynchronous SAR ADC (3bits/cycle) | 36 |
| 3.2.5. 6-bit flash ADC | 38 |
| 3.2.6. Pipeline of 4-bit asynchronous SAR ADC (2bits/cycle) and 2-bit flash ADC | 39 |
| 3.2.7. Pipeline of 2-bit flash ADC and 4-bit asynchronous SAR ADC (2bits/cycle) | 41 |
| 3.2.8. Pipeline of 2x 3-bit SAR ADCs..... | 42 |
| 3.2.9. Pipeline of 3x 2-bit SAR ADCs..... | 44 |
| 3.2.10. Pipeline of 2x 3-bit flash ADCs | 45 |
| 3.3. Architectures with inter-stage amplifier | 46 |
| 3.3.1. Architecture of a fast inter-stage amplifier | 46 |

| | | |
|--------|--|-----|
| 3.3.2. | Pipeline of 2-bit flash ADC and 5-bit SAR ADC (1 bit of redundancy) | 48 |
| 3.3.3. | Pipeline of 3-bit flash ADC and 3-bit flash ADC (1 bit of redundancy)..... | 50 |
| 3.4. | Conclusions | 51 |
| 4. | Operating Principle..... | 54 |
| 4.1. | ADC Architecture | 54 |
| 4.2. | Asynchronous signal generation..... | 55 |
| 4.3. | Treatment of metastable events | 56 |
| 4.4. | Verilog-A simulation results..... | 61 |
| 5. | Transistor-Level Design and Simulation..... | 66 |
| 5.1. | Track-and-Hold..... | 66 |
| 5.1.1. | Specifications | 66 |
| 5.1.2. | Possible Topologies | 67 |
| 5.1.3. | Conclusion | 70 |
| 5.2. | Digital-to-Analog Converter..... | 72 |
| 5.2.1. | Architecture | 72 |
| 5.2.2. | Schematic used for simulations | 76 |
| 5.2.3. | Control Logic | 76 |
| 5.2.4. | Study of the Reset and Settling Times of the DAC..... | 77 |
| 5.2.5. | Reset Time..... | 79 |
| 5.2.6. | Reference Generation Settling Time | 83 |
| 5.2.7. | Design Procedure and Sizing | 87 |
| 5.3. | Comparator and Preamplifier | 89 |
| 5.3.1. | Comparator | 89 |
| 5.3.2. | Speed | 90 |
| 5.3.3. | Noise | 92 |
| 5.3.4. | Reset time | 93 |
| 5.3.5. | Preamplifier..... | 94 |
| 5.3.6. | Offset for both the comparator and the preamplifier..... | 98 |
| 5.3.7. | Metastability..... | 101 |
| 5.3.8. | Layout | 103 |
| 5.3.9. | Conclusion | 105 |
| 5.4. | SAR Logic..... | 106 |

| | | |
|--------|--|-----|
| 5.4.1. | Comparator clock generator for the first set of comparators | 106 |
| 5.4.2. | Comparator clock generator for the second set of comparators | 107 |
| 5.4.3. | Ready generator | 108 |
| 5.4.4. | DAC control for the first segment..... | 109 |
| 5.4.5. | DAC control for the second segment..... | 110 |
| 5.4.6. | NAND4 and NAND2 gates for global ready generation | 110 |
| 5.4.7. | Synchronization blocks | 111 |
| 5.4.8. | Conclusion | 111 |
| 5.5. | Reference Generation..... | 112 |
| 5.6. | Calibration | 113 |
| 5.7. | Noise performance | 114 |
| 5.8. | Considerations on the total loop delay..... | 115 |
| 6. | Top Level Simulations | 118 |
| 6.1. | Summary of Performance..... | 118 |
| 6.2. | Linearity | 121 |
| 6.3. | Comparison with similar state-of-the art ADCs | 124 |
| 7. | Conclusion..... | 127 |
| | Key Contributions..... | 127 |
| | Future Work..... | 128 |
| 8. | Bibliography..... | 129 |

List of Figures

| | |
|--|----|
| Figure 2.3.1 - Average bandwidth (BW) published at ISSCC and VLSI (1997-2012) vs. CMOS node, courtesy of [6] | 18 |
| Figure 2.3.2- Evolution of sampling speed and ENOB vs. CMOS node, courtesy of [3]. The red line represents the noise limitation of the different technologies..... | 19 |
| Figure 2.4.1 - Average ENOB published at ISSCC and VLSI (1997-2012) vs. CMOS node, courtesy of [5] | 19 |
| Figure 2.5.1 -Walden FoM vs. ENOB, courtesy of [6] | 21 |
| Figure 2.5.2 - Walden FoM vs. CMOS technology node, courtesy of [6]..... | 21 |
| Figure 2.6.1.1 - Interleaving scheme, courtesy of [10] | 22 |
| Figure 2.6.2.1 - 6-bit ADCs with sampling speeds greater than 2GS/s published at ISSCC and VLSI Symposium in the past 6 years..... | 24 |
| Figure 3.2.1.1 – Block diagram of a synchronous SAR ADC..... | 31 |
| Figure 3.2.1.2 – Timing diagram of a 6-bit synchronous SAR ADC (1-bit/cycle) | 31 |
| Figure 3.2.2.1 - Timing diagram of a 6-bit asynchronous SAR ADC (1bit/cycle)..... | 33 |
| Figure 3.2.3.1 – Timing diagram of a 2-bit asynchronous SAR ADC (1bit/cycle)..... | 35 |
| Figure 3.2.4.1 - Block diagram of a 6-bit asynchronous SAR ADC (3bits/cycle)..... | 37 |
| Figure 3.2.4.2 – Timing diagram of a 6-bit asynchronous SAR ADC (3bits/cycle)..... | 37 |
| Figure 3.2.5.1 - Block diagram of a 6-bit flash ADC | 38 |
| Figure 3.2.5.2 – Timing diagram of a 6-bit flash ADC..... | 38 |
| Figure 3.2.6.1 - Block diagram of a pipeline of 4-bit asynchronous SAR ADC (2bits/cycle) and 2-bit flash ADC (with interpolation)..... | 39 |
| Figure 3.2.6.2 – Timing diagram of a pipeline of 4-bit asynchronous SAR ADC (2bits/cycle) and 2-bit flash ADC..... | 40 |
| Figure 3.2.6.3 - Block diagram of a 4-bit asynchronous SAR ADC (2bits/cycle)..... | 40 |
| Figure 3.2.6.4 - Block diagram of a 2-bit flash ADC | 40 |
| Figure 3.2.7.1 - Block diagram of a pipeline of 2-bit flash ADC and 4-bit asynchronous SAR ADC (2bits/cycle) | 41 |
| Figure 3.2.7.2 – Timing diagram of a pipeline of 2-bit flash ADC and 4-bit asynchronous SAR ADC (2bits/cycle) | 42 |
| Figure 3.2.8.1 - Block diagram of a pipeline of 2x 3-bit SAR ADCs | 43 |
| Figure 3.2.8.2 – Timing diagram of a pipeline of 2x 3-bit SAR ADCs | 43 |
| Figure 3.2.9.1 - Block diagram of a pipeline of 3x 2-bit SAR ADCs | 44 |
| Figure 3.2.9.2 – Timing diagram of a pipeline of 3x 2-bit SAR ADCs | 44 |
| Figure 3.2.10.1 - Block diagram of a pipeline of 2x 3-bit flash ADCs | 45 |
| Figure 3.2.10.2 – Timing diagram of a pipeline of 2x 3-bit flash ADCs | 45 |
| Figure 3.3.1.1 – Single-stage integrator dynamic residue amplifier, image courtesy of [34] | 47 |
| Figure 3.3.1.2 – Cascoded integrator dynamic residue amplifier (CIDRA), image courtesy of [34] | 47 |

| | |
|---|----|
| Figure 3.3.2.1 - Block diagram of a pipeline of 2-bit flash ADC and 5-bit SAR ADC (1 bit of redundancy)..... | 48 |
| Figure 3.3.2.2 – Timing diagram of a pipeline of 2-bit flash ADC and 5-bit SAR ADC (1 bit of redundancy)..... | 49 |
| Figure 3.3.3.1 - Block diagram of a pipeline of 3-bit flash ADC and 3-bit flash ADC (1 bit of redundancy)..... | 50 |
| Figure 3.3.3.2 – Timing diagram of a pipeline of 3-bit flash ADC and 3-bit flash ADC (1 bit of redundancy)..... | 50 |
| Figure 4.1.1 – Block diagram of a 6-bit asynchronous SAR ADC with a 3-bit/cycle resolving scheme..... | 54 |
| Figure 4.2.1 - Timing diagram of asynchronous signal generation | 55 |
| Figure 4.2.2 - Reference generation example | 56 |
| Figure 4.3.1 - Block diagram of asynchronous signal generation | 57 |
| Figure 4.3.2 - Outcome of a metastable event in the first set of comparators..... | 58 |
| Figure 4.3.3 - Metastable event in the first set of comparators that leads to an error | 58 |
| Figure 4.3.4 - Metastable event in the first set of comparators that does not lead to an error | 58 |
| Figure 4.3.5 - Outcome of a metastable event in the second set of comparators..... | 59 |
| Figure 4.3.6 - Error signal without the "force" circuit for a ramp input..... | 60 |
| Figure 4.3.7 - Error signal with the "force" circuit for a ramp input..... | 60 |
| Figure 4.4.1 - Comparison between track-and-hold signal (blue) and reconstructed output (pink)..... | 61 |
| Figure 4.4.2 - Error (Difference between reconstructed signal and delayed T&H signal).... | 62 |
| Figure 4.4.3 - Reference generation for a section of a sinusoidal input signal | 63 |
| Figure 4.4.4 - FFT of the reconstructed signal | 64 |
| Figure 4.4.5 - Asynchronous signal generation..... | 65 |
| Figure 5.1.1.1 - Interleaved ADC without front-end sampler [35]..... | 66 |
| Figure 5.1.1.2 - Interleaved ADC with front-end sampler [35] | 66 |
| Figure 5.1.2.1 - Clock-boosting circuit..... | 68 |
| Figure 5.1.2.2 - Differential complementary MOS track-and-hold circuit | 68 |
| Figure 5.1.2.3 -Differential NMOS track-and-hold circuit | 68 |
| Figure 5.2.1.1 – Segmented 6-bit DAC with a 3-voltage switching scheme..... | 72 |
| Figure 5.2.1.2 - Timing diagram of a segmented 6-bit DAC with a 3-voltage switching scheme..... | 73 |
| Figure 5.2.1.3 - Example of the switching scheme | 73 |
| Figure 5.2.1.4 - Segmented 6-bit DAC with a 2-voltage switching scheme | 74 |
| Figure 5.2.1.5 - Two-voltage Switching Scheme | 75 |
| Figure 5.2.2.1 - Transistor-level implementation of the DAC..... | 76 |
| Figure 5.2.3.1 - Example of DAC logic for 8Cu (or 4Cu)..... | 77 |
| Figure 5.2.3.2 - Example of DAC logic for Cu | 77 |
| Figure 5.2.4.1 - Maximum sampling frequency vs. DAC settling time | 78 |
| Figure 5.2.4.2 - Equivalent circuit for N identical DAC branches | 79 |

| | |
|---|-----|
| Figure 5.2.5.1 - Reset time – complete equivalent circuit..... | 80 |
| Figure 5.2.5.2 - Reset time - intermediate equivalent circuit | 80 |
| Figure 5.2.5.3 - Comparison of the single-ended reset waveforms for models of different complexity | 82 |
| Figure 5.2.5.4 - Reset time - simplified model | 82 |
| Figure 5.2.6.1 - Reference generation - simple model..... | 83 |
| Figure 5.2.6.2 - Reference generation - complex model | 85 |
| Figure 5.2.6.3 - Reference settling (simple model)..... | 85 |
| Figure 5.2.6.4 - Model comparison for reference settling (m=32) | 86 |
| Figure 5.2.6.5 - Waveforms for various codes (transistor-simulation)..... | 87 |
| Figure 5.2.7.1 - Settling time for different switch sizes (m=32)..... | 88 |
| Figure 5.3.1.1 - Transistor-level implementation of the Strong-Arm comparator | 89 |
| Figure 5.3.2.1 - Comparator delay vs. input signal..... | 91 |
| Figure 5.3.2.2 - Comparator delay vs. input common mode voltage ($V_{in} = V_{LSB4}$) | 92 |
| Figure 5.3.3.1 – Input-referred noise of the comparator vs. input common-mode voltage ($\Delta V_{in} = V_{LSB4}$)..... | 93 |
| Figure 5.3.4.1 - Residual voltage at the output of the comparator for various reset times .. | 94 |
| Figure 5.3.5.1 - Transistor-level implementation of the preamplifier | 95 |
| Figure 5.3.5.2 - Constant $G_m R$ bias circuit used for the preamplifiers..... | 96 |
| Figure 5.3.5.3 – Amplifier A1 used in the constant $G_m R$ bias circuit..... | 97 |
| Figure 5.3.5.4 - Preamplifier gain ($\Delta V_{in} = V_{LSB4}$) | 98 |
| Figure 5.3.6.1 - Offset of the comparator only | 99 |
| Figure 5.3.6.2 - Offset of the preamplifier only | 100 |
| Figure 5.3.6.3 - Offset of the comparator preceded by the preamplifier..... | 100 |
| Figure 5.3.7.1 - Comparator delay vs. V_{in} ($V_{in} = 2\text{half_diff}$) for different m factors | 102 |
| Figure 5.3.7.2 - Comparator power vs. V_{in} ($V_{in} = 2\text{half_diff}$) for different m factors..... | 102 |
| Figure 5.3.7.3 - Minimum input voltage for 94ps (54ps comparator only) | 103 |
| Figure 5.3.8.1 – Layout of the comparator | 104 |
| Figure 5.3.8.2 - Post-layout parasitic capacitances associated to the layout..... | 104 |
| Figure 5.3.8.3 - Delay of the comparator simulated using the schematic and the extracted layout | 105 |
| Figure 5.4.1 - Block diagram of the digital logic blocks involved in the SAR loop..... | 106 |
| Figure 5.4.1.1 – Waveforms corresponding to generation of the clock of the first set of comparators..... | 107 |
| Figure 5.4.2.1 – Waveforms corresponding to the generation of the clock of the second set of comparators..... | 108 |
| Figure 5.4.3.1 – Waveforms corresponding to the generation of the ready signal of the first set of comparators..... | 109 |
| Figure 5.4.4.1 - Example of control signal generation for the first segment of the DAC | 109 |
| Figure 5.4.4.2 – Example of the control signals of one bit of the first segment of the DAC..... | 109 |
| Figure 5.4.5.1 - Example of control signal generation for the second segment of the DAC | 110 |

| | |
|--|-----|
| Figure 5.4.5.2 - Example of the control signals of one bit of the second segment of the DAC | 110 |
| Figure 5.5.1 - Reference buffer with low output resistance [44], [8] | 113 |
| Figure 5.8.1 - Minimum loop delay vs. input signal | 116 |
| Figure 5.8.2 - Probability of metastability vs. maximum sampling frequency..... | 117 |
| Figure 6.1.1 - Distribution of simulated power consumption among circuit blocks | 119 |
| Figure 6.1.2 - FFT of the reconstructed output signal simulated at schematic level (full-scale input, f_{in} @ Nyquist, $f_s=2.5GS/s$)..... | 120 |
| Figure 6.1.3 - Error signal for a full-scale input with f_{in} @Nyquist, $f_s=2.5GS/s$, based on schematic simulations..... | 121 |
| Figure 6.2.1 - SNDR, SNR, SFDR, THD vs. sampling frequency from schematic simulations | 123 |
| Figure 6.2.2 - SNDR, SNR, SFDR, THD vs. sampling frequency expected after layout | 123 |
| Figure 6.2.3 - ENOB vs. sampling frequency from schematic simulations | 123 |
| Figure 6.2.4 - ENOB vs. sampling frequency expected after layout..... | 123 |
| Figure 6.2.5 - ENOB vs. input frequency from schematic simulations | 123 |
| Figure 6.3.1 - Positioning with respect to the Murmann Plot..... | 124 |

Abstract

Chapter 1 offers insight on the motivation that led to this research, as well as the application background and the objectives of the thesis. The next chapter is devoted to analyzing the design difficulties that arise when developing high-speed ADCs and to describing a few existing solutions. The third chapter presents the results of an extensive architecture study aimed at revealing the best converter topology for the speed required. The next section is aimed at explaining the operating principle of the architecture selected in detail in order to reveal the control signals that make its asynchronous functioning possible. The fifth chapter deals with the transistor-level implementation of the main building blocks of the ADC (track-and-hold, comparator, preamplifier, DAC and digital logic) and is focused on exposing the existing trade-offs in terms of power, speed and area. The subsequent section presents the top-level simulation results and discusses to what extent the specifications imposed at the beginning of the work were achieved. Last, but not least, chapter 7 summarizes the findings of this research while emphasizing the main contributions of the author and proposing a couple of possible improvements.

1. Introduction

1.1. Motivation

Two main trends can be identified with respect to data traffic: ever increasing speeds and a shift towards heavy processing in the digital domain. Moore's Law is useful in understanding both; on one hand, digital circuits become faster and require less area with every new technology node, which makes digital processing a lower-cost alternative compared to its analog counterpart. On the other hand, Moore's Law is a good predictor of data traffic needs, so a doubling of the existing volume is to be expected every few years.

In this context, high-performance transceivers have become increasingly reliant on high-speed analog-to-digital converters (ADCs), which tend to limit their processing capabilities. It is thus of great scientific and commercial interest to design ADCs operating in the multi-GHz range while maintaining an acceptable power consumption level.

A possible application of high-speed ADCs can be found in the vast field of high-speed serial links, of which wired transmissions are an important segment. Serial links are advantageous because of their small form factor, which owes to the need of a single transmission line and to lower pin counts, which decreases their cost and reduces the influence of various interference sources, such as crosstalk. The emergence of high-speed transmissions such as 40Gbps or 100Gbps Ethernet further motivate the need for extensive research in the area of high-speed ADCs.

1.2. Research Objectives

The purpose of this work is to describe a top-down design of a 6-bit 2.5GS/s analog-to-digital converter in 40nm CMOS technology, tailored to serve as a sub-ADC in a 20GS/s time-interleaved ADC, starting from architecture selection and leading up to the transistor implementation. The main points studied are related to:

1. Comparing various ADC topologies in order to select the best candidate architecture for achieving the desired sampling speed and estimating its power consumption,
2. Modelling the chosen architecture using Verilog-A in order to design the asynchronous control loop,
3. Exposing the trade-offs related to the different ADC blocks with respect to speed, power and area and choosing a suitable design for each in the broader context of the application such that an ADC with a linearity better than 4 bits and with a figure-of-merit of less than 450fj/conv-step results,
4. Analyzing the technological limitations that set the maximum speed of an ADC designed in 40nm CMOS technology, such as the parasitic elements,

5. Implementing the various circuit blocks and demonstrating that the ADC can indeed achieve the target performance,
6. Proposing directions for the improvement of the design.

Table 1.2.1 lists the main specifications of the design.

Table 1.2.1 - Summary of ADC Specifications

| Parameter | Desired Value |
|--------------------|---|
| Sampling frequency | 2.5GS/s |
| Input bandwidth | $\geq 10\text{GHz}$ |
| Supply voltage | 1.1V |
| Linearity | $SNDR \geq 28 \text{ dB}$ |
| Power consumption | As low as possible, without sacrificing speed |

2. Considerations on the design of high-speed ADCs in CMOS technology

Bearing in mind the applications presented in the previous chapter, it is easily understandable that high-speed, low-to-medium resolution ADCs are of great interest to the semiconductor industry, which justifies an increase of the research efforts pointed in this direction.

2.1. The impact of technology scaling on ADC design

As the minimum gate length continues to decrease in newer technology nodes, the performance of digital circuits is improved in terms of speed and power consumption. In order to accommodate this, scaling laws are applied to the voltage supply also, which leads to limited voltage headroom for analog circuits. Due to the fact that the cost of chips is driven by digital circuits and because signal processing in both signal domains is often required on the same chip, analog designers need to adapt their techniques to the ever greater miniaturization in order to ensure that their creations remain relevant and economically feasible in the future.

Competition is fierce on the ADC market and engineers are struggling to come up with ingenious solutions that push performance limits towards physical ones; in the meantime, performance demands steadily increase and the constrained design window continues to shrink with every gate length reduction. This section aims to establish which aspects of nanometer CMOS are the most problematic for high-speed data converter design. The purpose is not to be exhaustive, but rather to briefly familiarize the reader with the context in which the ADC design problem occurs.

The current scaling rules applied to mixed-signal circuits are summarized in Table 2.1.1, [1]. The scaling of this type of circuits can be classified according to the main limitation that arises from it. Thus, some mixed-signal circuits are limited by the noise level (SNR-limited), while others cannot go beyond the matching constraint (Matching-limited). It is important to note here that noise is a fundamental limitation, whereas the consequences of imperfect matching can be mitigated through the use of digital techniques, such as calibration [1].

For the sake of the argument, let us assume that an SNR-limited data converter has been designed in technology A and needs to be ported to technology B, whose feature size is S times smaller. According to Table 2.1.1, the designer can expect the area of the circuit to be S^2 larger, partly because larger capacitors are needed to keep the noise to the same relative level, which means there will also be an increase in power dissipation proportional to S [1]. One can easily understand that for analog circuits, in which noise is the biggest limitation, scaling rules do not bring any benefit, as the design window tends to shrink in newer

processes. Their matching-limited counterparts, however, experience a reduction of both area and power dissipation by a factor of S^2 [1]. Naturally, the performance, area and power consumption of digital circuits will improve, but this has already been assumed as the driving force of scaling.

Table 2.1.1- Scaling of Mixed-Signal Circuits, reproduced from [1]

| Scaling Parameter | SNR-Limited | Matching-Limited | Digital Circuits |
|-------------------|-----------------------|---------------------|------------------|
| Dynamic Range | $\propto \sqrt{C/kT}$ | $\propto \sqrt{WL}$ | Word length |
| Supply Voltage | $1/S$ | $1/S$ | $1/S$ |
| Speed | 1 | 1 | S |
| Area | S^2 | $1/S^2$ | $1/S^2$ |
| Power | S | $1/S^2$ | $1/S^2$ |

2.2. The effect of scaling on supply voltages

From an analog/mixed-signal design perspective, perhaps the most ‘dangerous’ scaling trend is the lowering of the supply voltage, which incurs a reduced signal swing and a lower amplification capability of individual transistors [2]. Another cause leading to reduced gain owes to the drain-induced barrier lowering (DIBL) effect, which lowers the output resistance of the transistors. Due to the reduced signal swing, the importance of the noise floor increases, although the absolute value of its RMS voltage remains more or less the same from one technological node to another [3]. As Table 2.1.1 shows, the dynamic range (DR) of SNR-limited circuits decreases whenever the noise level becomes higher compared to the available signal swing.

Enz and Vittoz [4] show that the power consumption of an analog circuit is directly proportional to the desired SNR (Equation (2.1), where k is Boltzmann’s constant, T represents the absolute temperature, f is the operating frequency, V_B denotes the bias voltage and V_{pp} is the peak-to-peak amplitude required at the output of the circuit). When moving to a smaller technology, this translates to a power penalty, as more transistors are needed in order to achieve the same SNR level. It is also clear from the equation that a reduced signal swing at the output negatively impacts the power consumption.

$$P = 8kT \cdot f \cdot SNR \cdot \frac{V_B}{V_{pp}} \quad (2.1)$$

Additionally, when supply voltages decrease, some of the existing design topologies are no longer practicable because they rely on cascoded transistors for which voltage headroom is no longer sufficient. The issue of porting a design to a more advanced technology thus

becomes increasingly problematic and novel circuit techniques and tricks need to be employed for this purpose.

2.3. The effect of scaling on sampling rate

Nevertheless, the design landscape shaped by technology scaling is not home only to a bleak future, as the miniaturization trend is known to improve the frequency performance of transistors. Due to the smaller area that is occupied by a single transistor, the intrinsic capacitances of the devices are reduced, which allows them to be operated at higher frequencies [2]. From this point of view, it appears that if a designer is able to extract enough gain from the transistors provided, higher speed (i.e.: increased bandwidth) is almost ‘for free’ in newer technology nodes.

Smaller devices bring an increase in f_T , the transition frequency of the transistor, and as a consequence, higher analog bandwidths (and implicitly higher sample frequencies) can be supported in applications designed in lower technology nodes [5]. Published designs confirm this theory, as Figure 2.3.1 illustrates.

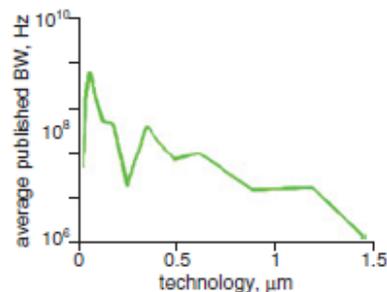


Figure 2.3.1 - Average bandwidth (BW) published at ISSCC and VLSI (1997-2012) vs. CMOS node, courtesy of [6]

In 2010, Jonsson [3] analyzed the effects of scaling on the performance of ADCs by comparing converters with the same effective number of bits (ENOB) and sampling rate. The study was performed on ADCs with effective resolutions of 4, 8, 12 and 15 bits and its results can be seen in Figure 2.3.2.

Although an increase in sampling speed can be observed for all cases, the graphs show that miniaturization is more beneficial to lower-resolution data converters than to their high-resolution counterparts, in the sense that higher sampling rates are possible in the low-ENOB case for the same technology node [3].

Thinking back on Equation (2.1), with every increase in operation frequency, more power needs to be spent to achieve the same SNR. This frequency can increase because

newer applications require higher bandwidth and in turn, this leads to more noise being integrated.

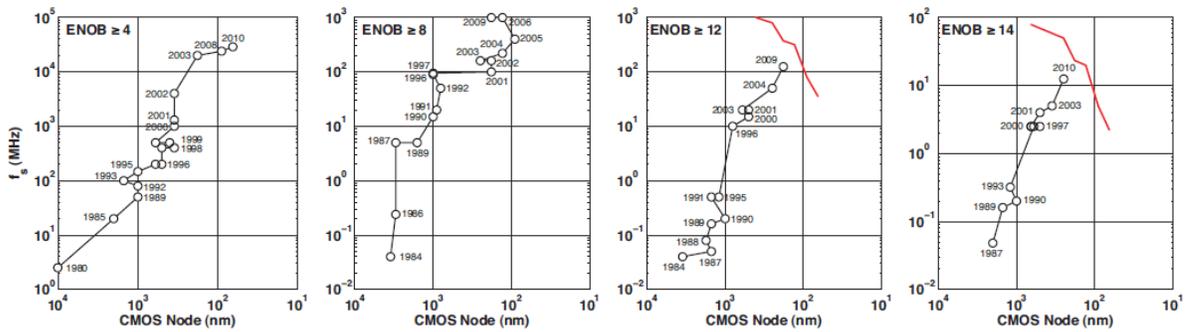


Figure 2.3.2- Evolution of sampling speed and ENOB vs. CMOS node, courtesy of [3]. The red line represents the noise limitation of the different technologies

2.4. The effect of scaling on area and power dissipation

Verhelst and Murmann [5] have surveyed converters published at the ISSCC and VLSI conferences between 1997 and 2012 and reached the same conclusion regarding the relation between technology node and resolution. As Figure 2.4.1 shows, there is a clear preference to implement low-resolution and medium-resolutions ADCs with small-feature transistors.

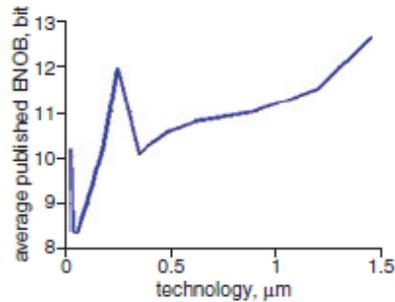


Figure 2.4.1 - Average ENOB published at ISSCC and VLSI (1997-2012) vs. CMOS node, courtesy of [5]

Equations (2.2) and (2.3), where λ is used to denote the feature size and f_N is the Nyquist sampling frequency, are used by Verhelst and Murmann [5] to model the area and the power in the surveyed ADC designs. Applying linear regression to these allows a fair comparison of scaling effects on different converter architectures. Table 2.4.1 and Table 2.4.2, reproduced here for the sake of completeness, illustrate the values obtained for the fitting coefficients.

$$A = 10^{\alpha_0} \lambda^{\alpha_1} 2^{\alpha_2 \cdot ENOB} f_N^{\alpha_3} \quad (2.2)$$

$$P = 10^{\beta_0} \lambda^{\beta_1} 2^{\beta_2 \cdot ENOB} f_N^{\beta_3} \quad (2.3)$$

The authors show that power consumption scales with $\sim\lambda^{1.7}$ for all converter types, and with 2^{ENOB} for all architectures except for flash, which scales twice as faster with respect to ENOB [5]. The tables also show that the power consumption of the SAR ADC is the one that benefits the most from a smaller feature size, due to its mostly digital components. Moreover, area is seen to scale with $\sim\lambda^{1.6}$ and much less than the power with respect to the ENOB. As regards the Nyquist sampling rate, it affects the power consumption in a linear fashion, while the area is mostly left unaltered.

Table 2.4.1 - Area fitting coefficients, courtesy of [5]

| Coeff | Flash | Fold | Pipe | Sigma | SAR | All |
|-----------------------|-------|------|------|-------|------|------|
| $\alpha_1 (\lambda)$ | 1.28 | 1.53 | 1.77 | 1.45 | 1.41 | 1.6 |
| $\alpha_2 (2^{ENOB})$ | 0.46 | 1.07 | 0.50 | 0.30 | 0.11 | 0.5 |
| $\alpha_3 (f_N)$ | -0.06 | 0.48 | 0.26 | 0.42 | 0.27 | 0.3 |
| Sigma | | | | | | |
| $\alpha_1 (\lambda)$ | 0.35 | 0.43 | 0.10 | 0.11 | 0.35 | 0.07 |
| $\alpha_2 (2^{ENOB})$ | 0.56 | 0.24 | 0.06 | 0.07 | 0.16 | 0.03 |
| $\alpha_3 (f_N)$ | 0.10 | 0.22 | 0.06 | 0.04 | 0.07 | 0.02 |

Table 2.4.2 - Power fitting coefficients, courtesy of [5]

| coeff | Flash | Fold | Pipe | Sigma | SAR | All |
|-----------------------|-------|------|------|-------|------|------|
| $\alpha_1 (\lambda)$ | 1.71 | 1.34 | 1.62 | 1.35 | 2.30 | 1.7 |
| $\alpha_2 (2^{ENOB})$ | 1.90 | 1.11 | 0.62 | 0.41 | 0.72 | 1.0 |
| $\alpha_3 (f_N)$ | 1.31 | 0.76 | 0.95 | 0.90 | 1.31 | 1.1 |
| Sigma | | | | | | |
| $\alpha_1 (\lambda)$ | 0.28 | 0.40 | 0.11 | 0.14 | 0.29 | 0.08 |
| $\alpha_2 (2^{ENOB})$ | 0.44 | 0.22 | 0.07 | 0.09 | 0.13 | 0.04 |
| $\alpha_3 (f_N)$ | 0.08 | 0.20 | 0.07 | 0.05 | 0.05 | 0.02 |

2.5. The effect of scaling on the figure of merit (FoM)

Jonsson [6] carried out a study aimed at revealing the most important features of Figures of Merit (*FoMs*) used in publications that report on data converters and to investigate whether the different FoMs used by different authors are biased towards one performance indicator or another. Aside from featuring an extensive list of possible FoMs, the paper also provides some insight into the characteristics of the most popular FoM, which will be explained in what follows.

Equation (2.4) represents the Walden FoM (also known as the ISSCC FoM due to its popularity at the conference), where P denotes the power dissipation, f_s is the Nyquist sampling frequency and *ENOB* is the effective number of bits.

$$FoM = \frac{P}{2^{ENOB} \times f_s}, \text{ referred to as } F_{A1} \text{ by Jonsson} \quad (2.4)$$

Figure 2.5.1 plots the Walden FoM vs. ENOB in black, the low-resolution energy plateau in red and the thermal noise slope in blue. It is interesting to observe that this particular figure of merit has a “sweet spot” [6] around the intersection of the low-resolution energy plateau with the thermal noise slope, which means that if these two boundaries continue to exist, the smallest FoMs will continue to occur around an ENOB of 8-9 bits [6]. As Jonsson points out, the existence of this “sweet-spot” means that in order to correctly assess converter designs with the Walden FoM, only ADCs with the same ENOB should be

compared [6]. In a way, this means doing an application-centric comparison (supposing that an application requires a fixed ENOB), rather than a general one.

Using Figure 2.5.2 to plot different FoMs vs. the technology nodes in which the converters were implemented, Jonsson shows that generally, the average FoM decreases when moving towards smaller transistor features [6]. This fact also suggests that it would be more insightful to use the Walden FoM to compare designs that are realized in the same technology node before attempting to draw any conclusions about their overall efficiency.

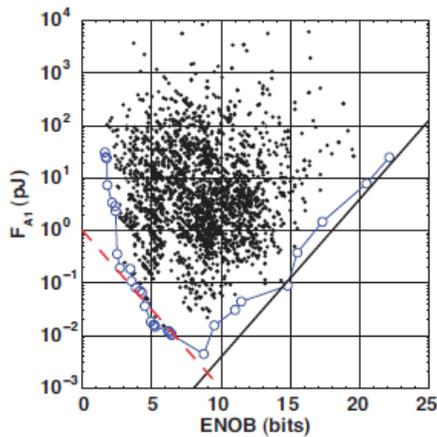


Figure 2.5.1 -Walden FoM vs. ENOB, courtesy of [6]

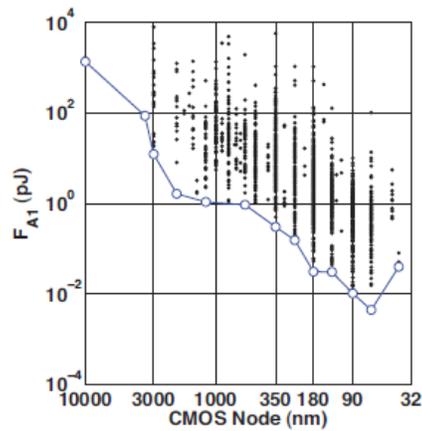


Figure 2.5.2 - Walden FoM vs. CMOS technology node, courtesy of [6]

2.6. Turning to prior art for inspiration

After having discussed the characteristics of modern technologies, the natural next step is of surveying a few successful attempts in designing high-speed analog-to-digital converters in order to understand the inherent trade-offs of such an effort. This section will explain why the time-interleaved (TI) successive-approximation register (SAR) ADC is the most promising architecture for the implementation of a low-to-medium resolution converter for high-speed applications.

2.6.1. The Time Interleaving Technique

Designing a high-speed ADC without exceeding an imposed power budget is no trivial task; Kull proposes one possible solution in [7] and demonstrates with measurement results that the methods chosen are appropriate for the task at hand. By using a combination of time interleaving and comparator alternation, the design achieves a sampling frequency of up to 100GS/s in a 32nm SOI CMOS technology. Interleaving is exploited by introducing four samplers and 64 SAR sub-ADCs. The SAR architecture was chosen for its low-power

operation and mostly-digital circuitry, which helps take advantage of the fact that submicron technologies have been optimized for digital design [7]. The designer has chosen to alternate comparators in order to trade area and power dissipation for speed. By providing two of comparators, one array can ensure the correct ADC operation, while the other is reset to eliminate any residual information from the previous conversion [7]. This technique shortens the critical path with an amount equal to the comparator reset time.

The fact that successful designs operating at speeds that can even reach 100GS/s exist is an indicator that some circuits are better suited for the task than others. It appears that the architecture of the ADC is restricted by the high speed specification, in the sense that a single-channel converter cannot exceed sampling speeds above 25GS/s even in the most advanced CMOS technology node [7].

This technique allows the designer to trade area and power consumption for speed by replicating some of the blocks in the design and operating them in a parallel fashion (Figure 2.6.1.1). This can be done to a single component - the comparators, for example [8] or to the ADC itself - [9]. Unfortunately, time interleaving requires the generation of N clock phases, where N is the interleaving factor (equal to the number of channels), which also adds to the total power consumption of the circuit [10]. Errors such as offset and mismatch can sometimes grow out of proportion in this approach, but seeing that they are not fundamental limits, they can be removed through calibration.

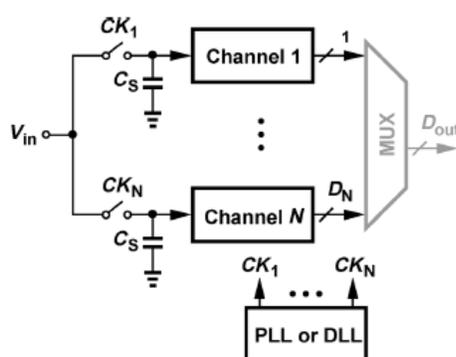


Figure 2.6.1.1 - Interleaving scheme, courtesy of [10]

Despite incurring some penalties, interleaving is also a means of improving the figure-of-merit (FoM) of an ADC; in the absence of this technique, any speed enhancement leads to significantly more power dissipation, which lowers the FoM, but interleaving can combat this trend by ensuring that the speed increase is more important than the additional power dissipation, which makes the converter a more efficient design overall [10]. Additionally, by adopting an interleaving scheme, the probability that the comparators will run into metastability issues also decreases to a large extent [10], which means that speed gains can

also be acquired from the analog components which are usually regarded as being the bottlenecks in many data converter designs.

Having decided to opt for an interleaving structure, there are a couple of requirements that the sub-ADC needs to fulfill in order to ensure that the overall design remains within the performance specifications. Kull advises that two different ratios can be used to estimate the efficiency of the sub-ADCs, namely the $\frac{\text{power}}{\text{area}}$ and $\frac{\text{speed}}{\text{area}}$ ratios [8]. The designer's goal should be to minimize the first, while attempting to increase the second. Despite the fact that speed is the most sought-after goal in this case, a low-power and low-area solution is still desirable because the sub-ADC will be replicated several times in the design, meaning that the total power and area budget could be exhausted without achieving the target interleaving factor. [8] summarizes the requirements for the sub-ADC and explains the consequences of not meeting them.

As Razavi points out [10], before the design of the sub-ADC is finalized, the engineer must find the optimal number of interleaved stages. He quotes [11] as one of the first papers to acknowledge that an optimum number of sub-ADCs can lead to the minimum power consumption, while additional stages bring no speed benefit at high dissipation costs.

2.6.2. Choosing the best sub-ADC

In order to better understand the state-of-the-art research in the field of ADCs, the Murmann plots were surveyed in order to find the 6-bit designs that operate at frequencies higher than 2GS/s, regardless if this is achieved in a time-interleaved fashion or not. Table 2.6.2.1 summarizes the most promising designs published at ISSCC and the VLSI Symposium in the past years which fit the search criteria.

Table 2.6.2.1 - 6-bit ADCs with sampling speeds greater than 2GS/s published at ISSCC and VLSI Symposium in the past years

| Reference and year | Architecture | Speed | Speed of sub-ADC | SNDR | Power | FoM (f/conv-step) |
|--------------------|-----------------------------|---------|------------------|-------|-------|-------------------|
| [12] 2015 | SAR, TI | 5GS/S | 1.25GS/s | 30.8 | 5.5m | 39 |
| [13] 2015 | SAR, TI | 10GS/S | 312.5MS/s | 30.3 | 79m | 295 |
| [14] 2015 | Two-step, binary search, TI | 25GS/s | 3.125GS/s | 29.7 | 88m | 141 |
| [15] 2014 | Flash, TI | 20GS/s | 2.5GS/s | 30.7 | 69.5m | 124 |
| [16] 2013 | Flash | 5GS/s | 5GS/s | 30.9 | 8.5m | 59 |
| [17] 2013 | Flash, TI | 10GS/s | 2.5GS/s | 29.2 | 139m | 59 |
| [18] 2010 | Pipe, Folding, TI | 2.2GS/s | 550MS/s | 31.1 | 2.6m | 40 |
| [19] 2009 | Flash, TI | 10GS/s | 2.5GS/s | 31.6 | 390m | 1219 |
| [20] 2008 | Flash | 5GS/s | 5GS/s | 32 | 320m | 1968 |
| [21] 2007 | Flash | 3.5GS/s | 3.5GS/s | 31.18 | 98m | 950 |

| | | | | | | | |
|------|------|-------|-------|-------|----|------|------|
| [22] | 2004 | Flash | 4GS/s | 4GS/s | 30 | 990m | 9581 |
| [23] | 2003 | Flash | 2GS/s | 2GS/s | 30 | 310m | 6000 |

The data show that single-channel ADCs running at sampling speeds higher than 2GS/s have been successfully implemented using flash ADCs, which are traditionally employed in high-speed application at the expense of prohibitive power consumption. A single binary-search ADC manages to exceed the target speed, but its dissipation is higher than the flash solutions, rendering it more inefficient.

Figure 2.6.2.1 shows the figure-of-merit achieved by the designs in Table 2.6.2.1 vs. their overall sampling speed. The plot shows that for time-interleaved ADCs, using a flash as sub-ADC generally results in designs that do not maximize efficiency, while a SAR and pipeline-based solution tends to be less power hungry and thus renders a lower FoM.

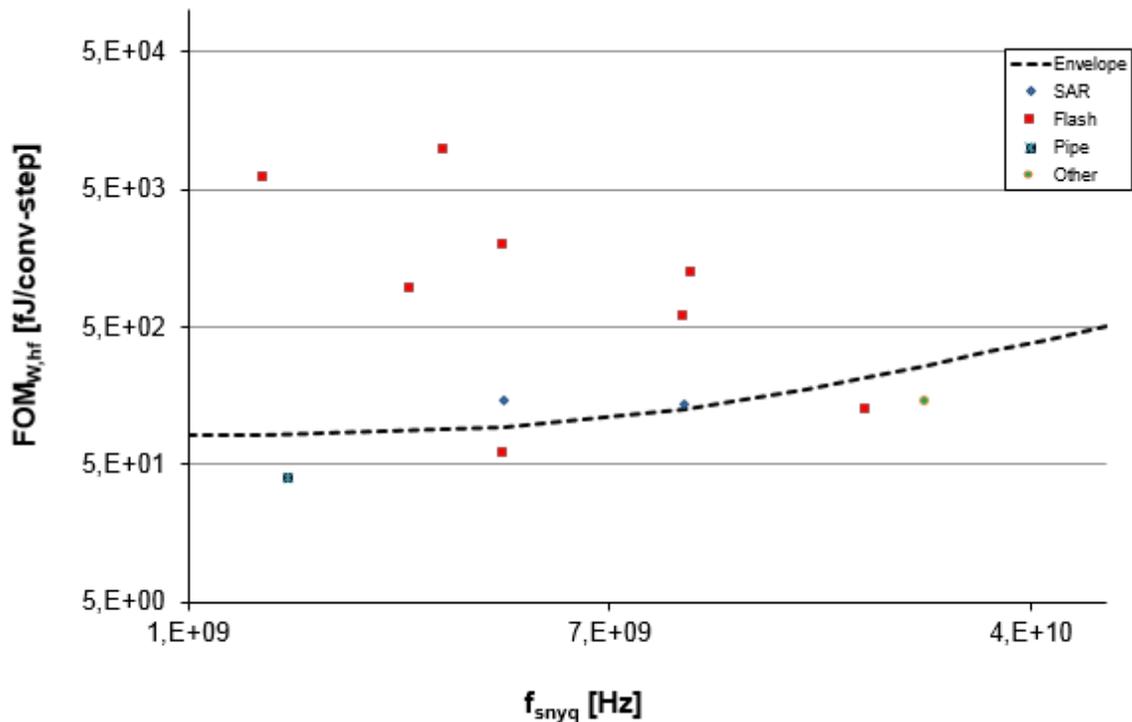


Figure 2.6.2.1 - 6-bit ADCs with sampling speeds greater than 2GS/s published at ISSCC and VLSI Symposium in the past years

In what follows, a few high-speed techniques collected during the literature review will be explained.

Huang et al., [24], demonstrate an 8-times interleaved ADC, achieving 6-bit resolution and a sampling rate of 16GS/s. The sub-ADC is built using a flash topology and aiming for the same resolution and requires 63 comparators for 6-bit accuracy. The authors opted for comparators without preamplifiers in an attempt to reduce the area, and solved the problem of offset between channels using on-chip background digital calibration [24]. Implemented

in 65nm CMOS, the converter consumes 435mW (FoM = 2.6pJ/conversion-step) and occupies an area of 1.47mm² [24]; these values prohibit a possible increase in the number of channels and thus sampling rate. The main disadvantage of this design is the fact that it requires 63 comparators/sub-ADC and this value increases exponentially if the resolution is augmented, incurring a high power and area penalty. Huang et al. manage to show that although the flash architecture is known to be the fastest among ADC types, this local optimum does not necessarily lead to a global optimum when included in a time-interleaved data converter. As Kull pointed out [8], the purpose is not only to include the fastest sub-ADC available, but also to ensure that this ADC has a small enough area to allow massive interleaving without leading to grotesquely large circuits. It is easily noticeable from the flash architecture that this is not the case.

The design in [12] is also completed in a 65nm CMOS technology, which allows an easy comparison with the previous one presented. Chan et al. opted for the same 6-bit resolution, but for a sampling speed of only 5GS/s, which is more than 3 times lower than that of [24]. Intuitively, it is expected that the power consumption and area will be at least 3 times lower. This is based on employing the same architecture, which is not what the authors decide for. Instead, they choose an interleaving factor of only 4 and a 3-bit/cycle resolving fashion for a SAR ADC. The resulting power consumption, including calibration is of 10.6mW (FoM = 39fJ/conversion-step), while the silicon area occupied is about 0.09mm² [12], which is a lot less than what could be expected if the design in [24] were used for the design of a slower ADC. This goes to show that the SAR architecture, if combined with a few tricks such as a multi-bit/cycle approach, is not a power-and area-hungry architecture that can easily achieve GS/s sampling rates. Due to its small form factor and significant energy savings, this ADC could also be used as a sub-ADC in a larger interleaving scheme.

A similar idea as the one based on a SAR sub-ADC can be found in [25], where Spagnolo et al. use a pipelined binary search (PLBS) algorithm and a few 2-bit flash converters to implement a 6-bit ADC clocked at 3.5GS/s in a 40nm low-power (LP) CMOS technology. Calibration is also needed here, and the authors use an interleaving factor of 4 to achieve a power consumption of 4.1mW (FoM = 48fJ/conversion-step) and an active area of 0.03mm². Thinking back on the design in [12], which achieves a higher sampling rate - albeit at a slightly higher power and area consumption, it appears that still the SAR architecture is the most likely candidate for efficient high-speed ADCs running at GS/s rates.

Another interesting design is featured in [12], where a 6-bit SAR ADC with a 3-bit/cycle resolving scheme is introduced. In contrast to [8], this paper describes a circuit implemented in an older technology (65nm), so more architectural innovation is needed to obtain a comparable FoM (39fJ/conv-step). The design is 4x interleaved and is comprised of sub-ADCs running at 1.25GS/s. The idea behind the architecture is to combine the high-speed operation of the flash ADC with the power efficiency of the SAR, so a multi-bit approach is adopted to achieve this. Two comparator stages are employed to determine the output bits while reducing the loop delay compared to the conventional SAR approach. The

3-bit flash ADC that results in each stage has a low enough resolution to not lead to excessive power dissipation, and has the advantage that it can determine 3 bits in the time required for a single comparator decision, which is roughly three times less than what is needed to perform the same operation in a traditional SAR ADC.

In order to keep the power dissipation as low as possible, interpolation - a technique typically applied to flash ADC to reduce hardware - is done at the outputs of the dynamic preamplifiers situated before the comparators. This choice is dictated by an expected increase in kickback, if interpolation were to be done at the outputs of the comparators themselves, which might compromise the accuracy of the result.

Since the time available for each comparison set is prescribed by the synchronous clock, a digital error-correction mechanism is included to correct the errors that appear for small input signals that require more decision time than is given.

The research landscape indicates a scarcity of viable solutions for sub-ADCs running in the 2-2.5GS/s range at reasonable power consumption levels, so it is worthwhile to attempt to bring the speed of a single-channel ADC towards this frequency. In the past 6 years, most of the designs running at high-speed rely on flash converters to function, and for this reason their power dissipation verges on the extreme. A more efficient ADC topology, such as the SAR or the pipelined ADC should be considered for implementation in the GHz range, despite the fact that these were traditionally used for medium-speed applications. In what follows, the results of an architecture study aimed at revealing the best architecture for realizing a single-channel ADC running at 2.5GS/s will be presented.

3. Architecture Study

The aim of the architecture study is to assess the speed and power performance of several ADC topologies in order to determine which of them are capable of reaching a speed of 2.5GS/s while maintaining a reasonable level of power consumption. Both high-speed and low-power ADCs were investigated for a resolution of 6 bits in order to expose the trade-offs between them.

3.1. Assumptions

To allow a fair comparison, a couple of assumptions were made concerning the building blocks of these ADCs, which were considered to have the same requirements in all cases. The calculations that follow are based on the performance specifications of a previously simulated 8-bit ADC in 40nm technology [26] and on previous design experience from NXP using this technology. The speed and power of the 8-bit design constitute the starting point for the speed and power estimations of the converter which needs to be designed, unless otherwise specified.

For the converter architectures that require a DAC, it is assumed that a unary capacitive DAC is used, where the unit capacitor is the same as that in [26] and has a value of 300aF. The input-capacitance of the preamplifier used in the 8-bit design is measured and found to be 12fF and constitutes the working assumption for this block. Last, but not least, the value of 100Ω is taken as the guideline for the on-resistance of the switches used to operate the circuit, as this value allows short settling times to be achieved with a reasonable transistor size. These assumptions are summarized in Table 3.1.1, together with the notations adopted to designate them.

Table 3.1.1 – Assumptions regarding various circuit elements

| Parameter | Assumption |
|---|------------|
| On-resistance of the sampling switch (R_{on}) | 100Ω |
| On-resistance of the DAC bottom-plate switch (R_{sw}) | 100Ω |
| Parasitic capacitance of a switch with an $R_{on} = 100\Omega$ (C_{para}) | 15fF |
| Input capacitance of the preamplifier (C_{pre}) | 12fF |
| Total DAC capacitance (C_{DAC}) | $64 * C_u$ |
| Unit capacitance for the DAC (C_u) | 300aF |
| MSB Capacitance of the DAC | $32 * C_u$ |

The first approximation for the on-resistances was chosen to be 100Ω because this value can be achieved with a switch that is not excessively large. The value of the unit capacitor assumed corresponds to the one used in [26].

The StrongArm comparator is the speed bottleneck in the 8-bit design, so it is important to understand what design parameters affect its speed. Wicht et al. [27] provide a formula that exposes all the contributors to the total delay:

$$t_{comp} = t_{inv} + t_{int} + \tau_{reg} \ln \left(\frac{V_{out}}{G\Delta V_{in}} \right) \quad (3.1)$$

It is assumed that the outputs of the comparators are connected to an inverter with a delay of t_{inv} and because of this, a valid output is obtained as soon as $V_{out} = \frac{V_{DD}}{2}$. The integration and regeneration time are accounted for by t_{int} and τ_{reg} , respectively. A high gain of the comparator and preamplifier (G) tends to increase the speed of the device, while a low input signal (ΔV_{in}) tends to decrease it, leading to a metastable event. For the linearity requirements of a 6-bit design and the scope of this architecture study, it is appropriate to assume that the minimum input signal is $\Delta V_{in} = \frac{V_{LSB}}{4}$ and that any value below this threshold leads to a metastable event. It is less important at this stage which circuit conditions influence each contributor, and assumptions made at this level of abstraction are sufficient for architecture comparison purposes. These assumptions are summarized in Table 3.1.2 and their viability will be addressed in a later chapter.

Table 3.1.2 – Timing Assumptions for the Comparator

| Parameter | Assumption |
|---|---------------------|
| Inverter delay (t_{inv}) | 30ps |
| Integrating delay (t_{int}) | 16ps |
| Regeneration delay (τ_{reg}) | 12ps |
| Valid output voltage (V_{out}) | $\frac{V_{DD}}{2}$ |
| Gain (G_{comp}) | 2 |
| Preamplifier gain (G_{pre}) | 3 |
| Total gain ($G = G_{comp}G_{pre}$) | 6 |
| Minimum input voltage (ΔV_{in}) | $\frac{V_{LSB}}{4}$ |

Based on these values, a general set of assumptions can be used to estimate the speed of several fundamental ADC blocks, as listed in Table 3.1.3:

Table 3.1.3 – Timing Assumptions

| Parameter | Assumption |
|--|------------|
| Digital logic delay (t_{logic}) | 80ps |
| Preamplifier delay (t_{pre}) | 10ps |
| Maximum comparator delay obtained for an input of $\frac{V_{LSB}}{4}$ (t_{comp}) | 84ps |

The power performance reported by [26] can be used to determine the energy consumption of each block if the frequency, number of blocks and number of times they are activated are taken into account.

For example, the total power of the comparators and associated preamplifiers in the 8-bit design is $P_{comp_pre_8b} = 12.7\text{mW}$ and the 3-2-3-bit resolving scheme implies that the comparators and preamplifiers are used 17 times at a frequency of 2GHz. It follows that the energy per comparison is equal to $E_{comp_pre} = \frac{12.7\text{mW}}{2\text{GHz} \cdot 17} = 0.374\text{pJ}$. Similar reasoning can be applied to the DAC, digital logic and clock generation circuits, leading to the values in Table 3.1.4.

The track-and-hold circuit, however, has quite different requirements in the case of a 6-bit design, so the energy estimation assumes that it is comprised of a simple CMOS switch and that there are three tapered inverters driving it. For a single NMOS transistor in the 40nm technology, it holds that $C_G R_{on} = 2.7\text{ps}$, which means that for an on-resistance $R_{on} = 100\Omega$, a capacitance of $C_G = 27\text{fF}$ is associated. The power needed to drive this switch at a frequency of 2GHz is thus given by:

$$P_{T\&H} = V_{DD}^2 f C_G \left(1 + \frac{1}{2} + \frac{1}{4}\right) \cong f * 52 * 10^{-15} \text{W} \quad (3.2)$$

This corresponds to an energy of $E_{T\&H} = 0.052\text{pJ}$ per conversion.

Table 3.1.4 – Power Assumptions

| Parameter | Assumption |
|--|------------|
| Comparator and preamplifier energy (E_{comp_pre}) | 0.374pJ |
| DAC switching energy (E_{DAC}) | 0.05pJ |
| SAR logic switching energy (E_{logic}) | 0.333pJ |
| Energy required for clock generation (E_{clk}) | 0.033pJ |
| Track and Hold energy ($E_{T\&H}$) | 0.052pJ |
| Latch energy (E_{latch}) – assumed 5x lower than E_{logic} | 0.0666pJ |

It is assumed that all of the designs surveyed are able to achieve a linearity of 4.5 bits, so in order to combine the speed and power performance into a single number, a simplified figure-of-merit is used:

$$FoM = \frac{\text{power consumption}}{\text{speed}} \quad (3.3)$$

This simplified figure-of-merit can be easily converted to the Walden figure-of-merit by dividing the value by $2^{ENOB} - 2^{4.5}$ in this case:

$$FOM_{Walden} = \frac{FOM}{2^{ENOB}} \quad (3.4)$$

3.2. Architectures without inter-stage amplifier

Some ADCs compute the output code by determining part of the bits first, and then subtracting the reconstructed analog value from the output signal to generate a residue. Depending on the implementation chosen, a gain stage may be inserted between the two ADC sections such that the residue signal has the same dynamic range as the input, which lowers the constraints placed on the second ADC section in terms of noise and distortion. In this first section, only architectures without this gain stage are discussed.

3.2.1. 6-bit synchronous SAR ADC (1 bit/cycle)

The idea behind the Successive-Approximation Register Analog-to-Digital Converter (SAR ADC) first appeared in an age when less sophisticated problems needed to be solved, specifically that of determining the weight of an unknown object (henceforth denoted by X) when binary-weighted weights and a scale were available [28]. The algorithm implies comparing the object with each known weight, starting from the heaviest. If the current comparison shows that the weight is lighter than the object, then the next comparison will be done by adding the second-heaviest weight left. Otherwise, the current weight will be removed from the scales before adding the next one.

An extrapolation to electronics easily proves that if n binary-weighted weights are available, there are n quantization levels that specify the unknown value and thus allow the experimenter to represent X with an accuracy of n bits. Similarly, an unknown voltage X can be quantized if $2^n - 1$ voltage references and a comparator are available.

An intuitive block diagram of a 6-bit SAR ADC is presented in Figure 3.2.1.1 and reveals the essential components of the architecture. A track-and-hold circuit is responsible for sampling the input signal (V_{in}) and retaining its value during the conversion process, while a comparator is in charge of deciding the output bits sequentially. A digital-to-analog converter is controlled by digital logic in order to generate the voltage references for the next comparison based on the current output of the comparator. The loop is evaluated several times (once per bit) and thus the internal operating frequency of a SAR ADC needs to be much higher than its sampling frequency.

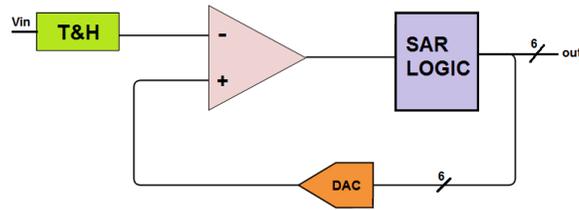


Figure 3.2.1.1 – Block diagram of a synchronous SAR ADC

Supposing that the internal clock period is $t_{internal_clk}$ and that sampling can be accurately done in one clock period, the conversion time of an N -bit SAR ADC can be approximated by:

$$t_{conv} = (N + 1)t_{internal_clk} \tag{3.5}$$

The formula prescribes a resolving time of $t_{internal_clk}$ for each evaluation of the loop and is thus equal to the critical path. In order to determine the minimum period of the internal clock, the timing diagram in Figure 3.2.1.2 is useful.

| T&H | TRACK | HOLD | | | | | | | | | | | | | |
|---------|----------------|----------------|---|----------------|---|----------------|---|----------------|---|----------------|---|----------------|---|----------------|---|
| CMP+PRE | | C_max | | C_max | | C_max | | C_max | | C_max | | C_max | | C_max | |
| LOGIC | | | P | | P | | P | | P | | P | | P | | P |
| DAC | | | S | | S | | S | | S | | S | | S | | S |
| | t_{int_clk} | t_{int_clk} | |

Figure 3.2.1.2 – Timing diagram of a 6-bit synchronous SAR ADC (1-bit/cycle)

In principle, the loop delay needs to accommodate the comparator resolving time (C_{max}), the propagation delay through the digital logic (P) and the settling time of the DAC (S). In the case of a synchronous design (i.e.: a well-defined internal clock frequency, derived – for example - using a multiple of the sampling clock), the worst-case delays need to be considered, as it is difficult to predict in which cycle they will occur and whether their incidence is simultaneous. It is assumed that the DAC is reset during the tracking phase and that the delay of the logic and the settling time of the DAC match the comparator reset time, such that a previous decision does not affect the current one.

The worst-case delay for the comparator depends on the minimum allowable input voltage, while the worst-case settling for the DAC happens for the MSB. Bearing this in mind and using the notations at the beginning of the chapter, the worst-case conversion time can be calculated as:

$$t_{conv} = t_{track} + 6(t_{comp_{LSB}} + t_{preamp} + t_{DAC_{MSB}} + t_{logic}) \quad (3.6)$$

Seeing that the noise requirements are fulfilled even with a very small sampling capacitance, the track-and-hold circuit can sample the input signal directly on the input capacitance of the comparator, which means the tracking time is imposed by this capacitance and is further limited by the parasitic capacitance of the tracking switch. If an error of $\frac{1}{2}V_{LSB}$ is considered tolerable and the assumptions at the beginning of the chapter are used, then the tracking time is given by:

$$t_{track} = (N + 1) \ln 2 R_{on} (C_{pre} + C_{para}) = 19ps, \quad (3.7)$$

where N is the number of bits defining the accuracy, R_{on} is the on-resistance of the switch, C_{pre} is the input capacitance of the preamplifier and C_{para} is used to denote the parasitic capacitance of the switch.

Such a narrow pulse width cannot be propagated easily in the 40nm CMOS technology that is at the designer's disposal, because its width tends to shrink and disappear when passed through digital blocks, so the minimum pulse width of 50ps will be considered for the tracking time.

Similar considerations regarding the accuracy can be made concerning the DAC settling time. As such, the MSB settling is given by:

$$t_{DAC_{MSB}} = (N + 1) \ln 2 R_{on} (C_{MSB} + C_{pre} + C_{para}) = 24ps \quad (3.8)$$

It follows that the minimum achievable conversion time is $t_{conv} = 1.238ns$, which points to a sampling frequency of $f_{conv} = 0.807GHz$.

The assumptions for the energy of each block and the fact that we use the comparator, logic and DAC 6 times for each conversion, while the track-and-hold is used only once, reveals the power of the 6-bit synchronous SAR ADC architecture:

$$P_{total} = f (6(E_{comp_{pre}} + E_{clk} + E_{DAC} + E_{logic}) + E_{T\&H}) = 3.867mW \quad (3.9)$$

The efficiency of the design can be evaluated and compared to the others based on the figure of merit defined previously:

$$FoM = \frac{3.867mW}{807MHz} = 4.792pJ \quad (3.10)$$

3.2.2. 6-bit asynchronous SAR ADC (1bit/cycle)

A closer look at the previous architecture raises an important question – since the minimum input voltage linked to the slowest comparison time does not occur in every cycle, can the speed of the ADC be improved if the actual conversion time is taken into account?

The answer can be found by turning to asynchronous processing, a technique which allows the conversion to progress as soon as each block has completed its task, rather than prescribing a fixed internal clock period [29], [30]. It is important to notice that although internally the ADC operates asynchronously, the sampling rate used is uniform. Although this technique invites a variable sampling rate, the possibilities and limitations of such an ADC lie outside the scope of this work and will not be discussed here. Adopting a non-uniform sampling rate creates difficulties in the design of the subsequent as well as the preceding blocks of the ADC and is unsuitable for time-interleaving due to the timing uncertainties that it introduces.

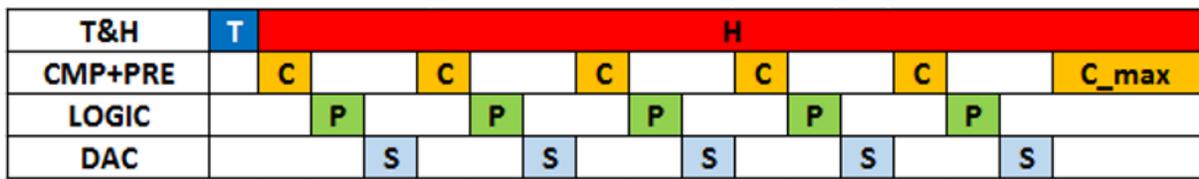


Figure 3.2.2.1 - Timing diagram of a 6-bit asynchronous SAR ADC (1bit/cycle)

Figure 3.2.2.1 shows the timing diagram of an asynchronous 6-bit SAR ADC. Compared to the synchronous version (Figure 3.2.1.2), the maximum comparison time occurs only once per conversion [29] and the other comparison delays need not be equal. Because there is no fixed clock period, the conversion can end after value of the last bit has been determined, without waiting for the logic to propagate and the DAC to settle. These two operations would be useless anyway, seeing that no new reference has to be prepared after the 6th bit is known.

Again, it is assumed that the DAC can be reset during the tracking phase and that there is sufficient time to reset the comparator while the result propagates through the logic and the DAC settles. However, the comparator reset time needs to be shorter than in the synchronous design because the trigger for the next comparison may arrive earlier than the internal clock period would have imposed. A trade-off between accuracy and speed is already visible: depending on the implementation of the logic, the conversion time might need to be increased to accommodate the comparator reset time, in case this operation cannot be performed in the available time. Supposing that this is not an issue, the conversion time is given by:

$$t_{conv} = t_{track} + \sum_1^N t_{comp} + (N - 1)(t_{DAC} + t_{logic}) \quad (3.11)$$

In order to assess the delay for each comparison, it is assumed that the smallest input voltage is of $\frac{V_{LSB}}{4}$ and that all the other inputs are twice their predecessor. In other words, $\Delta V_{in} = [8V_{LSB}, 4V_{LSB}, 2V_{LSB}, V_{LSB}, \frac{1}{2}V_{LSB}, \frac{1}{4}V_{LSB}]$ and $\sum_1^N t_{comp}$ is the sum of the corresponding comparator delays, which is equal to 438.6ps. The DAC settling time will also vary in the asynchronous design, but for the sake of simplicity it is assumed to be equal to the MSB settling for all bits.

The capacitances seen by the track-and-hold circuit and the DAC remain the same as in the synchronous version, so a tracking time of 50ps and a DAC settling time of 24ps also apply here.

Adding the delays points to a minimum conversion time of roughly $t_{conv} = 1ns$, which corresponds to a sampling frequency of $f_{conv} = 1GHz$. An increase of 200MHz is already visible compared to the previous architecture. The energy consumption is also reduced because the last cycle avoids activating the logic and the DAC again:

$$P_{total} = f \left(6 \left(E_{comp_{pre}} + E_{clk} \right) + 5 \left(E_{DAC} + E_{logic} \right) + E_{T\&H} \right) = 4.4mW \quad (3.12)$$

The overall efficiency of the architecture is slightly higher than for the synchronous case, as shown by the reduced figure of merit:

$$FOM = \frac{4.4mW}{1GHz} = 4.414pJ \quad (3.13)$$

According to Chen & Brodersen [29], the worst case improvement of an asynchronous design compared to a synchronous one (computed for an input signal that changes its polarity from one comparison to the next) is given by:

$$\frac{T_{asynch}}{T_{synch}} = \frac{(N - 1)ln3 + ln2 + \frac{N}{2}(N + 1)ln2}{N(N + 1)ln2} = 0.713 \text{ if } N = 6 \text{ bits} \quad (3.14)$$

This implies that if $T_{synch} = 1.238ns$, as was calculated in the previous section, then $T_{asynch} = 0.882ns$. Comparing this with the 1ns obtained for the computed T_{asynch} demonstrates that the assumptions made are slightly pessimistic because it was assumed

that the input signal halved from one comparison to the next, whereas the work in [29] assumes a uniform distribution of the input signal.

3.2.3. 2-bit asynchronous SAR ADC (1bit/cycle)

A 2-bit SAR ADC might not appear practical at first, but if it is part of a pipelined ADC, it might prove to be beneficial for the speed. SAR ADCs are particularly interesting for inclusion in a pipelined ADC because they already contain a DAC which is required for determining the residue. Because of this, calculating its minimum conversion time is a useful exercise and Figure 3.2.3.1 is used for this purpose.

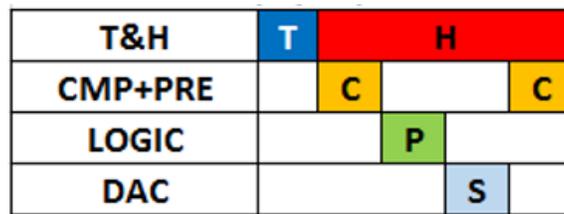


Figure 3.2.3.1 – Timing diagram of a 2-bit asynchronous SAR ADC (1bit/cycle)

The conversion time can be expressed as:

$$t_{conv} = t_{track} + 2t_{comp} + t_{DAC} + t_{logic} \quad (3.15)$$

The DAC settling time and the digital logic delay appear only once because of the asynchronous operation. It is important to note that in this case, choosing an asynchronous design does not yield significant speed savings because only one of the two comparisons will be performed for a large input signal. Because the difference is rather small, it is assumed that both comparisons take the same amount of time (maximum delay).

Depending on the implementation of the internal clock signals and how they are generated, it might turn out that the delays required to trigger the next comparator are longer than the time saved, making an asynchronous operation less appealing. For the scope of this architecture study it is assumed that this is not the case, which leads to a conversion time equal to $t_{conv} = 342\text{ps}$, which corresponds to an operating frequency of $f_{conv} = 2.923\text{GHz}$. At this frequency, the power consumption is given by:

$$P_{total} = f \left(E_{T\&H} + 2 \left(E_{cmp_{pre}} + E_{clk} \right) + E_{logic} + E_{DAC} \right) = 3.651\text{mW}, \quad (3.16)$$

which corresponds to a figure-of-merit of:

$$FoM = \frac{3.651\text{mW}}{2.923\text{GHz}} = 1.249\text{pJ} \quad (3.17)$$

The 2-bit SAR ADC is very efficient, which confirms that it would be a good candidate for inclusion in a pipelined ADC.

3.2.4. 6-bit asynchronous SAR ADC (3bits/cycle)

The previous design studied proves that speeds of over 2 GHz are possible with the SAR ADC (typically regarded as a medium-resolution ADC) as long as the number of comparator delays in the critical path remains low. This prompts the next architecture idea – using a 3-bit/cycle approach to keep the same critical path length as the efficient 2-bit asynchronous SAR, albeit compromising slightly on the efficiency due to the increased number of components required. A similar technique is applied successfully by [31].

Figure 3.2.4.1 shows the architecture modifications required to achieve the target critical path. There are two sets of comparators that operate alternately. The first set (light pink) compares the sampled input voltage with the first set of references provided by 7 DACs to quantize the signal with a 3-bit accuracy. Afterwards, the SAR logic takes the switches the DAC in order to generate the set of references for the second set of comparators. After the last set of decisions is complete (dark pink), the bits can be fed to the output and a new conversion can take place. The timing diagram (Figure 3.2.4.2) is identical to the one of the 2-bit SAR and it is assumed that the same logic delay is present.

Despite the striking similarity with the previous architecture, the input capacitance seen by the track-and-hold is much larger in this case, which means that the requirements for this input stage are more stringent. The DAC also sees twice the input capacitance of one comparator, but this speed penalty can be circumvented if 14 DACs are used instead of only 7. However, because the DAC settling time is relatively small, 7 blocks seem sufficient.

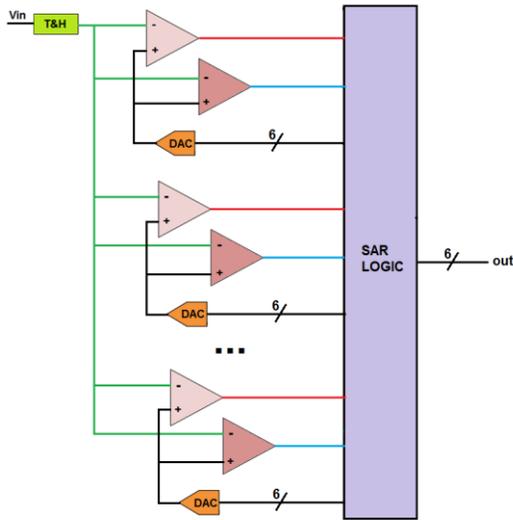


Figure 3.2.4.1 - Block diagram of a 6-bit asynchronous SAR ADC (3bits/cycle)

| | | | |
|---------|---|---|---|
| T&H | T | H | |
| CMP+PRE | | C | C |
| LOGIC | | P | |
| DAC | | | S |

Figure 3.2.4.2 – Timing diagram of a 6-bit asynchronous SAR ADC (3bits/cycle)

The conversion time is given by:

$$t_{conv} = t_{track} + 2(t_{comp_{LSB}} + t_{preamp}) + t_{DAC_{MSB}} + t_{SAR_{logic}} \quad (3.18)$$

The tracking time is 55ps if an $R_{on} = 50\Omega$ is assumed for the sampling switch, while the DAC settling has increased to 30ps, leading to a total conversion delay of $t_{conv} = 0.353ns$, which corresponds to a sampling frequency of $f_{conv} = 2.83GHz$. A smaller on-resistance is considered in this case because the tracking time of 50ps is desirable in order to maintain a high-speed operation. The high input capacitance of this architecture requires a lower on-resistance was for the sampling switch if a 50ps tracking time is aimed for.

The power dissipation and figure-of-merit further characterize this solution:

$$P_{total} = f \left(2E_{T\&H} + 14 \left(E_{cmp_{pre}} + E_{clk} \right) + 2E_{logic} + 7E_{DAC} \right) = 19.3mW \quad (3.19)$$

$$FoM = \frac{19.3mW}{2.83GHz} = 6.818pJ \quad (3.20)$$

This is the first 6-bit architecture that satisfies the speed requirement of the envisioned application. This can be explained by the fact that the individual stages are 3-bit flash ADCs, which are known for their high power consumption. Compared to a 6-bit flash ADC (which requires 63 comparators, as we will see in the next section), this one requires only 14, which

translates to a better area vs. speed optimization as well as a better power vs. speed (FoM) ratio.

3.2.5. 6-bit flash ADC

A quest for a fast ADC could not leave the epitome of converter speed untried. Due to its very short critical path – comprised only of the track-and-hold, a comparator and a latch as Figure 3.2.5.1 illustrates – the flash ADC seems the most promising architecture in terms of achieving a sampling frequency exceeding 2GHz [16], [17], [15].

The principle of the flash ADC requires $2^n - 1$ voltage references that help determine all the output bits in parallel [32]. For this operation, $2^n - 1$ comparators are needed, which makes the area and power consumption prohibitive to some applications.

The timing diagram of the flash ADC is very straightforward (Figure 3.2.5.2) and for the scope of this calculation, does not include any delays pertaining to reference generation. The minimum conversion time can be thus written as:

$$t_{conv} = t_{track} + t_{comp_{pre}} + t_{latch} \quad (3.21)$$

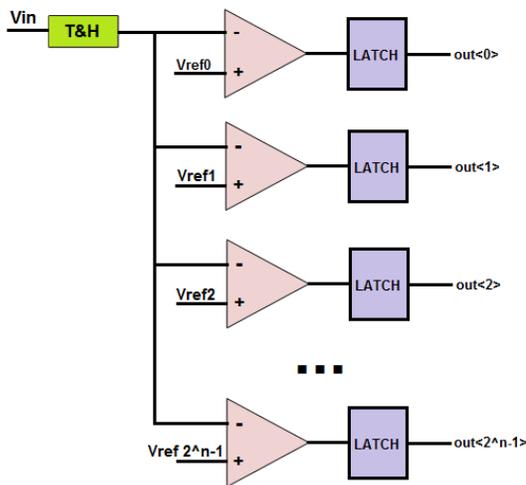


Figure 3.2.5.1 - Block diagram of a 6-bit flash ADC

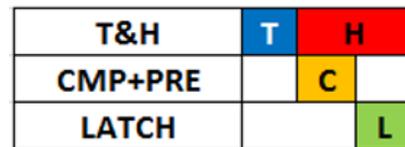


Figure 3.2.5.2 – Timing diagram of a 6-bit flash ADC

The tracking time is dependent on the input capacitance of the 63 comparators in the design and reaches a value of 100ps if $R_{on} = 50\Omega$ (the same consideration mentioned previously applies), leading to a conversion time of only $t_{conv} = 223ps$. This means that the maximum operating frequency is of $f_{conv} = 4.46GHz$ and the corresponding power consumption and FoM are:

$$P_{total} = f \left(2E_{T\&H} + 2^{N-1} \left(E_{cmp_{pre}} + E_{clk} + E_{latch} \right) \right) = 67.97mW \quad (3.22)$$

$$FoM = \frac{67.97mW}{4.46GHz} = 15.24pJ \quad (3.23)$$

The actual figure of merit will be slightly lower because the power dissipation of the resistive ladder that is assumed to generate the voltage references was not taken into account here. It can be concluded that even though the flash ADC is almost twice as faster as the promising SAR configuration studied previously, the power and area penalty make it the least attractive topology.

3.2.6. Pipeline of 4-bit asynchronous SAR ADC (2bits/cycle) and 2-bit flash ADC

The pipeline ADC is another attractive topology for high-speed applications because it trades latency for speed [33]. By increasing area, the same function can be split into two or more processing units such that more samples can be processed at the same time, allowing a higher sampling speed at its input. In order to accommodate this, additional track-and-hold circuits are needed in front of each sequential block.

The most straightforward pipeline design is limited to two stages, which divides the conversion into a step that determines the MSBs and one that determines the LSBs. Before passing the output to the second stage, a residue needs to be computed, which generally implies that a digital-to-analog conversion needs to take place to eliminate the MSBs from the input signal.

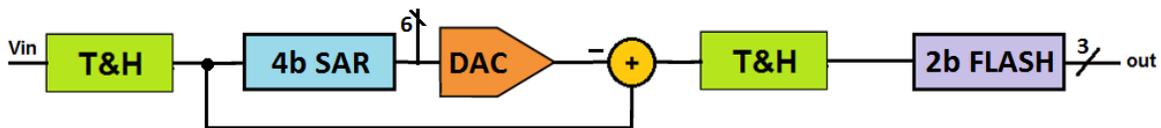


Figure 3.2.6.1 - Block diagram of a pipeline of 4-bit asynchronous SAR ADC (2bits/cycle) and 2-bit flash ADC (with interpolation)

Bearing this in mind, it seems like an efficient approach to perform this subtraction using a DAC that is already part of the first ADC, which motivates the choice for the next architecture: a pipeline between a 4-bit SAR ADC and a 2-bit flash ADC, whose block diagram can be found in Figure 3.2.6.1. The DAC depicted is not necessarily a separate block as a 4-bit DAC is already part of the 4-bit SAR, but has been represented in order to remain true to the operating principle. The timing diagram in Figure 3.2.6.2 reveals that the first stage is the

one limiting the maximum sampling frequency, which is also slightly larger than that of an independent 4-bit SAR ADC because it needs to include the residue computation and an additional tracking phase. In order to maximize the speed of the loop, a 2-bit/cycle approach was chosen.

| | | | | | | |
|-----------|---|---|---|---|-----|--|
| T&H 1 | T | H | | | | |
| CMP+PRE 1 | | C | | C | | |
| LOGIC | | P | | P | | |
| DAC | | | S | | S | |
| T&H 2 | | | | | T H | |
| CMP+PRE 2 | | | | | C | |
| LATCH | | | | | L | |

Figure 3.2.6.2 – Timing diagram of a pipeline of 4-bit asynchronous SAR ADC (2bits/cycle) and 2-bit flash ADC

Figure 3.2.6.3 and Figure 3.2.6.4 represent the block diagrams of the sub-ADCs and reveal that a tracking time of 50ps (corresponding to the minimum pulse width) is sufficient for both stages. An MSB settling time of 30ps renders the desired accuracy of 6-bits.

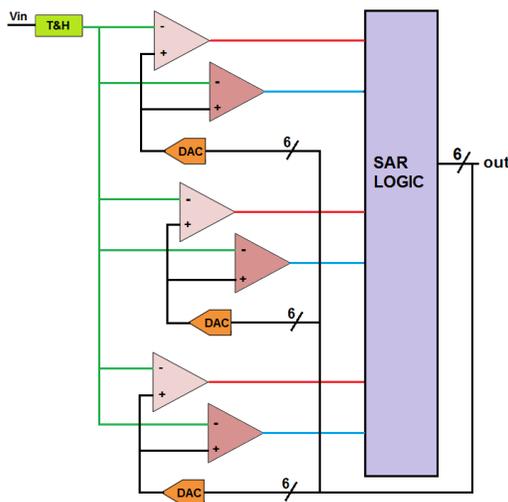


Figure 3.2.6.3 - Block diagram of a 4-bit asynchronous SAR ADC (2bits/cycle)

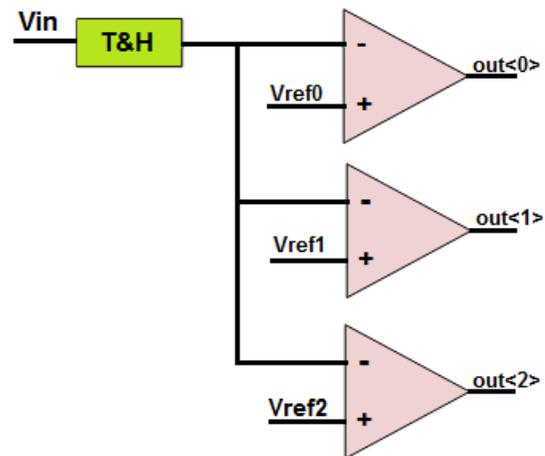


Figure 3.2.6.4 - Block diagram of a 2-bit flash ADC

It follows that the delay of the SAR, which gives the speed of the ADC, is given by:

$$t_{SAR} = t_{track1} + t_{track2} + 2 \left(t_{comp_{LSB}} + t_{preamp} + t_{DAC_{MSB}} + t_{SAR_{logic}} \right) \quad (3.24)$$

The calculation reveals $t_{SAR} = 0.636\text{ns}$ which implies a maximum sampling frequency of $f_{sample} = 1.572\text{GHz}$. At this frequency, the power consumption is of $P_{total} = 6.64\text{mW}$, with the SAR consuming 76% of this amount and the flash accounting for the rest. The figure of merit for this design is comparable to that of the asynchronous 6-bit SAR and reaches a value of $FoM = 4.22\text{pJ}$.

A general frequency-dependent expression of the power details the contribution of each stage:

$$P_{total} = f \left(6E_{cmp_{pre}} + 6E_{clk} + 2E_{logic} + 2E_{T\&H} \right) + f \left(2E_{cmp_{pre}} + 2E_{clk} + 3E_{latch} \right) \quad (3.25)$$

An important conclusion can be drawn from this architecture: the first stage in a pipelined ADC will always include two tracking times and an extra delay due to residue generation. Because of this, it makes sense to include the slowest sub-ADC in the second stage, as this would provide the greatest speed benefit.

3.2.7. Pipeline of 2-bit flash ADC and 4-bit asynchronous SAR ADC (2bits/cycle)

Simply changing the order of the two ADCs in the previous design is expected to give a substantial speed improvement. The only drawback of including the 2-bit flash in the first stage is that an additional DAC is required to perform the subtraction between the input signal and the signal reconstructed from the MSBs. Since this is a 2-bit DAC, the area and power penalty are marginal. Figure 3.2.7.1 shows the updated block diagram.

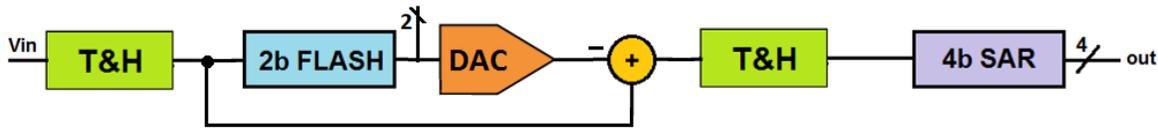


Figure 3.2.7.1 - Block diagram of a pipeline of 2-bit flash ADC and 4-bit asynchronous SAR ADC (2bits/cycle)

The timing of the circuit has changed slightly compared to the previous design. First, the SAR delay no longer needs to include the logic delay and the DAC settling for its second stage, because it does not need to compute the residue. The tracking time of the SAR, together with the extra DAC settling time are now part of the delay of the first stage which can easily accommodate them because of the fast 2-bit flash. The speed will, in this case, be limited by the flash stage, but is higher than that in the previous architecture because the individual stage delays are better balanced.

| | | | | |
|-----------|---|---|---|---|
| T&H 1 | T | H | | |
| CMP+PRE 1 | | C | | |
| LATCH | | L | | |
| DAC 1 | | | S | |
| T&H 2 | | | T | H |
| CMP+PRE 2 | | | C | C |
| LOGIC | | | P | |
| DAC 2 | | | S | |

Figure 3.2.7.2 – Timing diagram of a pipeline of 2-bit flash ADC and 4-bit asynchronous SAR ADC (2bits/cycle)

The minimum clock period is given by:

$$t_{clk} = t_{track1} + t_{comp+pre} + t_{latch} + t_{DAC1} + t_{track2} = 379ps \quad (3.26)$$

It follows that this ADC can operate at a frequency of $f_{clk} = 2.63GHz$, which leads to a power consumption of:

$$P_{total} = f \left(8E_{cmppre} + 8E_{clk} + 2E_{logic} + 2E_{T\&H} + 3E_{latch} \right) = 11.109mW \quad (3.27)$$

Since essentially the blocks involved in the conversion (and the number of times they are used) is the same, the energy per conversion will remain unchanged, rendering the same figure-of-merit as before: $FoM = 4.22pJ$.

3.2.8. Pipeline of 2x 3-bit SAR ADCs

The idea of reusing the same ADC for the pipeline stages has the potential to converge to a functional ADC with less design time and effort. It is thus worthwhile to explore this possibility when a power-efficient sub-ADC is employed.

In what follows, the performance of a pipeline ADC consisting of two 3-bit asynchronous SAR stages will be surveyed and compared to the pervious architectures considered. Figure 3.2.8.1 shows the block diagram pertaining to this design. As was the case with Figure 3.2.6.1, the DAC in the first stage is represented for algorithm-visualization purposes and is, in fact, included in the 3-bit SAR.

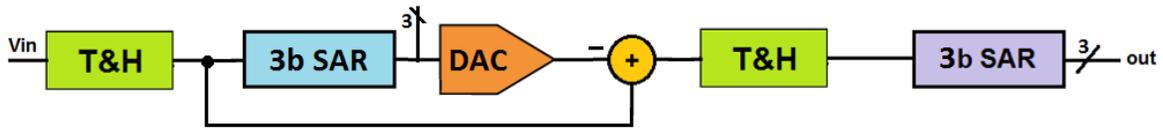


Figure 3.2.8.1 - Block diagram of a pipeline of 2x 3-bit SAR ADCs

The speed of this design will be lower than that of a single 3-bit SAR because the critical path includes an extra tracking time, as well as an extra logic delay and DAC settling. This is evidenced by the timing diagram in Figure 3.2.8.2.

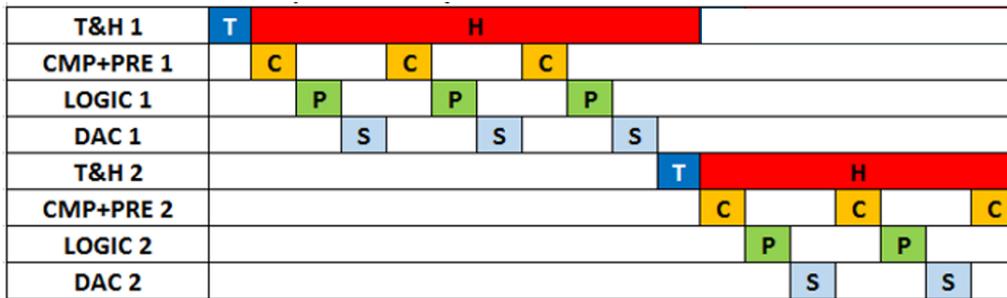


Figure 3.2.8.2 – Timing diagram of a pipeline of 2x 3-bit SAR ADCs

Having identified the critical path, the minimum clock period can be determined from:

$$t_{clk} = 2t_{track} + 3(t_{comp_{pre}} + t_{logic} + t_{DAC}) \tag{3.28}$$

It follows that this architecture can run at $f_{clk} = 1.252\text{GHz}$, while spending a power of:

$$P_{total} = f(6E_{cmp_{pre}} + 6E_{clk} + 5E_{logic} + 5E_{DAC} + 2E_{T\&H}) = 5.593\text{mW} \tag{3.29}$$

The efficiency of the design is approximately that of a single 6-bit asynchronous SAR ADC ($FoM = 4.461\text{p}$). This can be explained by the fact that using an internal asynchronous clock only yields significant speed improvements for high-resolution ADCs. The formula of Chen & Brodersen [29] described previously shows that for a 3-bit SAR, an asynchronous design is only 15% faster than its synchronous counterpart.

3.2.9. Pipeline of 3x 2-bit SAR ADCs

The previous pipeline idea revealed that a 3-bit SAR is too slow to allow a sampling speed exceeding 2GS/s. A different way to reuse the same ADC while achieving a slower delay could be the use of 2-bit SARs for the pipeline stages. It was previously shown that this sub-ADC has an excellent efficiency, which further supports its incorporation into a pipelined design. The block diagram of such a converter is shown in Figure 3.2.9.1. The two DACs shown are part of the SARs, but have been represented in order to illustrate the algorithm correctly.



Figure 3.2.9.1 - Block diagram of a pipeline of 3x 2-bit SAR ADCs

The corresponding timing diagram (Figure 3.2.9.2) shows that the number of comparisons per stage, as well as the logic delay was reduced by 33% compared to the previous case.

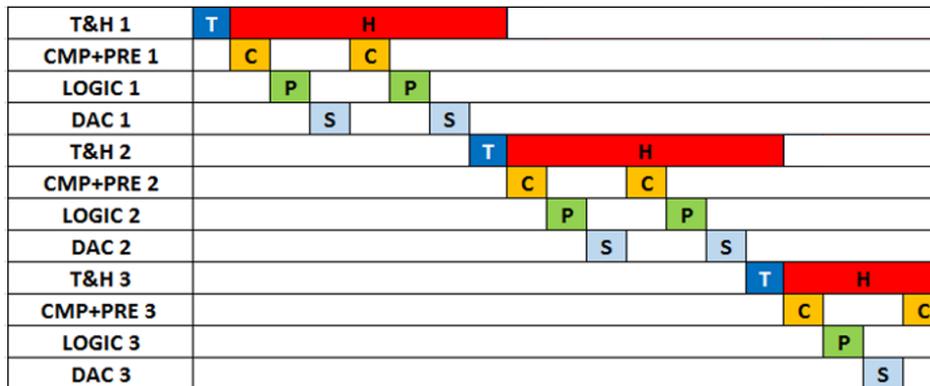


Figure 3.2.9.2 – Timing diagram of a pipeline of 3x 2-bit SAR ADCs

It thus follows that the minimum clock period and power consumption are given by:

$$t_{clk} = 2t_{track} + 2(t_{comp} + t_{DAC} + t_{logic}) \quad (3.30)$$

$$P_{total} = f * (3E_{T\&H} + 6E_{cmp} + 5E_{logic} + 6E_{clk} + 5E_{DAC}) \quad (3.31)$$

This architecture achieves a sampling speed of $f_{clk} = 2.016\text{GHz}$, at a power consumption of 9.097mW . The corresponding figure of merit is comparable to that of other SAR-based converters: $FoM = 4.51\text{pJ}$.

3.2.10. Pipeline of 2x 3-bit flash ADCs

Because the two previous architectures, although fairly efficient, are deemed too slow for the application envisioned, a pipeline between 2 3-bit flash ADCs is considered. The block diagram (Figure 3.2.10.1) illustrates the operating principle. As there is no DAC in the first stage that can be used, an explicit 3-bit DAC needs to be included.

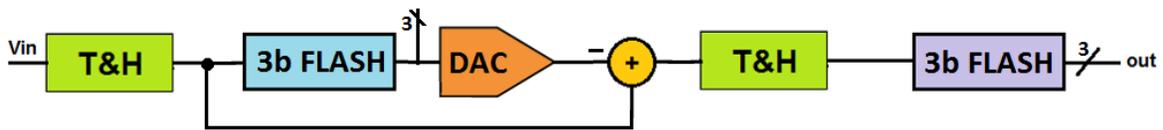


Figure 3.2.10.1 - Block diagram of a pipeline of 2x 3-bit flash ADCs

The timing diagram (Figure 3.2.10.2) shows that the critical path is slightly longer than in the case of a 3-bit flash ADC as it includes the DAC settling time needed to compute the residue and an additional tracking phase that allows the signal to be passed to the next stage.

| | | | | | |
|-----------|---|---|---|---|---|
| T&H 1 | T | H | | | |
| CMP+PRE 1 | | C | | | |
| LATCH 1 | | | L | | |
| DAC 1 | | | S | | |
| T&H 2 | | | T | H | |
| CMP+PRE 2 | | | C | | C |
| LATCH 2 | | | | L | L |

Figure 3.2.10.2 – Timing diagram of a pipeline of 2x 3-bit flash ADCs

A clock frequency of $f_{clk} = 3.436\text{GHz}$ is possible and limited by:

$$t_{clk} = 2t_{track} + t_{comp_{pre}} + t_{DAC_{MSB}} + t_{latch} \tag{3.32}$$

At this speed, the expected power consumption is 23.282mW, as detailed by the equation below. The inefficiency of the flash ADCs also impacts the figure-of-merit, limiting it to $FoM = 6.77pJ$.

$$P_{total} = f(2E_{T\&H} + 14E_{cmp_{pre}} + 14E_{clk} + 14E_{latch} + E_{DAC}) \quad (3.33)$$

3.3. Architectures with inter-stage amplifier

The pipelined ADCs that have been surveyed up to now assume that the residue that is generated after the conversion of the MSBs can be processed by the subsequent stages directly. This is true only if the performance requirements of all stages are the same as those for the first ADC. This owes mostly to the fact that noise levels remain the same, while the signal amplitude is decreased substantially (for example, in the case of a two 3-bit ADCs it is halved) which makes it more difficult to maintain the same SNR.

This problem can be approached in a different fashion if residue amplification is taken into account. Inserting an amplifier with a gain equal to 2^m between pipelined stages (where m represents the number of bits resolved by the previous stage), the SNR can be more easily maintained because the signal amplitude for the next stage will use the available headroom better than when no amplifier is present. The main drawback of opting for an inter-stage gain is that it requires a very accurate value of the amplification while maintaining a high bandwidth. Violating the first constraint implies a loss of conversion accuracy and can be compensated only if redundancy bits are incorporated into the system, which is a fairly accessible solution. However, failing to accommodate the second requirement incurs an undesired speed penalty and is a concern much harder to address in practice.

To assess the performance of a pipelined ADC with inter-stage gain and compare it to the previous designs analyzed, a fast amplifier topology is chosen to illustrate the achievable speed and power consumption level. In what follows, the characteristics of this amplifier will be described and then the results pertaining to architectures that employ this block are reported.

3.3.1. Architecture of a fast inter-stage amplifier

Astgimath [34] describes the principle and implementation of a high speed, low power amplifier designed especially for inclusion in a pipelined ADC. The circuit is based on an integrator and consumes no static power and is further referred to as a *cascoded integrator dynamic residue amplifier* (CIDRA). An advantage of the design is that it is a low-noise solution, but for a 6-bit design, which is not limited by thermal, but by quantization noise, this is a less interesting property.

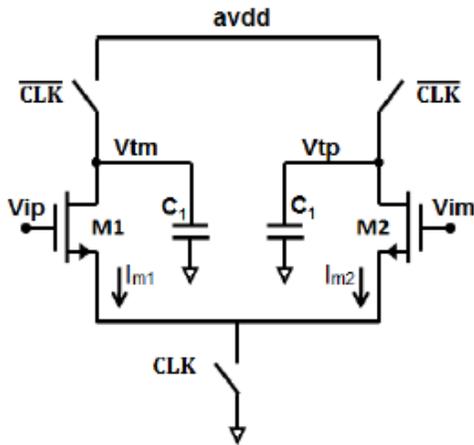


Figure 3.3.1.1 – Single-stage integrator dynamic residue amplifier, image courtesy of [34]

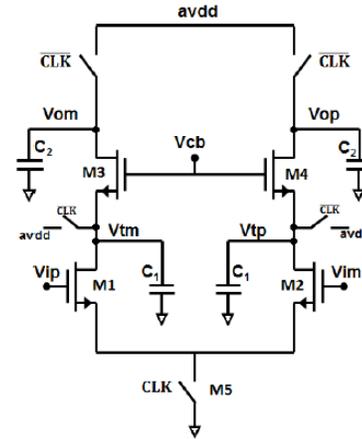


Figure 3.3.1.2 – Cascoded integrator dynamic residue amplifier (CIDRA), image courtesy of [34]

The simplest version of the circuit is shown in Figure 3.3.1.1 and represents the single-stage realization of the CIDRA [34]. Let us first assume that the clock signal is low; the amplifier is reset, while the capacitors are charged to the supply voltage. This phase prevents errors from previous clock cycles to occur by ensuring a known state for the outputs. When the clock becomes high, the sources of the transistors in the input pair are pulled to ground, which turns them on, leading to a slow discharge of the output capacitors. The rate at which the output nodes are discharged depends on the input signal, with a higher rate associated to a bigger input signal. The output needs to be stored before the capacitors discharge completely, causing the input pair to enter the triode region.

The gain of the amplifier depends on the integration time and is given by [34]:

$$A_0 = \frac{T_{int1}g_{m1}}{2C_1} = \frac{V_{ocm}}{V_{gt}} \quad (3.34)$$

The equation shows that the amplification can be increased by lowering the overdrive voltage of the transistors and increasing the output common mode. It is thus beneficial to size the input pair such that it operates in weak inversion. The design reported in the thesis, [34], uses a supply voltage of 1V to achieve a gain of 6.25 in case the transistors are in weak inversion and the output common mode voltage is half that of the supply [34]. This amplification is useful in a pipelined ADC only if the preceding stage resolves at most 2 bits, which prompts the cascoded design with higher gain in Figure 3.3.1.2.

The operating principle is identical to the one described before, except for the fact that two different integrations occur, owing to the two integrators connected in series.

In this case, the gain is given by [34]:

$$A_0 = \frac{g_{m1}}{2C_2} (T_{int1} + T_{int2}) = \frac{V_{tn}}{V_{gt}} \left(1 + \frac{C_1}{C_2} \right) \quad (3.35)$$

Despite the first term being limited by technology and operating region, the second provides enough flexibility to implement a large gain. The main disadvantage of the circuit is also obvious from the formula – as the threshold voltage of the transistor will change due to process variations, the accuracy of the gain cannot be guaranteed, and so redundant bits are unavoidable in the pipelined ADC.

3.3.2. Pipeline of 2-bit flash ADC and 5-bit SAR ADC (1 bit of redundancy)

The most promising pipelined ADC from earlier paragraphs is analyzed in case a high-speed gain stage is present between stages. The modified block diagram is depicted in Figure 3.3.2.1 and features a 5-bit SAR, which is intended to have one redundancy bit that can compensate for the inaccuracy of the inter-stage gain. Similar to the previous case, the SAR implementation is a 2-bit/cycle approach.

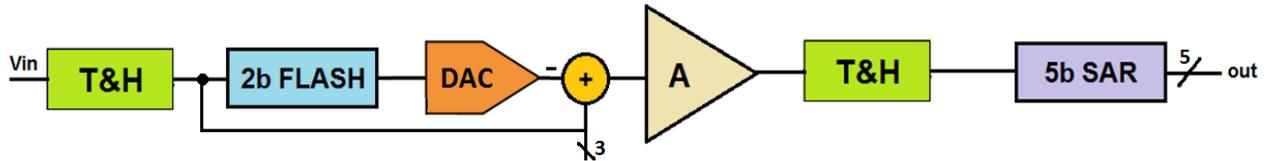


Figure 3.3.2.1 - Block diagram of a pipeline of 2-bit flash ADC and 5-bit SAR ADC (1 bit of redundancy)

The timing diagram is adjusted to account for the amplification time (Figure 3.3.2.2). The position of this delay is not limited to that shown in the picture, as the signal can be amplified during the processing done in the second stage. However, because the 2-bit flash ADC is the fastest stage in this case, the amplification time can easily fit in the available time difference.

| | | | | |
|-----------|---|---|---|---|
| T&H 1 | T | H | | |
| CMP+PRE 1 | | C | | |
| LATCH | | | P | |
| DAC 1 | | | | S |
| AMP | | | | A |
| T&H 2 | | | T | H |
| CMP+PRE 2 | | | C | C |
| LOGIC | | | | P |
| DAC 2 | | | | S |

Figure 3.3.2.2 – Timing diagram of a pipeline of 2-bit flash ADC and 5-bit SAR ADC (1 bit of redundancy)

Supposing that the input capacitance of the gain stage is kept the same as that of the first ADC (24fF) and that a gain of 4x is desired, a set of possible values for sizing the CIDRA is listed below:

$$\begin{cases} C_1 = C_2 = 6C_{inpre} = 72\text{fF} \\ V_{gt} = 125\text{mV} \end{cases} \quad (3.36)$$

The noise is much lower than half an LSB even with $C_1 = C_2 = 15\text{fF}$, so the current through the preamplifier is imposed by the delay and not the noise specification. Supposing that Δt_2 is the time difference between the delays of the two stages (around 140ps in this case), it follows that the current through the gain stage is given by:

$$I = \frac{V_T}{\Delta t_2} (C_1 + C_2) = 56.79\mu\text{A} \quad (3.37)$$

This current corresponds to a dynamic power consumption of:

$$P_{amp} = I * V_{DD} = 0.056\text{mA} * 1.1\text{V} = 62.469\mu\text{W} \quad (3.38)$$

The speed of the overall ADC is once again limited by the SAR stage:

$$t_{clk} = t_{SAR} = 2(t_{comp2} + t_{DAC2} + t_{logic}) = 380\text{ps} \quad (3.39)$$

The circuit operates at a maximum frequency of $f_{clk} = 2.63\text{GHz}$, and assuming that there are no preamplifiers in the second stage (modelled by accounting for half of the power of the same SAR with preamplifiers), consumes a total power of:

$$P_{total} = f(5E_{comp+pre} + 5E_{clk} + 4E_{DAC} + 3E_{logic} + 2E_{T\&H}) + P_{amp} = 6.358\text{mW} \quad (3.40)$$

Despite having a lower speed than the equivalent design without an inter-stage gain, the efficiency is much higher, owing to the relaxed specifications on the second ADC:

$$FoM = \frac{6.358mW}{2.63GHz} = 2.42pJ \tag{3.41}$$

3.3.3. Pipeline of 3-bit flash ADC and 3-bit flash ADC (1 bit of redundancy)

The speed of the previous circuit was limited due to the slow second stage, so a final design, featuring a faster second ADC is surveyed. The block diagram of a pipeline between a 3-bit flash and a 4-bit flash ADC is shown in Figure 3.3.3.1.

One bit of redundancy is included in the second stage to compensate for gain inaccuracy.

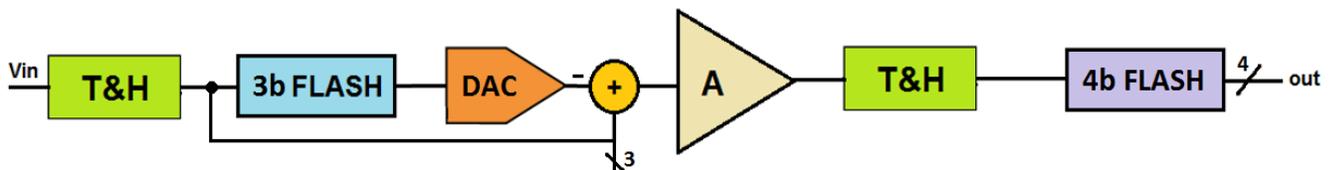


Figure 3.3.3.1 - Block diagram of a pipeline of 3-bit flash ADC and 3-bit flash ADC (1 bit of redundancy)

Because of the extra tracking time included in the delay of the first stage, the amplifier delay is assumed to be part of the second stage in this case, as the timing diagram in Figure 3.3.3.2 shows.

| | | | | |
|------------------|----------|----------|----------|----------|
| T&H 1 | T | H | | |
| CMP+PRE 1 | | C | | |
| LATCH 1 | | | P | |
| DAC 1 | | | S | |
| AMP | | | | A |
| T&H 2 | | T | H | |
| CMP+PRE 2 | | | C | |
| LATCH 2 | | | | P |

Figure 3.3.3.2 – Timing diagram of a pipeline of 3-bit flash ADC and 3-bit flash ADC (1 bit of redundancy)

A gain of 8 needs to be realized and the goal is within reach if the following values are adopted:

$$\begin{cases} C_1 = C_2 = C_{in_{ADC2}} = 96\text{fF} \\ V_{gt} = 62.5\text{mV} \end{cases} \quad (3.42)$$

Together with an amplification time of 140ps, a current of 322.1 μA will flow through the amplifier, which translates to the meagre power consumption of 0.354mW.

The speed of the pipeline is limited by the first stage:

$$t_{clk} = t_1 = t_{s1} + t_{cmp+pre} + t_{latch} + t_{DAC1} + t_{s2} \quad (3.43)$$

Considering that the second stage does not require the use of preamplifiers, modelled by assuming only 40% of the power consumption of the same flash ADC with preamplifiers, it follows that the design runs at maximum 3.86GHz and consumes a total of:

$$P_{total} = f(7.2E_{comp+pre} + 7.2E_{clk} + E_{DAC} + 7.2E_{latch} + 2E_{T\&H}) + P_{amp} = 17.647\text{mW} \quad (3.44)$$

The figure of merit is comparable to that of a synchronous SAR, with a value of $FoM = 4.578\text{pJ}$.

3.4. Conclusions

Reviewing the designs considered throughout the chapter leads to the performance summary in Table 3.4.1. The designs in bold represent architectures that are capable of achieving a sampling rate of over 2.5GHz with a single-channel ADC with an accuracy of 6 bits. The power consumption of the designs can be further improved by using interpolation, which has the potential to reduce the number of preamplifiers by half. This is true for multi-bit/cycle and flash ADCs, so a comparison with this design choice is excluded from the present study, as it is assumed that it does not affect the ability to compare circuits and can be applied later, if needed.

The 6-bit flash ADC is too inefficient to be considered for implementation in a time-interleaved ADC, as it would amount to roughly 340mW with a 5x interleaving factor. The pipelined ADCs with inter-stage gain, although very efficient, are also unattractive because of the design complications that the gain stage entails, especially that of compensating for its inaccuracies. From the three promising architectures left, the most efficient appears to be the pipeline between a 2-bit flash and a 4-bit SAR. However, the expected sampling rate marginally exceeds the target specification, which makes it difficult to achieve the desired performance in practice, not least because of the non-idealities that have not been taken into account during this phase.

Chapter 3 – Architecture Study

It follows that the natural choice is a 3bit/cycle asynchronous SAR ADC, which leaves a margin of over 300MHz to account for layout parasitic elements and other limitations not taken into account. With a target linearity of 4.5bits, the corresponding Walden figure-of-merit is expected to be around 301fF/conv-step. The chapters that follow will discuss the challenges and performance achieved by a 6-bit ADC using this architecture.

Table 3.4.1 - Summary of the performance of the architectures surveyed

| ARCHITECTURE | SAMPLING RATE | POWER | FOM |
|--|-----------------|-----------------|----------------|
| Synchronous 6-bit SAR 1b/cycle | 807MHz | 3.867mW | 4.792pJ |
| Asynchronous 6-bit SAR 1b/cycle | 1GHz | 4.4mW | 4.414pJ |
| Asynchronous 2-bit SAR 1b/cycle | 2.923GHz | 3.69mW | 1.249pJ |
| Asynchronous 6-bit SAR 3b/cycle | 2.83GHz | 19.3mW | 6.818pJ |
| 6-bit flash | 4.46GHz | 67.97mW | 15.24pJ |
| Pipeline of 4-bit SAR (2b/cycle, asynchronous) and 2-bit flash | 1.572GHz | 6.64mW | 4.22pJ |
| Pipeline of 2-bit flash and 4-bit SAR (asynchronous, 2b/cycle) | 2.636GHz | 11.109mW | 4.22pJ |
| Pipeline of 2x 3-bit SAR ADC | 1.254GHz | 5.59mW | 4.46pJ |
| Pipeline of 3x 2-bit SAR ADC | 2.016GHz | 5mW | 4.51pJ |
| Pipeline of 2x 3-bit flash | 3.436GHz | 23.282mW | 6.77pJ |
| Pipeline of 2-bit flash and 5-bit SAR (1b redundancy), with inter-stage gain | 2.63GHz | 6.385mW | 2.426pJ |
| Pipeline of 3-bit flash and 4-bit flash (1b redundancy) with inter-stage gain | 3.86GHz | 17.647mW | 4.578pJ |

The frequency dependence of the power consumption associated to each ADC considered is summarized in Table 3.4.2 for completeness.

Table 3.4.2 – Frequency dependence of the power consumption of the surveyed architectures

| ARCHITECTURE | DEPENDENCE |
|---|---|
| Synchronous 6-bit SAR 1b/cycle | $P_{total} = f\left(6\left(E_{comp_{pre}} + E_{clk} + E_{DAC} + E_{logic}\right) + E_{T\&H}\right)$ |
| Asynchronous 6-bit SAR 1b/cycle | $P_{total} = f\left(6\left(E_{comp_{pre}} + E_{clk}\right) + 5\left(E_{DAC} + E_{logic}\right) + E_{T\&H}\right)$ |
| Asynchronous 2-bit SAR 1b/cycle | $P_{total} = f\left(E_{T\&H} + 2\left(E_{comp_{pre}} + E_{clk}\right) + E_{logic} + E_{DAC}\right)$ |
| Asynchronous 6-bit SAR 3b/cycle | $P_{total} = f\left(2E_{T\&H} + 14\left(E_{comp_{pre}} + E_{clk}\right) + 2E_{logic} + 7E_{DAC}\right)$ |
| 6-bit flash | $P_{total} = f\left(2E_{T\&H} + 2^{N-1}\left(E_{comp_{pre}} + E_{clk} + E_{latch}\right)\right)$ |
| Pipeline of 4-bit SAR (2b/cycle, asynchronous) and 2-bit flash | $P_{total} = f\left(8E_{comp_{pre}} + 8E_{clk} + 2E_{logic} + 2E_{T\&H} + 3E_{latch}\right)$ |
| Pipeline of 2-bit flash and 4-bit SAR (asynchronous, 2b/cycle) | $P_{total} = f\left(8E_{comp_{pre}} + 8E_{clk} + 2E_{logic} + 2E_{T\&H} + 3E_{latch}\right)$ |
| Pipeline of 2x 3-bit SAR ADC | $P_{total} = f\left(6E_{comp_{pre}} + 6E_{clk} + 5E_{logic} + 5E_{DAC} + 2E_{T\&H}\right)$ |
| Pipeline of 3x 2-bit SAR ADC | $P_{total} = f\left(3E_{T\&H} + 6E_{cmp} + 5E_{logic} + 6E_{clk} + 5E_{DAC}\right)$ |
| Pipeline of 2x 3-bit flash | $P_{total} = f\left(2E_{T\&H} + 14E_{comp_{pre}} + 14E_{clk} + 14E_{latch} + E_{DAC}\right)$ |
| Pipeline of 2-bit flash and 5-bit SAR (1b redundancy), with inter-stage gain | $P_{total} = f\left(5E_{comp+pre} + 5E_{clk} + 4E_{DAC} + 3E_{logic} + 2E_{T\&H}\right) + P_{amp}$ |
| Pipeline of 3-bit flash and 4-bit flash (1b redundancy) with inter-stage gain | $P_{total} = f\left(7.2E_{comp+pre} + 7.2E_{clk} + E_{DAC} + 7.2E_{latch} + 2E_{T\&H}\right) + P_{amp}$ |

4. Operating Principle

The previous chapter has shown that a 6-bit asynchronous SAR ADC with a 3-bit/cycle resolving scheme is the most promising candidate for realizing an ADC that can run at 2.5GS/s while consuming a reasonable amount of power. Before any implementations of the blocks can be considered, their specifications need to be devised in relation to the broader context of the operating principle of the ADC.

In order to prove that the concept is a viable one, a Verilog-A model of the converter was created, which is also intended as a test-bench for the sub-circuits after they are implemented in the 40nm technology. The particularities of this model, together with the asynchronous signal generation that allows the circuit to run correctly are described in detail in this chapter. All simulation results presented were generated using the ideal model.

4.1. ADC Architecture

The SAR ADC features a track and hold front-end, 2 sets of 7 comparators that act like low-resolution flash ADCs and are responsible for determining 3 bits each, a set of 7 DACs that generate the references for the comparators to make their decision and a digital logic module which controls the reference generation (Figure 4.1.1).

The internal operation is asynchronous, so as soon as one of the blocks has finished its task, it needs to trigger the start of the next, in a domino-like scheme. For example, the first set of comparators can be clocked as soon as the hold mode of the track and hold begins.

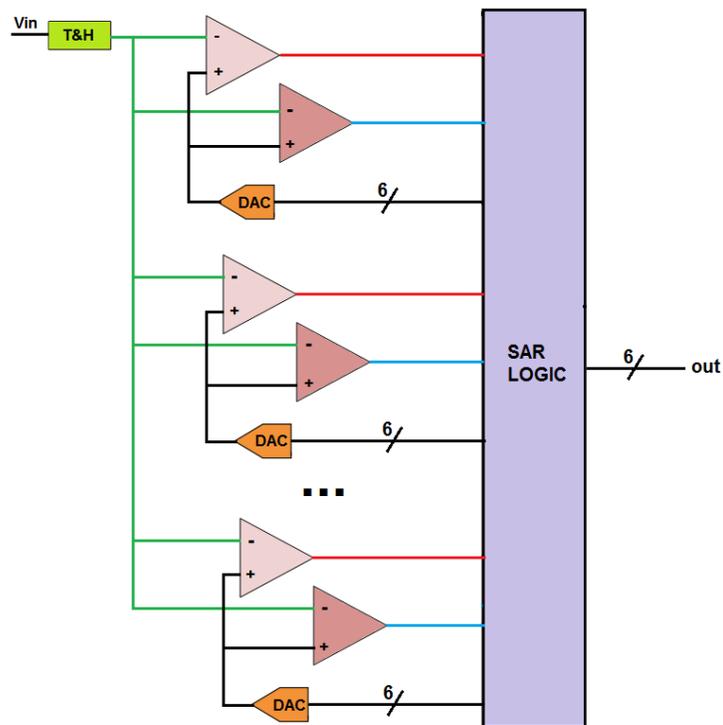


Figure 4.1.1 – Block diagram of a 6-bit asynchronous SAR ADC with a 3-bit/cycle resolving scheme

4.2. Asynchronous signal generation

Figure 4.3.1 details the way the asynchronous logic is generated based on an external clock signal with a pulse width of 50ps and a period equal to the sampling period (*rst_DAC_logic*). The operation of the ADC can be split into the following six phases:

1. Reset of the DAC,
2. Sampling of the input and pre-charging of the DAC according to the first set of references,
3. First set of comparisons,
4. Reference generation for the second set of comparisons,
5. Second set of comparisons,
6. Synchronization of the output bits.

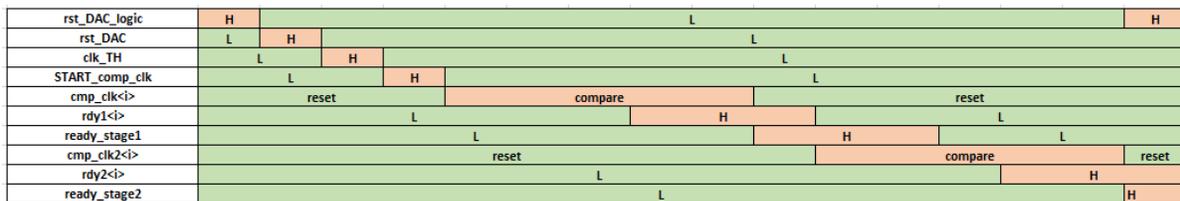


Figure 4.2.1 - Timing diagram of asynchronous signal generation

During the first phase, half of the DAC capacitors are connected to the positive reference, while the other half are connected to the negative reference. These switches are controlled via flip-flops with a data-to-Q delay, t_{del1} . This is compensated by introducing a delay equal to the propagation delay of a flip-flop (t_{del1}) between the reset signal of the DAC logic (*rst_DAC_logic*) and that of the DAC itself (*rst_DAC*). A delay equal to the reset time of the DAC (t_{del2}) is interposed between the reset signal and the track-and-hold clock (*clk_TH*), such that the negative edge of the DAC reset coincides with the positive edge of *clk_TH*. In a similar fashion, for the maximum speed benefit, the positive edge of the clocks of the first comparator set (*cmp_clk* < *i* >) need to coincide with the negative edge of *clk_TH*, which is realized in practice using the signal *START_cmp_clk* and t_{del3} . This choice also accounts for the delay of the flip-flop responsible for generating each comparator clock. During the tracking phase, the DAC switches are used to generate the correct references for each comparator.

The third phase begins as soon as the comparators are clocked. The NAND gate at the output of the comparators detects whether a valid decision has been reached by checking if its inputs are of different polarity, which then leads to the generation of the ready signal pertaining to that particular comparator. A set of AND gates are used to detect when the whole set of comparators has generated a valid output, which triggers the DAC switches and propagates the results of the first 7 comparisons so that the new references can be built on top of them (Figure 4.2.2).

The positive edge of the second set of comparators is then generated using the ready_stage1 signal which indicates that the first stage has completed its task. The DAC needs to have settled to $\frac{V_{LSB}}{2}$ accuracy before the clock of the second set of comparators is generated, otherwise the comparisons will use incorrect references.

The track-and-hold clock is used to trigger the reset of the last set of comparators, thus ensuring that they are always reset, even in case of a metastable event which would prevent the ready signals from being generated in case the same scheme adopted for the first set of comparators were used here.

Last, but not least, the output bits which are stored in SR-latches that follow the comparator, are synchronized by the synch signal so that they can be processed by the next block in the correct sequence. The delay of the output flip-flops is not part of the critical path, as the synchronization can be done while the circuit is performing the next conversion.

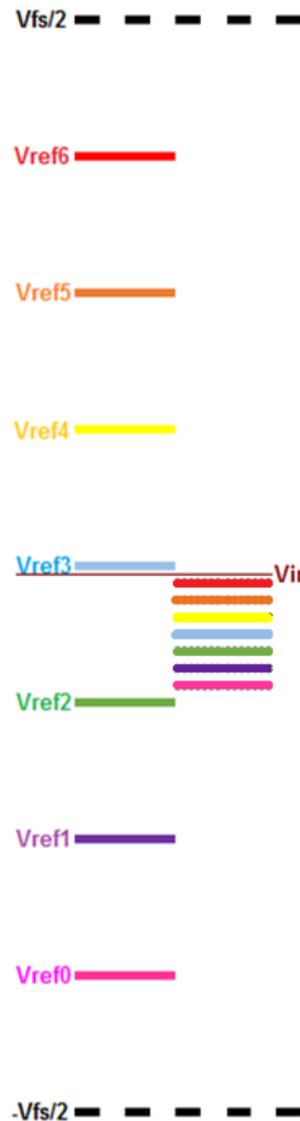


Figure 4.2.2 - Reference generation example

4.3. Treatment of metastable events

Owing to the internal clock generation scheme, a metastable event in the first set of comparators can prevent the ready signal from being generated, which in turn, will not trigger the second set of comparators. Such a situation will cause the resolution of the ADC to temporarily drop to 3 bits, which is unacceptable for the application envisioned. During the architecture study it was assumed that the minimum input that can be encountered by the comparator is of $\frac{V_{LSB}}{4}$, which corresponds to a delay of 94ps. In a practical design, this is unfortunately not always the case, and smaller inputs can occur. For the scope of this design, any comparison requiring more than 94ps to be completed is considered a metastable event.

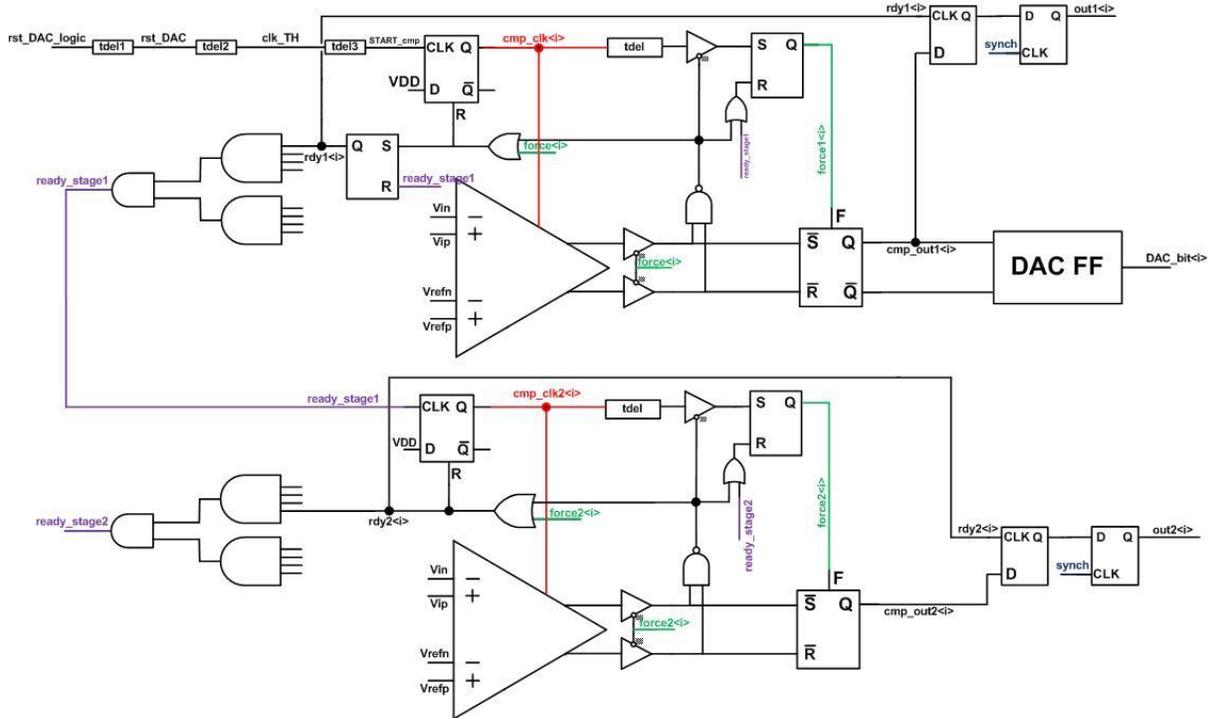


Figure 4.3.1 - Block diagram of asynchronous signal generation

Although metastability is a problem that can never be completely mitigated, a solution for ensuring that the operation of the converter can continue despite such an event is stringently needed. In this design, a circuit forcing the output of a metastable comparator is implemented and hereon referred to as the *output-forcing* circuit.

In Figure 4.3.1, the positive edge of the comparator clock is passed through a delay that generates a force signal ($force < i >$) after the time allotted for the comparison has elapsed. This signal commands the SR-latch and causes its output to go to V_{DD} . If the force signal was generated, the path connecting the SR-latch to the outputs of the comparators is disabled, such that the forced output is not overwritten by the comparator decision after it is completed. Conversely, the path that generates the force signal is disabled if the NAND detects a valid decision. The force signal is reset to ground as soon as $ready_stage1$ is charged to V_{DD} , signaling that forcing an output is no longer needed.

In addition to forcing an output for the metastable comparator, the force signal is also used to gate the ready signal in order to ensure that $ready_stage1$ is always generated and triggers the next stage.

Let us assume that the ADC is faced with the metastable event in the first set of comparators, as shown in Figure 4.3.3. The input voltage is close to V_{ref3} , so the comparators will take a very long time until they reach a correct decision. Waiting an infinite amount of time would lead the positive output of the comparator to be discharged to ground, meaning that the second set of references generated by the DAC would be those circled with red in

Chapter 4 – Operating Principle

Figure 4.3.2. If no other errors occur, the outputs of the first set of comparators are 1110000, while the second set gives 1111111.

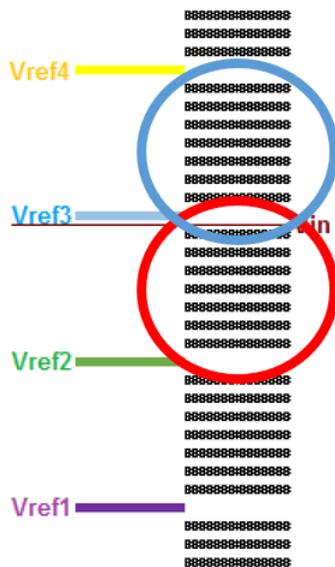


Figure 4.3.2 - Outcome of a metastable event in the first set of comparators

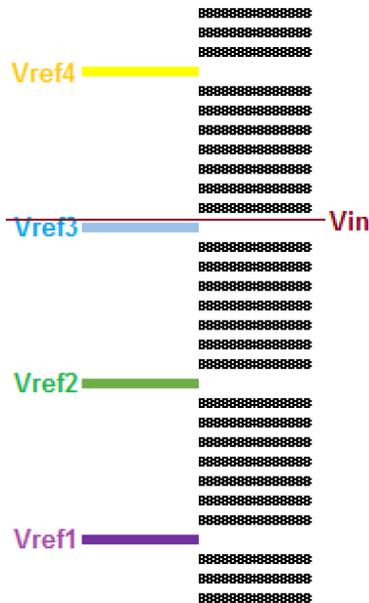


Figure 4.3.4 - Metastable event in the first set of comparators that does not lead to an error

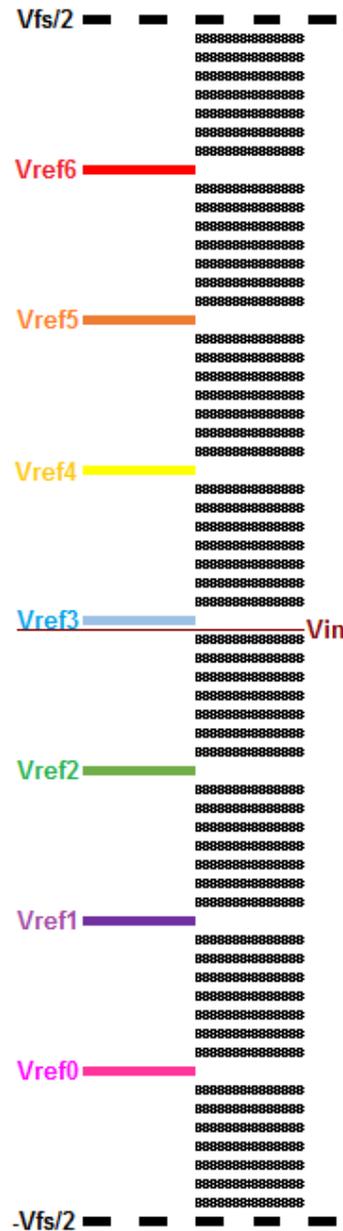


Figure 4.3.3 - Metastable event in the first set of comparators that leads to an error

If the output is forced to the opposite value (V_{DD}) by default, the set of references generated would be those circled in blue. In this case, the first set of comparisons leads to a code of 1111000, while the second gives 0000000. The error introduced by forcing the

output of the metastable comparator is of only one LSB, which is a small compromise to make to avoid losing 3 bits. It is important to note that this error does not occur in all metastable cases, as proven by Figure 4.3.4. In this case, the output of the comparator would have gone high even if it were not forced to the default value, so the correct set of references will be used for the second set of comparisons.

If the metastable event occurs in the second set of comparators, its outcome is also an error of at most 1LSB (Figure 4.3.5). Bearing in mind that the results of these decisions are fed to the output buffer directly and do not influence the reference generation, it is tempting to leave them untreated. However, failing to provide a ready signal in the second stage has the potential of interfering with the reset of the comparators and with that of some flip-flops or latches in the first stage, so it is important that every output bit is assigned.

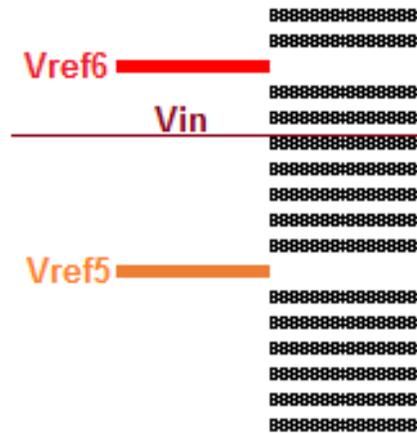


Figure 4.3.5 - Outcome of a metastable event in the second set of comparators

In order to prove the efficiency of the output-forcing circuit in dealing with metastable comparators, a ramp with a voltage confined to the input range and with a small slope was applied at the input. If the derivative (slope) of this signal is small enough, the ADC will generate all possible output codes and will cause a metastable event to occur in each comparator over the simulated time.

If this test is performed when the output-forcing circuit is disabled, comparing the analog signal reconstructed from the output bits with the input signal leads to the error profile shown in Figure 4.3.6. The waveform shows that once in a while, a very high error will occur, which means that the accuracy of the ADC was compromised.

Running the same simulation on the ideal model when the forcing circuit is activated reveals that the error remains within 1 LSB at all times, which proves that the solution described previously is viable and prevents dramatic losses in accuracy due to metastability.

It is important to mention that the asynchronous processing scheme is beneficial for reducing the incidence of metastability issues. If the input voltage is close to one of the references from the first set of comparators, then it will not trigger a metastable event in the second set also. This means that although the first decision may take longer, the clock of the second stage will also arrive later and the second decision will need less time to complete because it corresponds to a bigger input.

Chapter 4 – Operating Principle

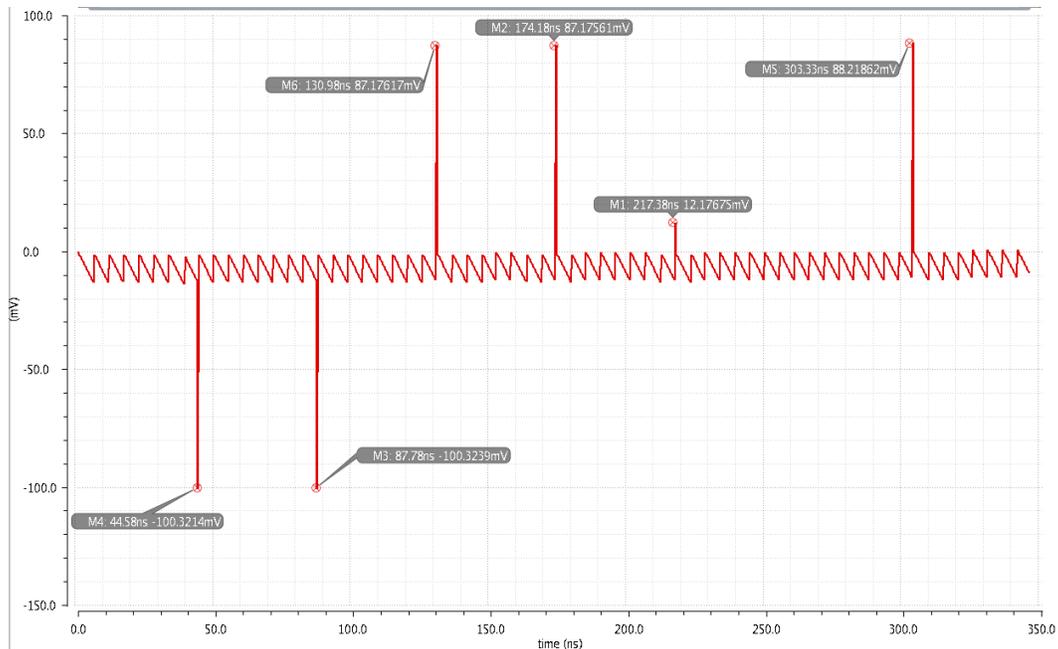


Figure 4.3.6 - Error signal without the "force" circuit for a ramp input

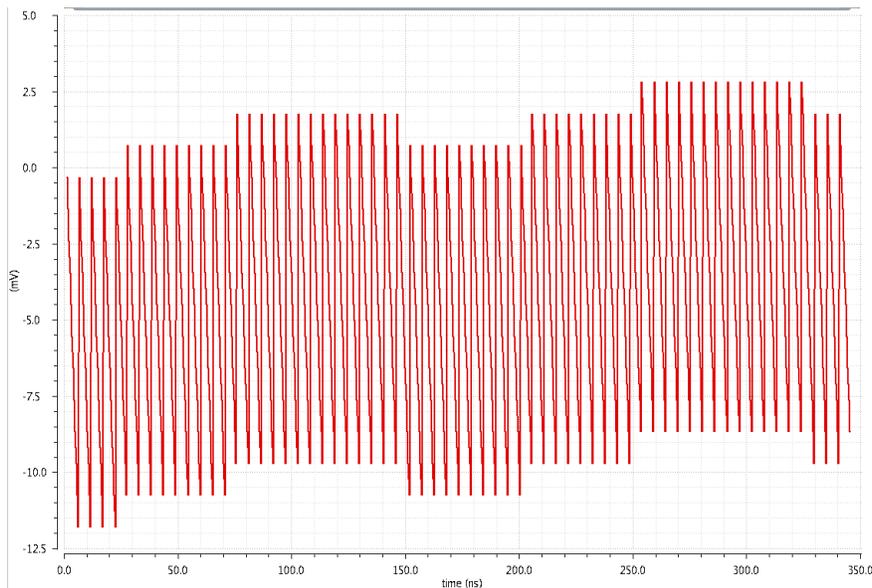


Figure 4.3.7 - Error signal with the "force" circuit for a ramp input

Furthermore, metastability was modelled by relating the input differential voltage to the delay of the comparator and was assumed to only appear in this block. In reality, the SR latch used to generate the force signal may become metastable. The viability of the output-forcing circuit will be further assessed using the complete transistor-level implementation and described in the next chapter.

4.4. Verilog-A simulation results

The RTL-level blocks described previously were modelled behaviorally in Verilog-A by using the assumptions made in the previous chapter, which allowed a proof-of-concept simulation to be run. Figure 4.4.3 shows the track-and-hold signal, together with the seven reference voltages generated by the DACs. The coarse-fine conversion principle is evident: the first set of comparisons uses the large references and based on the outputs of the comparators selects the interval which contains the input signal, which will be used to generate the next set of references. The plot is generated for a sampling frequency of 2.2GS/s, an input signal at Nyquist and a differential amplitude of 800mVpp.

If the output bits are reconstructed using an ideal DAC, the resulting waveform can be compared with a delayed version of the track-and-hold signal. Figure 4.4.1 shows precisely this and confirms that the ADC works as intended.

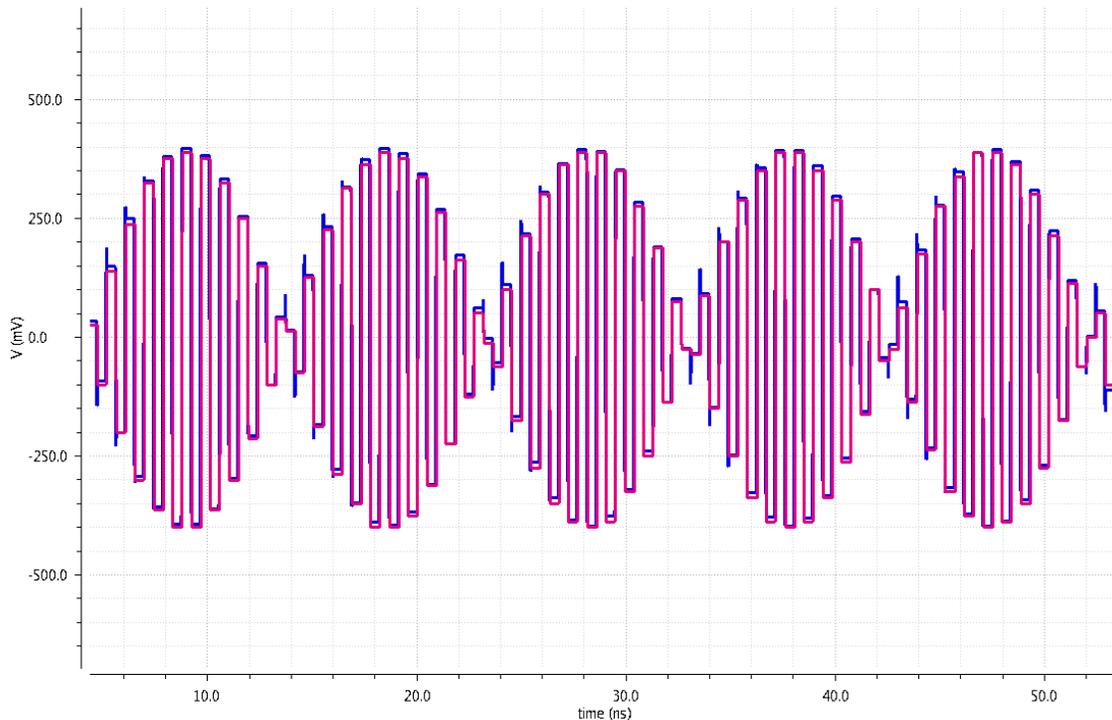


Figure 4.4.1 - Comparison between track-and-hold signal (blue) and reconstructed output (pink)

The error of the ADC is given by the difference of the waveforms and can be seen in Figure 4.4.2.

Chapter 4 – Operating Principle

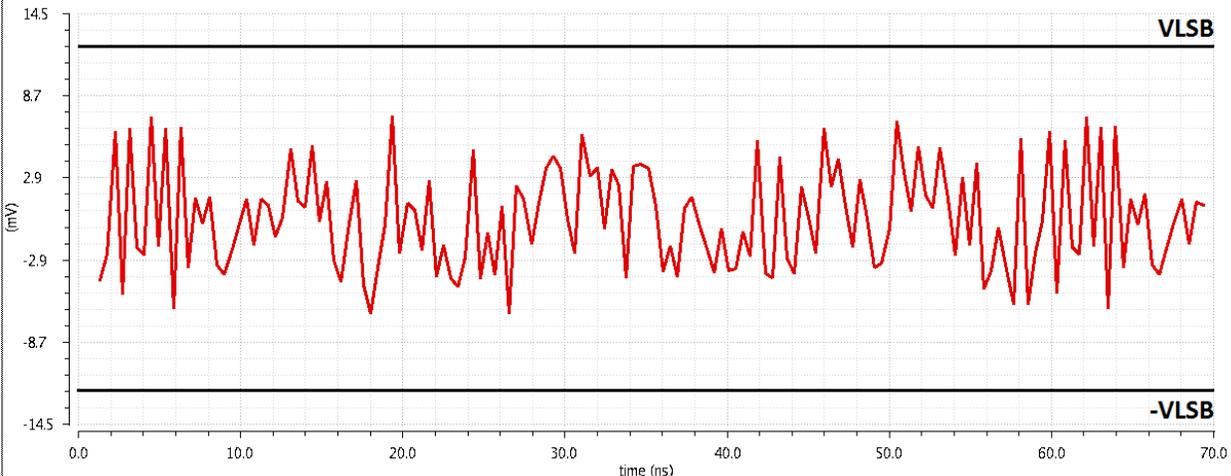


Figure 4.4.2 - Error (Difference between reconstructed signal and delayed T&H signal)

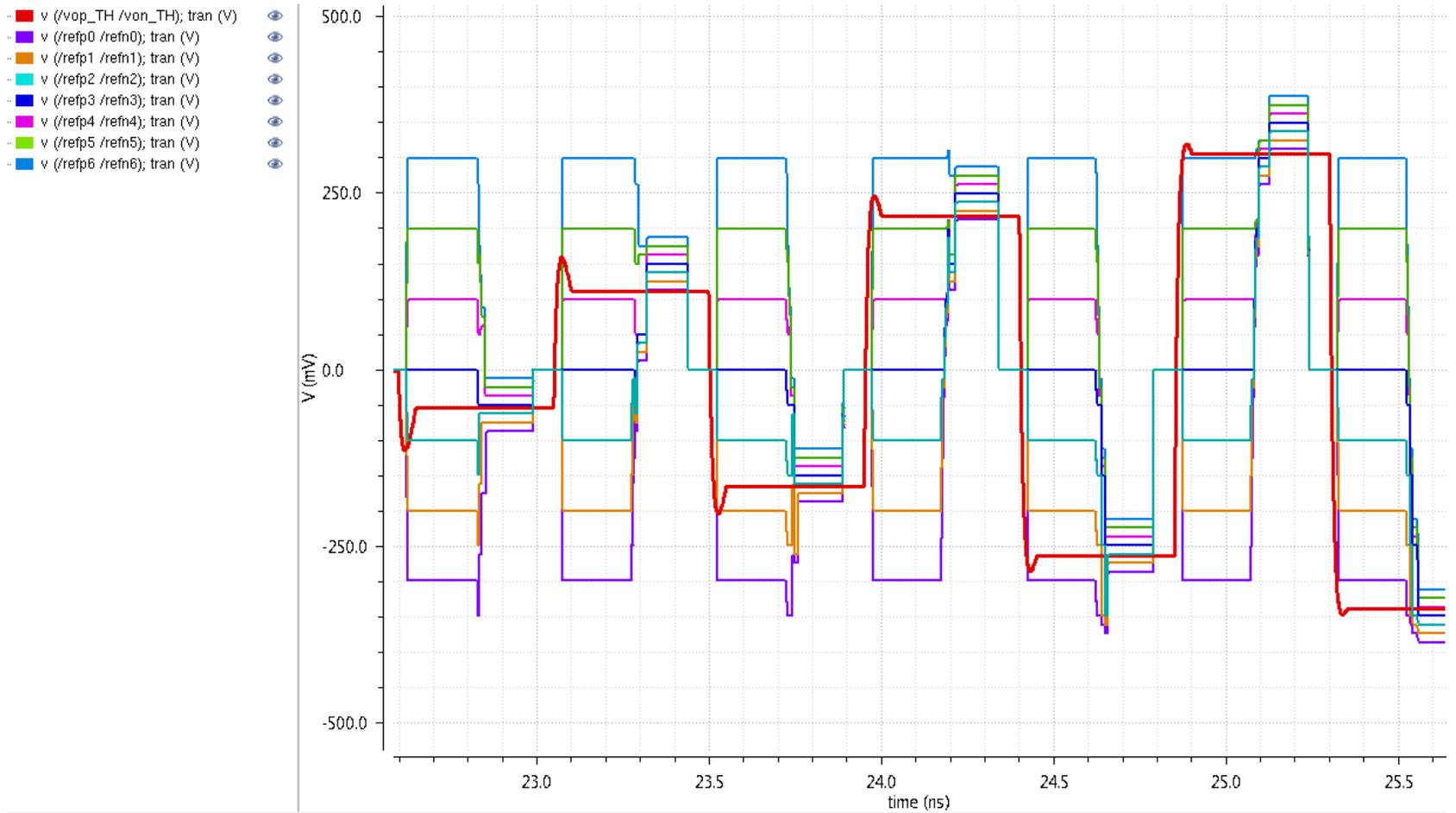


Figure 4.4.3 - Reference generation for a section of a sinusoidal input signal

Chapter 4 – Operating Principle

Figure 4.4.4 shows the FFT of the reconstructed output bits and reveals an ENOB of 5.92 bits (THD=-49.15dB) at a sampling frequency of 2.22GHz (corresponding to an input frequency of 1.059GHz). The accuracy of the result is guaranteed by coherent sampling, which imposes that the sampling frequency must be a prime integer of the input frequency. The parameters used to apply this technique to this ADC are summarized in Table 4.4.1. Figure 4.4.5 shows a sample of the asynchronous signal generation.

Table 4.4.1 – Parameters used to set up the FFT simulation

| Parameter | Notation | Value used for the FFT |
|--------------------------------------|-------------------------------------|------------------------|
| Number of points for the FFT | N_{FFT} | 128 |
| Sampling speed | f_s | 2.22GHz |
| Simulation time | $T_{sim} = N_{FFT} * \frac{1}{f_s}$ | 57.6ns |
| Frequency resolution | $f_{bin} = \frac{1}{T_{sim}}$ | 17.361MHz |
| An integer number of periods (prime) | M_{per} | 61 |
| Input frequency | $f_{in} = M_{per} * f_{bin}$ | 1.059GHz |

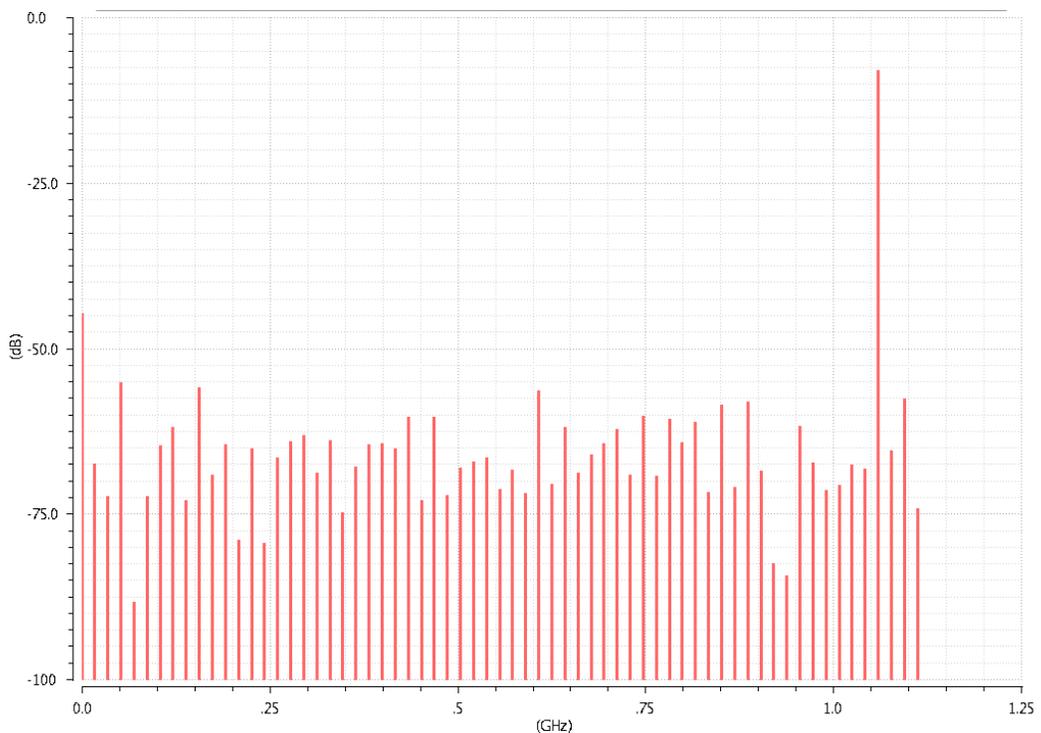


Figure 4.4.4 - FFT of the reconstructed signal

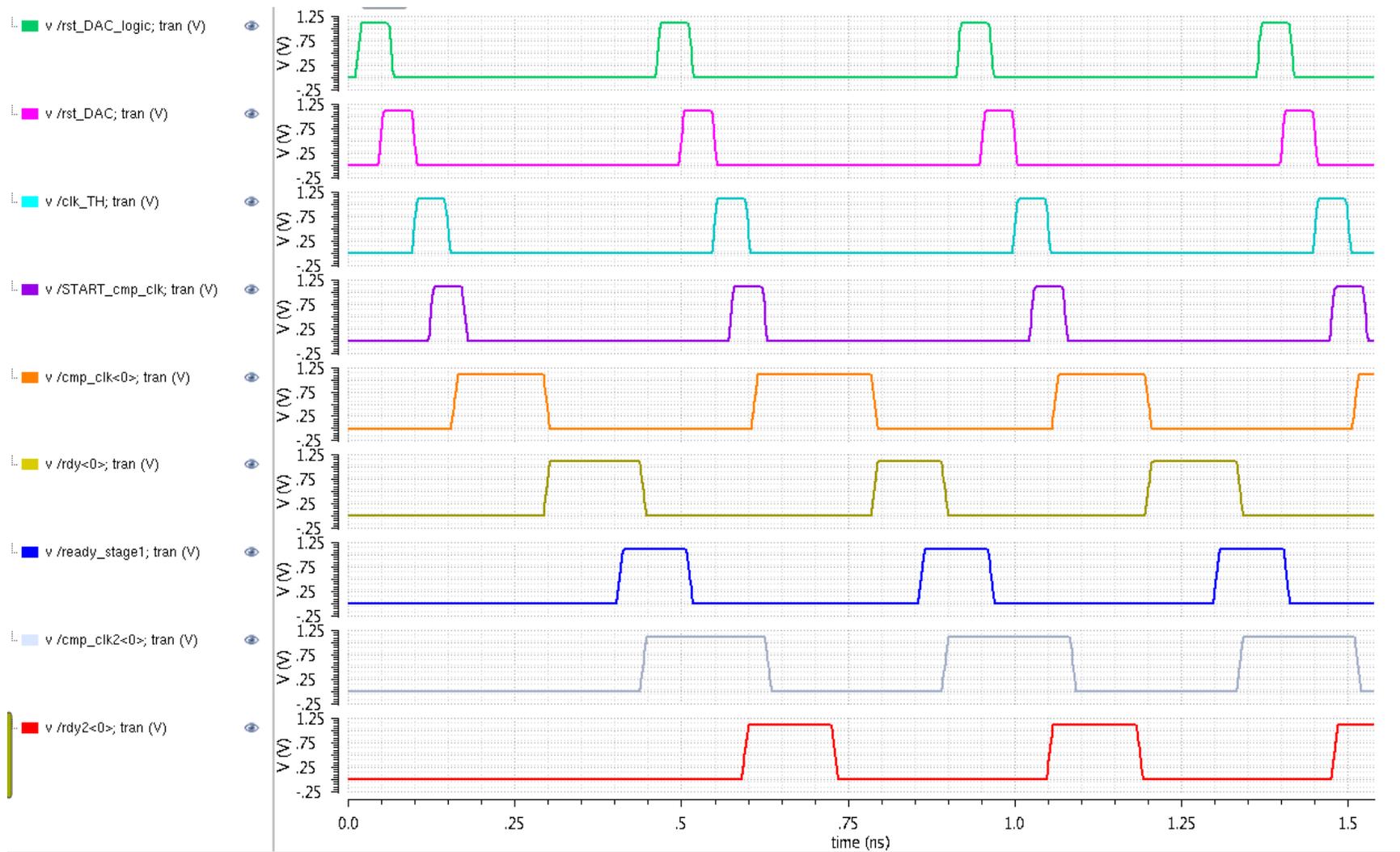


Figure 4.4.5 - Asynchronous signal generation

5. Transistor-Level Design and Simulation

The purpose of this chapter is to describe the transistor-level implementation of the main building blocks that comprise the 6-bit asynchronous 3-bit/cycle SAR ADC, specifically the track-and-hold, comparator and associated preamplifier, DAC and digital logic.

At the end of the chapter, a possible implementation of the reference generator and offset calibration are presented together with the performance achieved by other designs that employ them successfully. Specific design considerations imposed by the specifications of this ADC are presented, but owing to the limited time at the author's disposal, their actual design and simulation lie outside the scope of this thesis.

5.1. Track-and-Hold

5.1.1. Specifications

The application envisioned for the high-speed ADC is time-interleaved, which means that the input stage, which is responsible for the sampling of the circuit, can be implemented in two different ways.

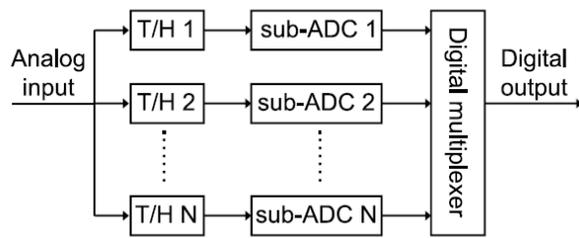


Figure 5.1.1.1 - Interleaved ADC without front-end sampler [35]

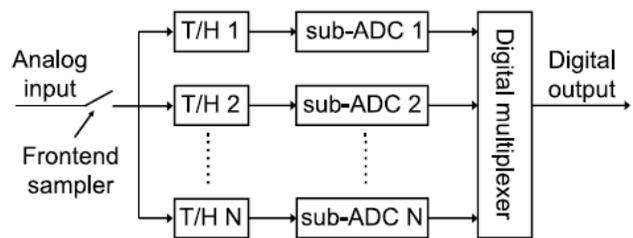


Figure 5.1.1.2 - Interleaved ADC with front-end sampler [35]

Both Figure 5.1.1.1 and Figure 5.1.1.2 feature a time-interleaved ADC (TI-ADC) with an interleaving factor of N . The difference between the two designs is that the one in Figure 5.1.1.2 features an extra block - the front-end sampler. The extra circuit is introduced to mitigate the errors that frequently occur in the architecture of Figure 5.1.1.1 due to timing mismatches appearing because of process variation existing between the different channels. Supposing that f_{sTI} is the sampling rate of the TI-ADC and that $f_{sSingleChannel}$ is the operating frequency of each sub-ADC, it holds for both designs that:

$$f_{sTI} = N f_{sSingleChannel} \quad (5.1)$$

The bandwidth of the TI-ADC is the same as that of the sub-ADC in Figure 5.1.1.1, while in the case of Figure 5.1.1.2 it is equal to the bandwidth of the front-end sampler. In this thesis we target a TI-ADC speed of 20 GHz, and an analog input bandwidth of 10 GHz. The track-and-hold topology chosen should thus be able to sample the input signal with an accuracy of 6-bits at 2.5 GS/s while maintaining this bandwidth.

In what follows, the simulation results pertaining to the track-and-hold of the sub-ADC will be described in detail. It is not the purpose of this project to assess which of the two implementations will be adopted in the TI-ADC.

5.1.2. Possible Topologies

The study performed on the different track-and-hold topologies is intended as a comparison between the sizes required to achieve an ENOB of 7 bits. If 6 bits were targeted instead of 7, a reduction of 0.5 bits (3 dB) would be visible in the SNDR. It is expected that this approach aptly exposes the trade-off between linearity and speed for this particular block. The track-and-hold is separated from the DAC and allows the sampling capacitor to be sized based on the noise requirements.

Table 5.1.2.1 summarizes the configurations explored, together with their corresponding schematic, for easier reference.

Table 5.1.2.1 – List of track-and-hold topologies that are compared in this chapter

| Topology Specifics | Figure |
|--|--|
| Differential NMOS with dummies on the sides | Figure 5.1.2.3 |
| Differential NMOS without dummies | Figure 5.1.2.3 |
| Differential complementary MOS with dummies on the sides | Figure 5.1.2.2 |
| Differential complementary MOS without dummies | Figure 5.1.2.2 without M_{dum1} – M_{dum4} |

A small sampling capacitor helps increase the input bandwidth of the ADC, so it is desirable to use the total input capacitance of the preamplifiers for this operation. The only drawback of this approach would be a limited achievable linearity due to the fact that this is a non-linear gate capacitance, but in the 6-bit case it will be shown that this is not an issue. In what follows, it is assumed that the sampling capacitor is equal to $C_S = 14 * C_{in_{pre}} = 14 * 12\text{fF} = 168\text{fF}$. The on-resistance associated with the sampling switch, together with its parasitic capacitance and the sampling capacitance limit the settling speed of the track-and-hold and are thus responsible for imposing the minimum theoretical tracking time that can

be used. For practical reasons, the tracking time cannot be lower than 50ps because signals with a shorter pulse width cannot be generated in the 40nm technology that is available. It is therefore desirable to push the tracking time towards this limit as much as possible in order to maximize the speed of the circuit. It is also assumed that the ADC input has a 50Ω termination and is driven by a 50Ω source, which can be accounted for by setting $R = 25\Omega$ in either of the topologies described below (Figure 5.1.2.2 and Figure 5.1.2.3).

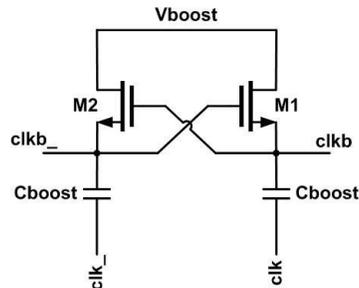


Figure 5.1.2.1 - Clock-boosting circuit

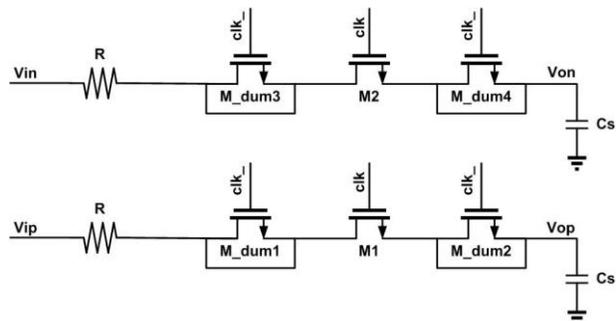


Figure 5.1.2.3 -Differential NMOS track-and-hold circuit

The design in Figure 5.1.2.3, inspired by [36], relies on an NMOS transistor to perform the sampling function. Having set the sampling capacitance, the next parameter that needs to be tuned in order to set the tracking time is the on-resistance of the switch, which for an NMOS transistor is given by [37]:

$$R_{on} = \frac{1}{\mu C_{ox} \frac{W}{L} (V_{GS} - V_{THn})}, \text{ where} \quad (5.2)$$

$$V_{GS} = V_G - V_{in} \quad (5.3)$$

The equation above exposes the dependence of the on-resistance on the input signal which causes non-linearity and also the fact that a high gate-to-source voltage is beneficial both for reducing the relative contribution of V_{in} and for obtaining a low resistance. Based on this observation, achieving a short settling time is possible if a charge pump is used to increase the clock voltage such that V_G increases.

The input common mode voltage is set to 700mV, a value that is beneficial for the comparator speed, as will be explained later. Figure 5.1.2.1 shows the implementation of choice, taken from [36]. Clk and $Clk_$ are the two opposite phases of a conventional clock with a pulse width equal to the tracking time, toggling between “0” and “ V_{DD} ”. The boosted version, denoted by $Clkb$ and $Clkb_$, respectively, will toggle between “ V_{boost} ” and “ $V_{boost} + V_{DD}$ ”. It is tempting to use a very high V_{boost} in order to maximize the speed of the track-and-hold circuit, but the maximum gate voltage that is allowed for an NMOS sets an upper limit at 600mV. The input swing is set at 800mV_{pp}, which means that even in the case when the minimum input signal ($V_{CM} - 200mV$) is present at the source of the switches, they can still be turned off. This design uses $V_{boost} = 500mV$ to ensure that the switch is indeed switched off during the hold time. The technology available limits the gate voltage to 1.2V due to reliability issues, so V_{boost} cannot be increased further.

Transistors M_{dum1} through M_{dum4} are included to mitigate charge-injection and clock-feedthrough errors. They are clocked by the opposite phase as the track-and-hold clock and should be sized two times smaller than the switch itself.

A different way of linearizing the on-resistance of the sampling switch is shown in Figure 5.1.2.2, where a PMOS transistor is connected in parallel to the NMOS and is clocked with the opposite phase. The on-resistance in this case is given by:

$$R_{on} = R_{onp} || R_{onn} \quad (5.4)$$

This method is useful for maintaining a constant on-resistance over a larger range of input voltages, which helps improve the accuracy of the circuit. The best results are achieved when the width of the PMOS is 2.5 times larger than that of the NMOS, which compensates for the difference in carrier mobility and ensures both switches have the same speed. When the PMOS turns off ($clk = 1$), the charge stored in its channel will be pushed out of the transistor. Since the NMOS is smaller, it cannot absorb this charge in its entirety, so dummy switches are used to absorb the rest of the charge (M_{dum1} through M_{dum4}). Similarly, when the NMOS is discharged ($clk = 0$), the extra charge absorbed by the longer channel of the PMOS will be provided by the dummies and not by the input and output voltages, which means that no errors will be introduced by the switching events. A dummy size 1.5 times larger than that of the NMOS switches secures the desired operation by providing equal channel lengths during either clock phase.

In order to allow a fair comparison for the different architectures, the input voltage was adapted such that the output signal amplitude would remain the same. This step was necessary because of the attenuation that occurs due to the parasitic capacitance of the switch.

5.1.3. Conclusion

Table 5.1.3.1 summarizes the simulation results obtained for each architecture in case 50ps or 75ps were adopted for the settling time. The main objective of the comparison is to determine the minimum sizing of the switch that leads to an ENOB of 7 bits, which is considered sufficient for the desired accuracy. Some of the architectures do not achieve this accuracy no matter how much their size is increased because the parasitic capacitance associated increases proportionately and limits the achievable linearity and bandwidth.

The table also lists the associated bandwidth and input amplitude required to compensate for the attenuation of the circuit. In order to characterize the designs using these elements, it was assumed that the preamplifiers used as load were not clocked. For reference purposes, each switch type was also simulated with an ideal capacitor with the same value as $14C_{pre}$ to assess the non-linearity due to the input capacitance of the preamplifiers.

When the preamplifiers are not clocked, their input capacitance appears quite linear and the kickback coming from the clock signal is not visible. Because of this, the simulations were performed with half the preamplifiers clocked at the falling edge of the track-and-hold clock, while the other half was clocked after half of the sampling clock period. This setup is useful in determining if the kickback seen on the input signal from the first set of preamplifiers is similar to the one seen by the second set. An error of less than $\frac{1}{2}V_{LSB}$ was deemed acceptable for the settling inaccuracy and neither topology exceeded it.

Considering a tracking time of 75 ps, as expected, smaller switches are needed in all architectures if an accuracy of 7 bits is to be obtained. Even with this compromise in speed, the complementary MOS architecture is generally undesirable because it cannot achieve a bandwidth of 10 GHz. The boosted NMOS topology appears the most promising and increasing the size of the NMOS switch reveals that the target bandwidth and accuracy can also be obtained with a shorter tracking time – 50 ps – which is the narrowest pulse width that can be passed through the digital logic without disappearing. The main disadvantage of the boosted NMOS - having to generate an extra voltage – keeps the complementary MOS architecture attractive, so the 7-bit accuracy is attempted with this topology also, but even with very large transistors (40um for the NMOS and 100um for the PMOS), it is clear that self-loading effects hinder any gain in accuracy and severely limit the bandwidth.

Table 5.1.3.1 - Sizing and achievable performance of the proposed topologies for a tracking time of 50ps and 75ps

| Topology | 50ps tracking time | | | | | 75ps tracking time | | | | |
|--|--------------------|----------|----------|----------|--|--------------------|----------|----------|----------|--|
| | Wn/Wp (um) | Vid (mV) | ENOB (b) | BW (GHz) | Difference between Vin seen by the two preamplifier sets <LSB/2? | Wn/Wp (um) | Vid (mV) | ENOB (b) | BW (GHz) | Difference between Vin seen by the two preamplifier sets <LSB/2? |
| Differential NMOS with dummies on the sides (Figure 5.1.2.3) | 41 | 440 | 7.06 | 16.76 | ✓ | 17 | 444 | 7.03 | 13.58 | ✓ |
| Differential NMOS without dummies (Figure 5.1.2.3) | 42 | 415.36 | 7.00 | 19.36 | ✓ | 17 | 414 | 7.1 | 14.14 | ✓ |
| Differential complementary MOS with dummies (Figure 5.1.2.2) | 40/ 100 | 460 | 6.55 | 9.626 | ✓ | 19/ 47.5 | 436.5 | 7.25 | 7.72 | ✓ |
| Differential complementary MOS without dummies (Figure 5.1.2.2 without $M_{dum1} - M_{dum4}$) | 40/ 100 | 460 | 6.66 | 11.59 | ✓ | 20/50 | 436.5 | 7.25 | 8.88 | ✓ |

5.2. Digital-to-Analog Converter

The digital-to-analog converter plays an important role in the SAR loop because it is responsible for generating the voltage references that will be compared to the input. Seven such circuits are part of the design, so a small form-factor is preferred. In what follows, the architecture choice will be motivated in the context of the 3-bit/cycle resolving scheme, the reset and settling times will be estimated and last, but not least, the sizing of the capacitors and switches will be shown.

5.2.1. Architecture

The DAC settling time is part of the critical path of the SAR ADC, so it is crucial to minimize it. For the same reasons, the propagation delay of the digital logic that controls the generation of the second set of voltage references needs to be reduced as much as possible. Aside from this, the 3-bit/cycle resolving scheme implies that a conventional unary or binary-scaled DAC cannot be adopted, which prompts a segmented architecture if delay and power dissipation are to be minimized by avoiding a decoder. Compared to its unary and binary counterparts, this design does not require a decoder to be placed between the outputs of the first set of comparators and the input of the DAC.

A segmented design suitable for this ADC architecture features 7 capacitors with a value of $8C_u$ and 8 capacitors with a value of C_u , which allows the LSB to be correctly generated for each comparator set. The total capacitance of the DAC is equal to $64C_u$, the same value that a unary or binary topology would require. The most straightforward switching scheme available entails the use of three voltages: the common-mode voltage V_{CM} and the positive and negative reference voltages, V_{refp} and V_{refn} , connected as shown in Figure 5.2.1.1.

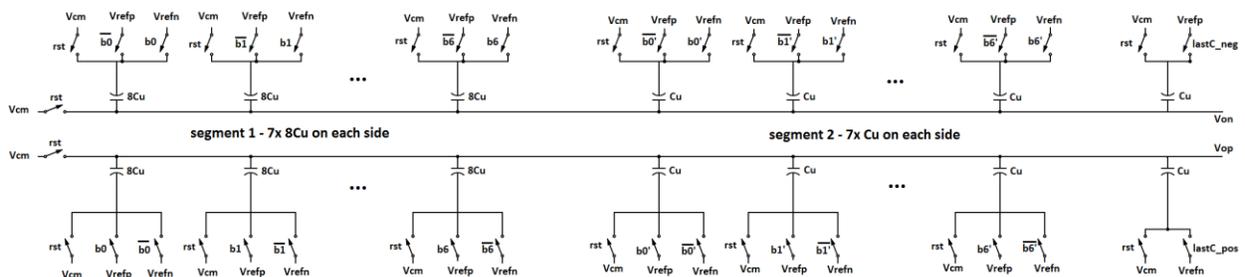


Figure 5.2.1.1 – Segmented 6-bit DAC with a 3-voltage switching scheme

For the positive side of the DAC, the operation is as follows: during the reset phase, both the bottom and the top plates of the capacitors are connected to V_{CM} in order to bring

the output voltage to a known, well-defined state before the next reference generation needs to take place. Afterwards, during the first pre-charge phase, the reset switches are opened and the bottom plates of the capacitors in the first segment are switched to either V_{refp} or V_{refn} , depending on which voltage reference needs to be generated. The bottom plates of the second segment are all connected to V_{refn} during this phase. After the first set of comparators has made a decision, the outputs switch the first segment of the DAC to select which reference level will be used for the next pre-charge phase. On top of the reference provided by the segment, the appropriate voltage level for each DAC will be formed by switching the capacitors in the second segment.

The operation of the negative side of the DAC is identical, except that V_{refp} and V_{refn} are reversed. The timing diagram corresponding to the positive side of the DAC can be found in Figure 5.2.1.2. **P** (**N**) signifies that the bottom plate of that particular capacitor is connected to V_{refp} (V_{refn}), while **CM** indicates a connection to the common-mode voltage.

| POSITIVE DAC | | | Segment 1 | | | | | | | Segment 2 | | | | | | | |
|-----------------|-------|------|-----------------------------|-----|-----|-----|-----|-----|-----|-----------|----|----|----|----|----|----|----|
| | | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Capacitor value | | | 8Cu | 8Cu | 8Cu | 8Cu | 8Cu | 8Cu | 8Cu | Cu | Cu | Cu | Cu | Cu | Cu | Cu | Cu |
| RESET | | | CM | CM | CM | CM | CM | CM | CM | CM | CM | CM | CM | CM | CM | CM | CM |
| Pre-charge 1 | -3/8 | DAC1 | P | N | N | N | N | N | N | N | N | N | N | N | N | N | N |
| | -2/8 | DAC2 | P | P | N | N | N | N | N | N | N | N | N | N | N | N | N |
| | -1/8 | DAC3 | P | P | P | N | N | N | N | N | N | N | N | N | N | N | N |
| | 0 | DAC4 | P | P | P | P | N | N | N | N | N | N | N | N | N | N | N |
| | 1/8 | DAC5 | P | P | P | P | P | N | N | N | N | N | N | N | N | N | N |
| | 2/8 | DAC6 | P | P | P | P | P | P | N | N | N | N | N | N | N | N | N |
| | 3/8 | DAC7 | P | P | P | P | P | P | P | N | N | N | N | N | N | N | N |
| Pre-charge 2 | -3/64 | DAC1 | Based on comparator outputs | | | | | | | P | N | N | N | N | N | N | N |
| | -2/64 | DAC2 | | | | | | | | P | P | N | N | N | N | N | |
| | -1/64 | DAC3 | | | | | | | | P | P | P | N | N | N | N | |
| | 0 | DAC4 | | | | | | | | P | P | P | P | N | N | N | |
| | 1/64 | DAC5 | | | | | | | | P | P | P | P | P | N | N | |
| | 2/64 | DAC6 | | | | | | | | P | P | P | P | P | P | N | N |
| | 3/64 | DAC7 | | | | | | | | P | P | P | P | P | P | P | N |

Figure 5.2.1.2 - Timing diagram of a segmented 6-bit DAC with a 3-voltage switching scheme

For example, if we consider the switching situation depicted in Figure 5.2.1.3, the single-ended output voltages are given by Equations (5.5-5.6), while the differential output voltage can be found in Equation (5.7).

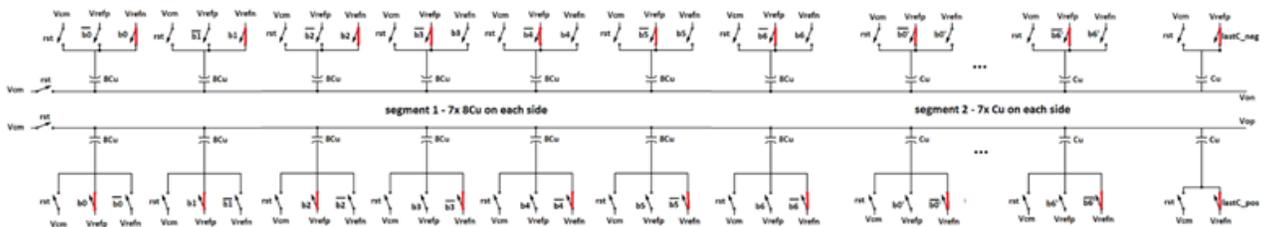


Figure 5.2.1.3 - Example of the switching scheme

$$V_{Op} = 3 * \frac{8C_u}{64C_u} (V_{refp} - V_{CM}) + 4 * \frac{8C_u}{64C_u} (V_{refn} - V_{CM}) + 7 * \frac{C_u}{64C_u} (V_{refn} - V_{CM}) \quad (5.5)$$

$$V_{On} = 3 * \frac{8C_u}{64C_u} (V_{refn} - V_{CM}) + 4 * \frac{8C_u}{64C_u} (V_{refp} - V_{CM}) + 7 * \frac{C_u}{64C_u} (V_{refp} - V_{CM}) \quad (5.6)$$

$$V_{Odiff} = V_{Op} - V_{On} = \frac{V_{refp} - V_{refn}}{64} \quad (5.7)$$

The main disadvantage of the switching scheme described previously is that it requires a very accurate common-mode voltage reference in order to prevent linearity errors from occurring due to the variation of V_{CM} . In order to avoid the need of an accurate V_{CM} , a scheme using only two reference voltages inspired by [12] was explored using the schematic shown in Figure 5.2.1.4. Compared to the previous topology, the last capacitor in the first segment needs to be split into two capacitors with a value of $4C_u$ for reset purposes.

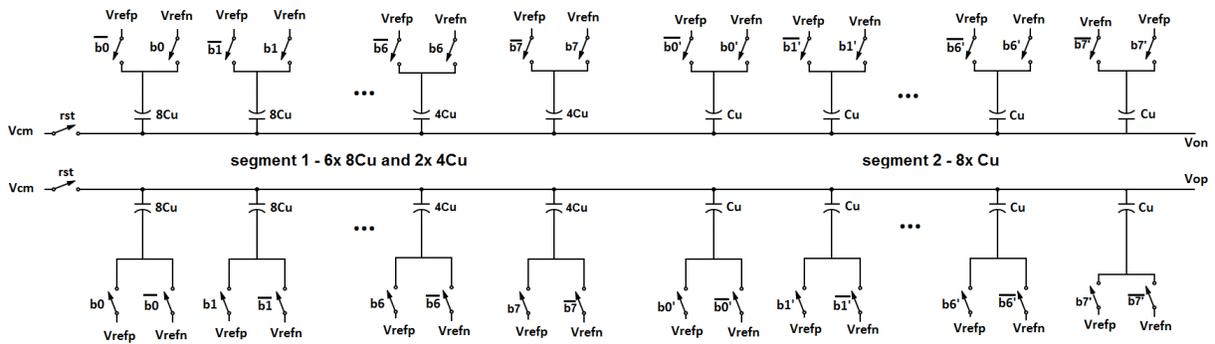


Figure 5.2.1.4 - Segmented 6-bit DAC with a 2-voltage switching scheme

During the reset phase, the top plates of the capacitors are connected to V_{CM} , while half of the bottom plates are switched to V_{refp} and half to V_{refn} . Therefore, half of the DAC capacitance is charged to $V_{CM} - V_{refp}$ and the other half is charged to $V_{CM} - V_{refn}$. In order to accommodate this requirement, the last capacitor in the first segment was split in two equal capacitors such that they can be connected to different references during the reset. The output of the comparator that has to control the $8C_u$ capacitance is connected to both halves after a decision has been made.

During the first pre-charge phase, at most three the bottom plates are switched from V_{refp} to V_{refn} or vice-versa, depending on the particular DAC that is performing this operation. The other bottom plates remain connected to the same voltage as in the reset phase. Due to this fact, less switching energy is required than in the 3-voltage scheme. The

fourth DAC in the ADC does not need to perform any switching during this time, as the reference it needs to provide is of 0V, so it will not consume any power.

After the comparators have delivered their outputs to the DAC, the appropriate reference level is selected as a base for generating the second set of references. A similar switching scheme as in the first segment is used for this purpose, as suggested by Figure 5.2.1.5. **P** (**N**) signifies that the bottom plate of that particular capacitor is connected to V_{refp} (V_{refn}), while \uparrow is used to denote a transition from V_{refp} to V_{refn} and \downarrow indicates switching from V_{refn} to V_{refp} . Using the notations in the previous chapter, the first pre-charge phase is commanded by the track-and-hold clock (clk_{TH}), while the second pre-charge phase begins as soon as the ready signal turns 1 ($rdy < i >$).

| POSITIVE DAC | | Segment 1 | | | | | | | Segment 2 | | | | | | | |
|-----------------|------------|-----------------------------|--------------|--------------|------------|------------|------------|-----|--------------|--------------|--------------|------------|------------|------------|----|----|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Capacitor value | | 8Cu | 8Cu | 8Cu | 8Cu | 8Cu | 8Cu | 4Cu | 4Cu | Cu | Cu | Cu | Cu | Cu | Cu | Cu |
| RESET | | P | P | P | N | N | N | N | P | P | P | P | N | N | N | N |
| Pre-charge 1 | -3/8 DAC1 | \downarrow | \downarrow | \downarrow | - | - | - | - | - | - | - | - | - | - | - | - |
| | -2/8 DAC2 | \downarrow | \downarrow | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | -1/8 DAC3 | \downarrow | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | 0 DAC4 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| | 1/8 DAC5 | - | - | - | \uparrow | - | - | - | - | - | - | - | - | - | - | - |
| | 2/8 DAC6 | - | - | - | \uparrow | \uparrow | - | - | - | - | - | - | - | - | - | - |
| | 3/8 DAC7 | - | - | - | \uparrow | \uparrow | \uparrow | - | - | - | - | - | - | - | - | - |
| Pre-charge 2 | -3/64 DAC1 | Based on comparator outputs | | | | | | | \downarrow | \downarrow | \downarrow | - | - | - | - | - |
| | -2/64 DAC2 | | | | | | | | \downarrow | \downarrow | - | - | - | - | - | |
| | -1/64 DAC3 | | | | | | | | \downarrow | - | - | - | - | - | - | |
| | 0 DAC4 | | | | | | | | - | - | - | - | - | - | - | |
| | 1/64 DAC5 | | | | | | | | - | - | - | \uparrow | - | - | - | |
| | 2/64 DAC6 | | | | | | | | - | - | - | \uparrow | \uparrow | - | - | |
| | 3/64 DAC7 | | | | | | | | - | - | - | \uparrow | \uparrow | \uparrow | - | |

Figure 5.2.1.5 - Two-voltage Switching Scheme

For example, assuming the outputs of the comparator in the first stage are all 0, DAC5 should have the differential output voltage in Equation (5.8) after the second pre-charge phase.

$$V_{DAC5_{precharge2}} = V_{LSB} \tag{5.8}$$

Due to its increased dynamic efficiency and lower complexity, the last switching scheme was chosen for implementation. All the design considerations and simulation results that follow pertain to this DAC version.

5.2.2. Schematic used for simulations

Figure 5.2.2.1 shows the schematic used for simulations.

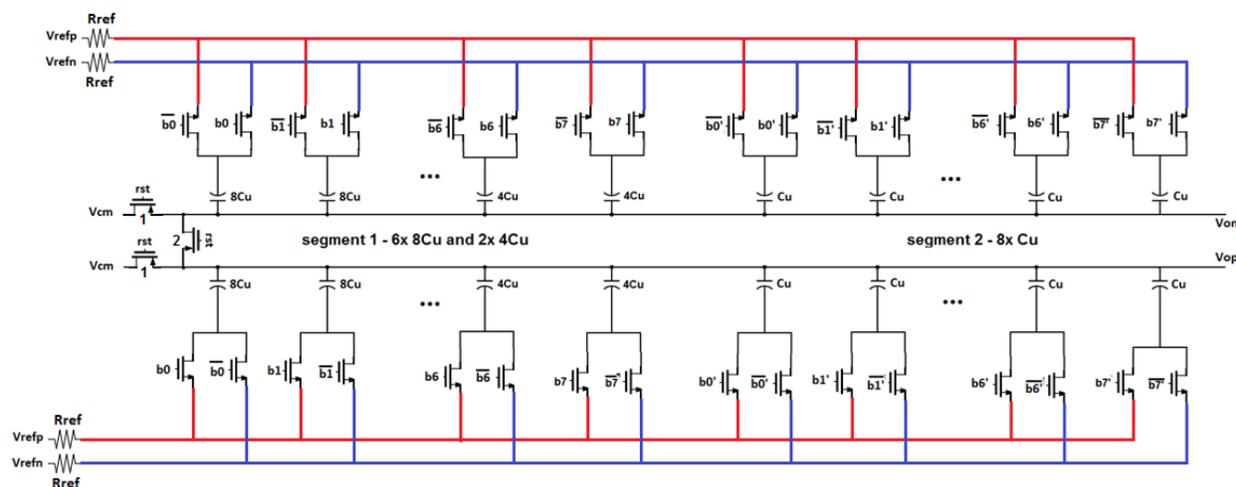


Figure 5.2.2.1 - Transistor-level implementation of the DAC

The total DAC capacitance should be high enough to ensure that the $\frac{kT}{C}$ noise it contributes is less than the quantization noise. For 6-bit accuracy, this translates to a unit capacitance of only 3.17aF, which is too small to be manufactured. If the unit capacitance of 300aF used in the 8-bit reference design, [26], were used, the total DAC capacitance would be $C_{DAC} = 19.2\text{fF}$. In a switching operation of the DAC, the charge sharing between the DAC capacitance and the input capacitance of the preamplifiers results in nonlinear behavior of the DAC, since the input capacitance of the preamplifiers is signal-dependent. In order to mitigate this effect, we adopted an unit capacitor $C_u = 1.7\text{fF}$, resulting in 10% attenuation from the charge sharing. This value was adopted for the simulations that follow.

In order to maximize the switching speed, NMOS switches are used for the bottom plates of the DAC capacitors connected to the reference voltages $V_{refp} = 515\text{mV}$ and $V_{refn} = 0\text{V}$, as shown in Figure 5.2.2.1. For the reset switches, PMOS transistors are chosen because of the common-mode voltage $V_{CM} = 700\text{mV}$. Aside from the two in-line switches, a differential switch controlled by the same reset signal is used to speed-up the reset by ensuring the two differential voltages are equal to each other.

5.2.3. Control Logic

The DAC architecture chosen allows the output of the comparator to be directly connected to the switches at the bottom plates of the capacitor that needs to be controlled. A D-flip-flop with set and reset inputs is sufficient for this task.

Figure 5.2.3.1 shows an example of how the switching is controlled for the first half of the DAC ($8C_u$ or $4C_u$ capacitors); the signal names correspond to those from the previous chapter. When the DAC is reset, the bottom plate of that particular capacitor needs to be connected to V_{refn} . This operation is triggered by $therst_DAC_logic$. When the references for the first set of comparisons need to be generated, the track-and-hold clock (clk_TH) connects the bottom plate of the capacitor to V_{refp} using the set input. The signals rst_DAC and clk_TH are non-overlapping to ensure that only one of the switches connected to the bottom plate is active at once. After the comparator has made its decision and its output is available after having been stored in an SR-latch ($cmp_out < i >$), the ready signal associated to it triggers the D-flip-flop and connects the capacitor to the appropriate reference. The generation of the second set of references on top of the interval selected is triggered by the same event, as shown in Figure 5.2.3.2.

Depending on the bit that is in use, the set and reset inputs may be connected to different signals, according to the reference that needs to be prepared before the next comparison.

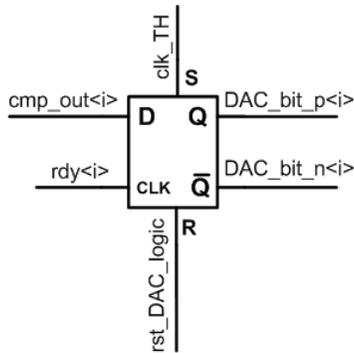


Figure 5.2.3.1 - Example of DAC logic for $8C_u$ (or $4C_u$)

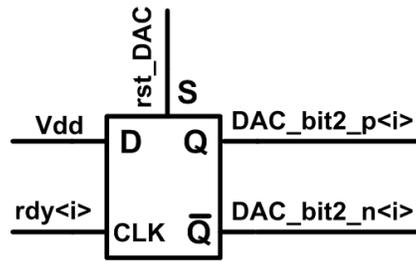


Figure 5.2.3.2 - Example of DAC logic for C_u

5.2.4. Study of the Reset and Settling Times of the DAC

From the operation of the ADC, it can be derived that:

$$t_{sample} = t_{DAC_{reset}} + t_{track} + 2(t_{comp} + t_{pre}) + t_{DAC_{MSB}} + t_{logic}, \quad (5.9)$$

where $t_{DAC_{reset}} = t_{track} = 50\text{ps}$ and represent the time needed to reset the DAC and the tracking time, $t_{comp} + t_{pre} = 94\text{ps}$ and represent the delay of the comparator and preamplifier, $t_{logic} = 120\text{ps}$ represents the logic delay, while $t_{DAC_{MSB}}$ is used to denote the settling time of the MSB capacitor of the DAC.

Taking into account the assumptions made previously, the influence of the DAC settling time on the maximum sampling frequency is the one shown in Figure 5.2.4.1. It is clear that a long settling time decreases the speed of the ADC significantly, so it is important that the switches are sized such that the target sampling frequency is still met. In order to achieve an operating frequency of 2.5GS/s, the DAC settling time should be lower than 100ps. This value is used as a guideline in designing the DAC.

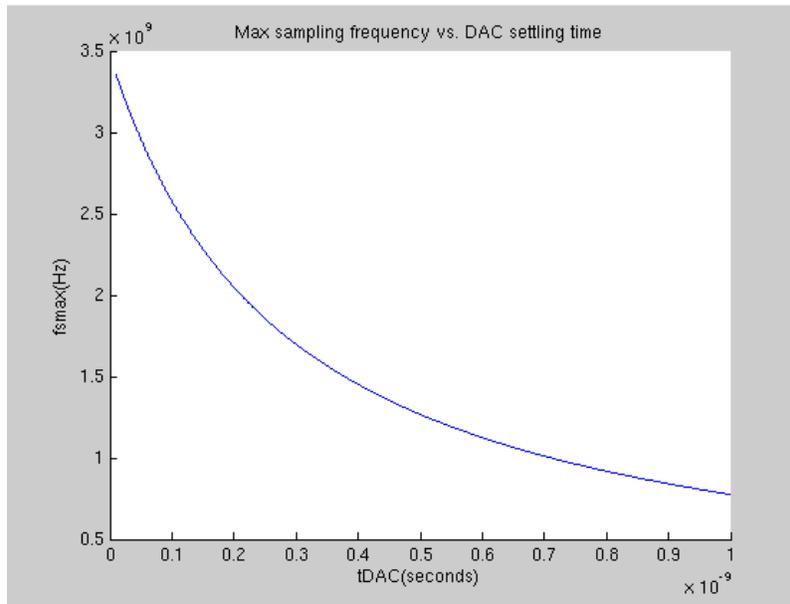


Figure 5.2.4.1 - Maximum sampling frequency vs. DAC settling time

In order to be able to size the DAC switches such that they lead to the desired reset and settling time, a mathematical analysis of these times is useful in determining which parameters impact these most. Despite the fact that the NMOS transistors used as switches are non-linear elements, their behavior can be approximated by their on-resistance as soon as switching has occurred. For the accuracy of a hand calculation, it is safe to consider that the on-resistance remains constant regardless of the input signal, which is the working assumption used throughout the sections that follow.

The previous assumption means that the DAC can be regarded as a linear, time-invariant circuit, and that as such, its step response can be determined using the Laplace transform [38].

Assuming that the time constants of the DAC branches are equal (which is true if the transistors connected to the bottom plates are sized in accordance to the capacitors they drive), each side of the DAC will have N such branches in parallel during the reset time (as shown on the left side of Figure 5.2.4.2), which can be formally replaced by the simpler branch shown on the right side of Figure 5.2.4.2. For the analysis on the settling time, the

segmented DAC can be treated as if it were a unary DAC with 64 unit capacitors driven by distinct transistors, which greatly simplifies the calculations.

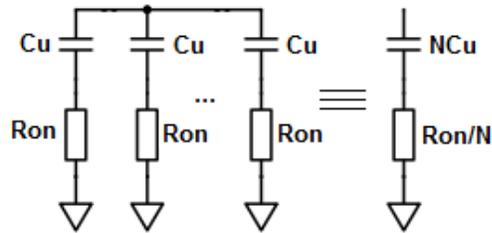


Figure 5.2.4.2 - Equivalent circuit for N identical DAC branches

5.2.5. Reset Time

Owing to the asynchronous signal generation, the reset of the DAC cannot be performed in parallel with another operation, so it impacts the sampling frequency directly and should be minimized accordingly. In order to avoid using a control pulse that is too short, a target reset time of 50ps is assumed.

Figure 5.2.5.1 shows the equivalent circuit that can be used to determine the reset time of the DAC. The circuit elements that influence it, together with their notations and the expected numerical value are listed in Table 5.2.5.1. Several iterations were necessary to size the DAC switches, but for the sake of clarity only the final values will be presented here. The assumptions in Table 5.2.5.1 correspond to the final sizes of the switches and they are used in the hand calculations in order to prove that the results match the simulation data.

Table 5.2.5.1 – Summary of notations used to analyze the reset time of the DAC

| Notation | Significance | Expected numerical value |
|-----------|--|--------------------------|
| R_{on1} | on-resistance of the in-line PMOS switch | 213 Ω |
| R_{on2} | on-resistance of the differential PMOS switch | 175 Ω |
| R_{on} | resistance of an NMOS bottom plate switch | 488.5 Ω |
| C_{p1} | parasitic capacitance associated with the in-line PMOS switch | 4.9 fF |
| C_{p2} | parasitic capacitance associated with the differential PMOS switch | 5.96 fF |
| C_p | parasitic capacitance associated with the NMOS bottom plate switch | 2.3 fF |
| C_u | unit capacitor used in the DAC | 1.7 fF |
| C_{pre} | input capacitance of the preamplifier | 20.3 fF |
| R_{ref} | resistance of the V_{CM} buffer | 5 Ω |

During the reset time, half of the bottom plates of the capacitors on each side are connected to V_{refp} , while the others are connected to V_{refn} . The top-plates of the capacitors are switched to V_{CM} and it is expected that each of the differential outputs will reach this value after the settling time has elapsed.

This situation is nothing other than a step-response in which the input voltage is a step with an amplitude equal to V_{CM} within a given tolerance, instead of the traditional Heaviside unit step and thus can be analyzed with the mathematical method explained previously.

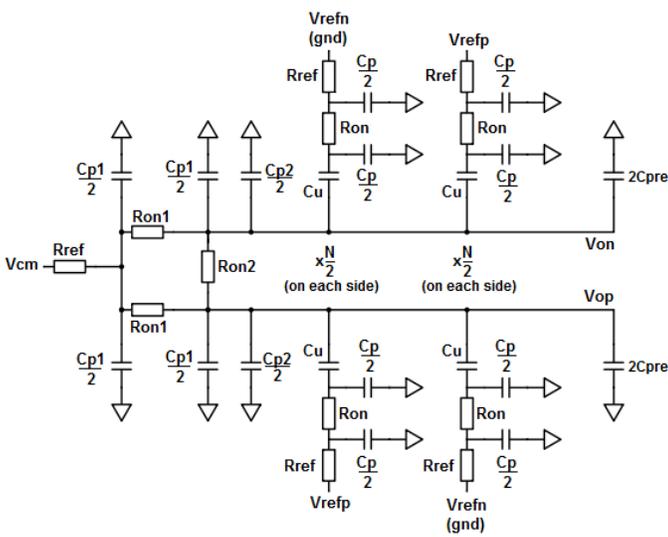


Figure 5.2.5.1 - Reset time – complete equivalent circuit

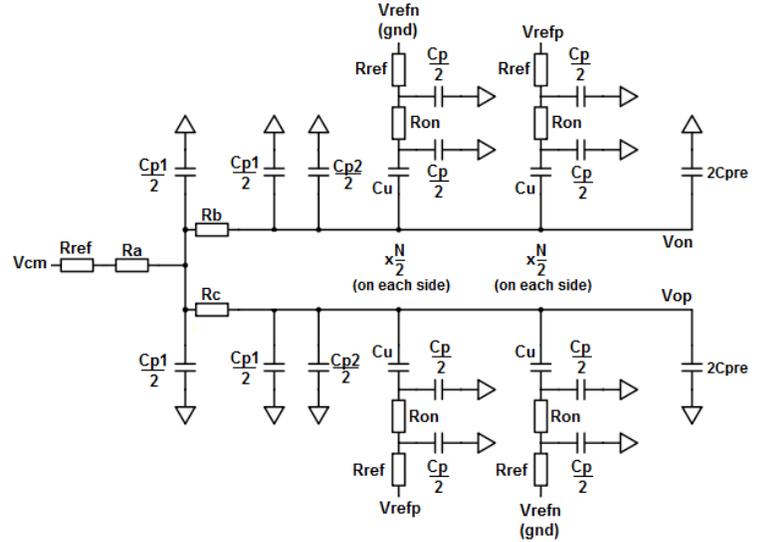


Figure 5.2.5.2 - Reset time - intermediate equivalent circuit

With respect to the V_{CM} switching, it is as if the bottom plates of the capacitors were connected to ground. Connecting half of the DAC capacitance to V_{refp} on each side will impact the output waveforms, but this effect can be overlooked during the hand calculations, as will be shown later.

Using the well-known $Y - \Delta$ network transform, the combination of the three on-resistances of the reset switches ($2xR_{o1}$ and R_{o2}) from Figure 5.2.5.1 can be replaced by their Y-equivalent (Figure 5.2.5.2) using (5.10).

$$\begin{cases} R_a = \frac{R_{on1}^2}{2R_{on1} + R_{on2}} \\ R_b = \frac{R_{on1}R_{on2}}{2R_{on1} + R_{on2}} \\ R_c = \frac{R_{on1}R_{on2}}{2R_{on1} + R_{on2}} \end{cases} \quad (5.10)$$

The circuit in Figure 5.2.5.2 is still not a trivial one to analyze, so it is useful to use the expected numerical values in Table 5.2.5.1 as a guideline for neglecting the elements that do not have a great impact on the settling time in order to simplify the existing configuration. For this purpose, the circuit in Figure 5.2.5.2 was simulated using ideal resistors and capacitors with various components omitted.

Several different incremental models were compared:

- The complete equivalent circuit (Figure 5.2.5.2)
- The model in Figure 5.2.5.2 without buffer resistors (R_{ref})
- The model in Figure 5.2.5.2 without R_{ref} and without the parasitic capacitance of the differential reset switch (C_{p2})
- The model in Figure 5.2.5.2 without R_{ref}, C_{p2} and without the parasitic capacitances of the in-line reset switches (C_{p1})
- The model in Figure 5.2.5.2 without R_{ref}, C_{p2}, C_{p1} and without the input capacitance of the preamplifier (C_{pre})

The positive single-ended output waveforms of the different circuits considered can be seen in Figure 5.2.5.3. It is easily observable that little variation exists between the models, which means that R_{ref}, C_{p2}, C_{p1} can be safely ignored for hand calculation purposes. Ignoring C_{pre} produces a slightly different waveform, so this element will be included in the simplified model. At the same time with the V_{CM} switching, half the capacitors are connected to V_{refp} . Depending on the values of the output voltages before the reset command is given to the DAC, this effect may help reduce the settling time. In order to derive the settling time mathematically, V_{refp} is passivized and the only switching that occurs is with respect to V_{CM} .

The reset time of the DAC can be defined as the time that needs to pass from the switching moment until the moment when the output has settled to a voltage within the range of V_{CM} that is of sufficient accuracy for the application. In this case, an error of at most $\frac{V_{LSB}}{6}$ ensures that the total error stemming from incomplete reference settling and the residual error from the reset is below V_{LSB} .

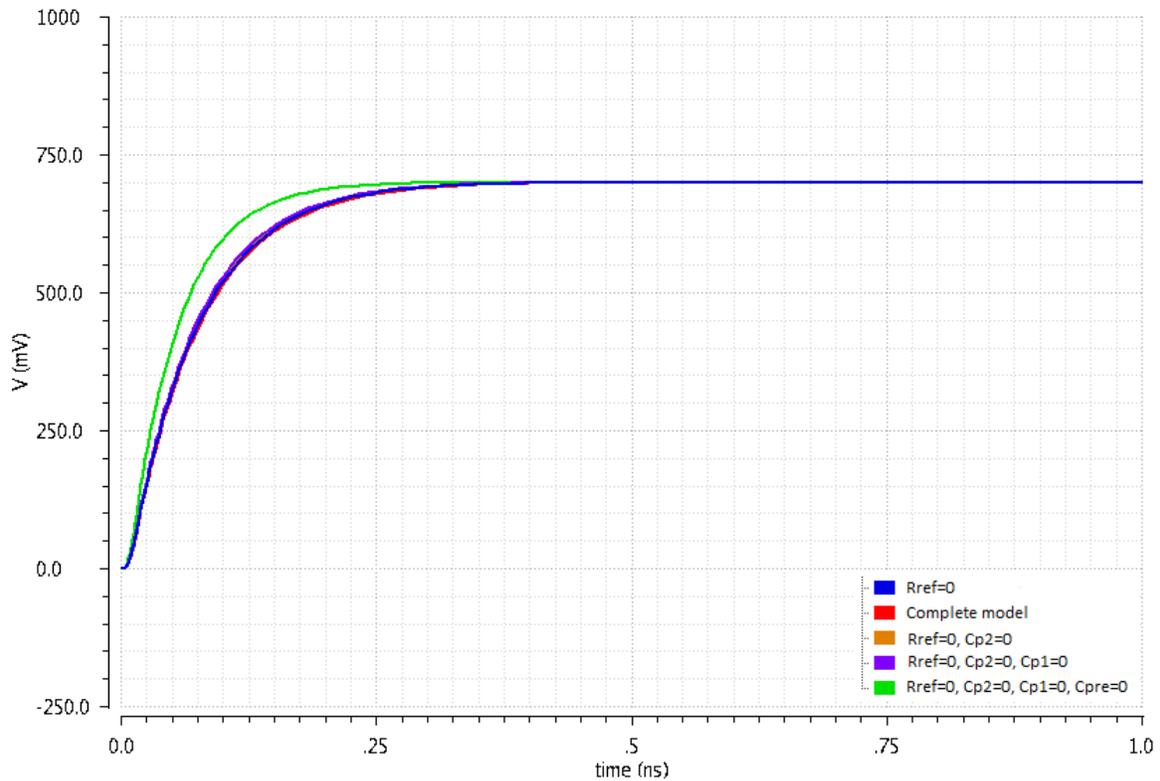


Figure 5.2.5.3 - Comparison of the single-ended reset waveforms for models of different complexity

It can be concluded that the simplified model in Figure 5.2.5.4, where V_{refp} is replaced by a ground connection, predicts the reset time with sufficient accuracy and in what follows, its step-response will be determined using the Laplace transform.

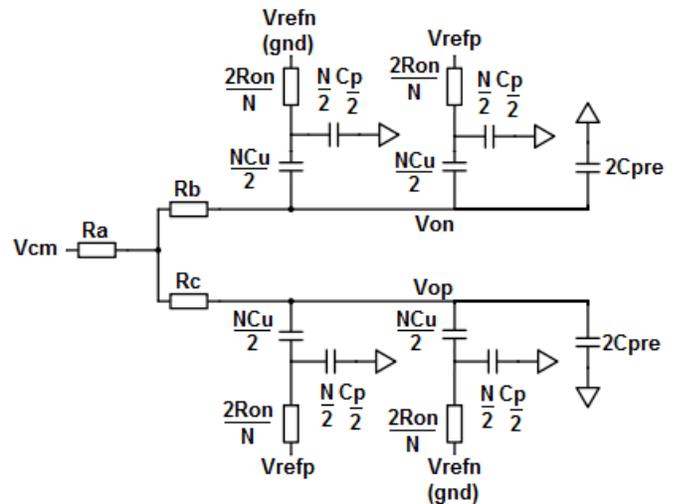


Figure 5.2.5.4 - Reset time - simplified model

In order to obtain the transfer function corresponding to the differential voltage, the ratio of $\frac{V_{op}}{V_{on}}$ is computed when $V_{CM} = 0$. Considering the notations in Equation (5.11), the result can be found in Equation (5.12).

$$\left\{ \begin{array}{l} G_4 = \frac{1}{R_a} \\ G_3 = \frac{1}{R_b} \\ G_2 = \frac{N}{2R_{on}} \end{array} \right. \text{ and } \left\{ \begin{array}{l} C_2 = \frac{N}{2}C_u \\ C_1 = 2C_{pre} \end{array} \right. \quad (5.11)$$

$$H(s) = V_{step} \frac{G_3^2 C_2 s + G_3^2 G_2}{s^2 C_1 C_2 (2G_3 + G_4) + s(C_2 G_3 (G_3 + 4G_2) + C_1 G_2 (2G_3 + G_4) + C_2 G_4 (2G_2 + G_3)) + G_2 G_3 (G_3 + G_4)} \quad (5.12)$$

The time needed for the differential output to settle with an accuracy of half of an LSB is $t_{settle_{formula}} = 47\text{ps}$, while the waveform from the Cadence simulation exhibits a settling time of $t_{settle_{transistor}} = 53.6\text{ps}$. The differences are accounted for by the signal-dependent variation of the on-resistances of the switches and by the elements that were neglected for the hand calculations. The sizing of the reset switches that corresponds to the assumptions for the resistances and capacitances is found to be appropriate to produce a reset time of about 50ps.

5.2.6. Reference Generation Settling Time

For the sake of clarity, the settling time analysis will be performed first using the simplified model in Figure 5.2.6.1, where both the parasitic capacitances of the switches and the input capacitance of the preamplifiers have been omitted. Let n denote the number of unit capacitors that are switched to V_{refp} , and N the total number of unit capacitors in the DAC. It follows that $(N - m)$ represents the number of unit capacitors that are connected to V_{refn} at that particular switching moment.

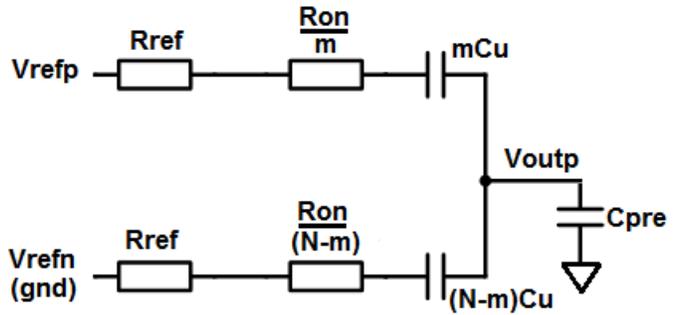


Figure 5.2.6.1 - Reference generation - simple model

The purpose of the ensuing computations is the study of the settling time of the output voltage (V_{outp}) if the input voltage is V_{refp} . This is the case for the positive side of the DAC,

but a similar formula can be derived for the negative single-ended voltage, seeing that the same number of unit capacitors will be switched on that side. The difference between the two resulting output voltages is only in terms of the amplitude of the step, as the configuration of the circuit is similar.

Let Z_1 denote the impedance between the node where V_{refp} is connected and the output, and Z_2 denote the impedance seen from V_{refp} to the output. It can be shown that:

$$\begin{cases} Z_1 = R_{ref} + \frac{R_{on}}{m} + \frac{1}{smC_u} \\ Z_2 = R_{ref} + \frac{R_{on}}{(N-m)} + \frac{1}{s(N-m)C_u} \end{cases} \quad (5.13)$$

The transfer function is depends on the two impedances and can be written as:

$$H(s) = \frac{Z_2}{Z_1 + Z_2} = \frac{sC_u m[(N-m)R_{ref} + R_{on}] + m}{sC_u [2m(N-m)R_{ref} + R_{on}N] + N} \quad (5.14)$$

Using partial fraction expansion on the output signal $Y(s)$ can be simplified and its time-domain equivalent can be computed:

$$y(t) = V_{step}u(t) \left[\frac{m}{N} + C_u(N-m)mR_{ref} \left[1 - \frac{2m}{N} \right] e^{-\frac{N}{C_u[2m(N-m)R_r + R_{on}N]}t} \right] \quad (5.15)$$

If the values in Table 5.2.5.1, together with the assumption that $m = 32$ are used in Equation (5.15), the output signal is:

$$y_{differential}(t) = y(t) - y_n(t) = 0.5(V_{refp} - V_{refn})u(t), \quad (5.16)$$

where:

$$y(t) = 0.5V_{refp}u(t) \quad (5.17)$$

$$y_n(t) = -0.5V_{refn}u(t) \quad (5.18)$$

Bearing in mind that $y(t)$ represents the single-ended output voltage, and that conversely, for the negative side of the DAC Equation (5.18) holds, this result is correct. Due

to the capacitive division, it is expected that when half of the DAC capacitance is switched, the differential reference voltage is halved (Equation (5.16)), just as the formula predicts.

In what follows, the complete model of the reference generation will be analyzed using the same technique. Figure 5.2.6.2 shows the complete equivalent model for the settling time of the DAC when it operates in reference-generation mode. The same notations and values in Table 5.2.5.1 are used for this calculation.

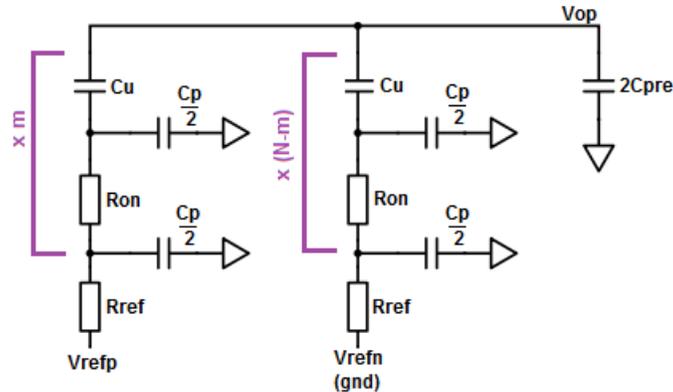


Figure 5.2.6.2 - Reference generation - complex model

As in the case with the reset time, the possibilities of neglecting some of the circuit elements for the scope of the hand calculations were investigated first by using ideal passive components.

Several models were considered:

- The complex model in Figure 5.2.6.2
- The complex model without the resistance of the reference buffers (R_{ref})
- The complex model without R_{ref} and without the parasitic capacitance of the switches - C_p (Figure 5.2.6.3)

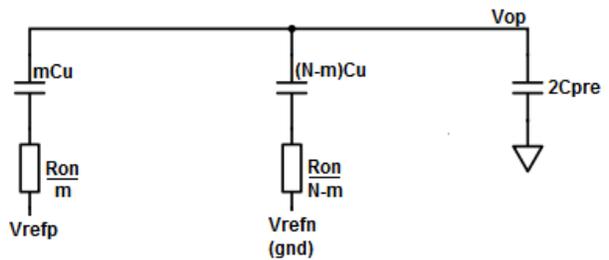
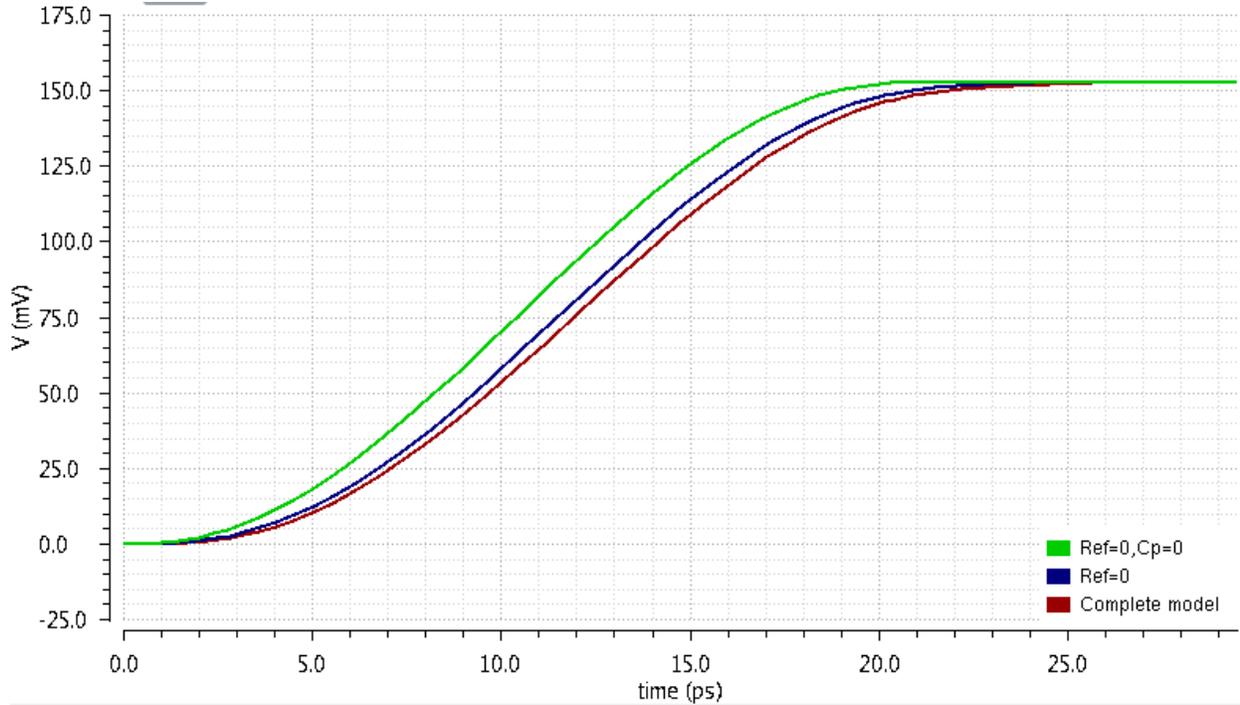


Figure 5.2.6.3 - Reference settling (simple model)

Figure 5.2.6.4 shows that all the models considered render similar accuracy levels, so the simplified one in Figure 5.2.6.3 is used for the hand calculations.


 Figure 5.2.6.4 - Model comparison for reference settling ($m=32$)

The transfer function of the circuit with respect to V_{refp} is given by Equation (5.19) using the notations in Equation (5.20):

$$H(s) = \frac{sC_6C_5R_6 + C_5}{s^2C_5C_6C_7R_5R_6 + s[C_5R_5(C_6 + C_7) + C_6C_7R_6 + C_5C_6R_6] + C_5 + C_6 + C_7}, \quad (5.19)$$

$$\text{where } \begin{cases} R_5 = \frac{R_{on}}{m} \\ C_5 = mC_u \\ R_6 = \frac{R_{on}}{N-m} \\ C_6 = (N-m)C_u \\ C_7 = 2C_{pre} \end{cases} \quad (5.20)$$

The calculations reveal a settling time of $t_{settle} = 18\text{ps}$ for the MSB ($m = 32$), while the transistor-level simulation points to a settling time of $t_{settle} = 23.5\text{ps}$. The difference stems mainly from the nonlinearity of the on-resistance of the switches, but the fact that several elements of the complete model were neglected also explains part of the difference.

Figure 5.2.6.5 shows the differential waveforms corresponding to different codes (modelled by a different value for the parameter m), while Table 5.2.6.1 summarizes the settling times of each of them. As expected, the MSB settling takes the longest time.

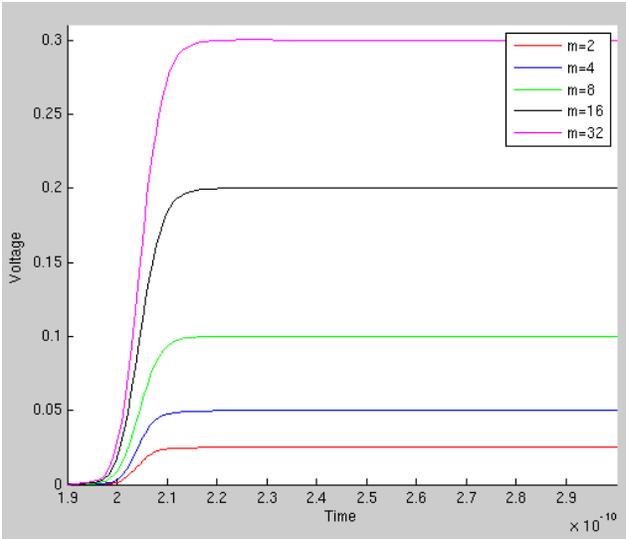


Figure 5.2.6.5 - Waveforms for various codes (transistor-simulation)

Table 5.2.6.1 - Settling time for different codes

| m | Settling time (ps) |
|----|--------------------|
| 2 | 16.1 |
| 4 | 17.8 |
| 8 | 20.3 |
| 16 | 22.4 |
| 32 | 23.5 |

5.2.7. Design Procedure and Sizing

The capacitors were sized according to noise constraints, as was explained at the beginning of the chapter, and their values were small enough not to constitute a problem for the settling time. An on-resistance of around 500Ω proved to be sufficient to achieve the desired speed.

The MSB settling time of the DAC is the longest because half of the total DAC capacitance is switched, so it is useful to examine how the size of the bottom-plate switch affects it. Let W_8 and L_8 represent the width and length of the NMOS switch that controls a capacitance with a value of $8C_u$. The transistors are always sized with minimum length, such that their on-resistance is minimized ($L_8 = L_{min} = 40\text{nm}$). Opting for different values for the widths of the NMOS switches reveals that a value of $W_8 = 8\mu\text{m}$ gives a settling time of around 20ps (Figure 5.2.7.1). The other switches (characterized by W_4, L_4 for a $4C_u$ capacitor and W_1 and L_1 for a C_u capacitor, respectively), are sized proportionally to the capacitor they control:

$$\begin{cases} W_8 = 8\mu\text{m}, L_8 = L_{min} = 40\text{nm} \\ W_4 = \frac{W_8}{2} = 4\mu\text{m}, L_4 = L_8 = 40\text{nm} \\ W_1 = \frac{W_8}{8} = 1\mu\text{m}, L_1 = L_8 = 40\text{nm} \end{cases} \quad (5.21)$$

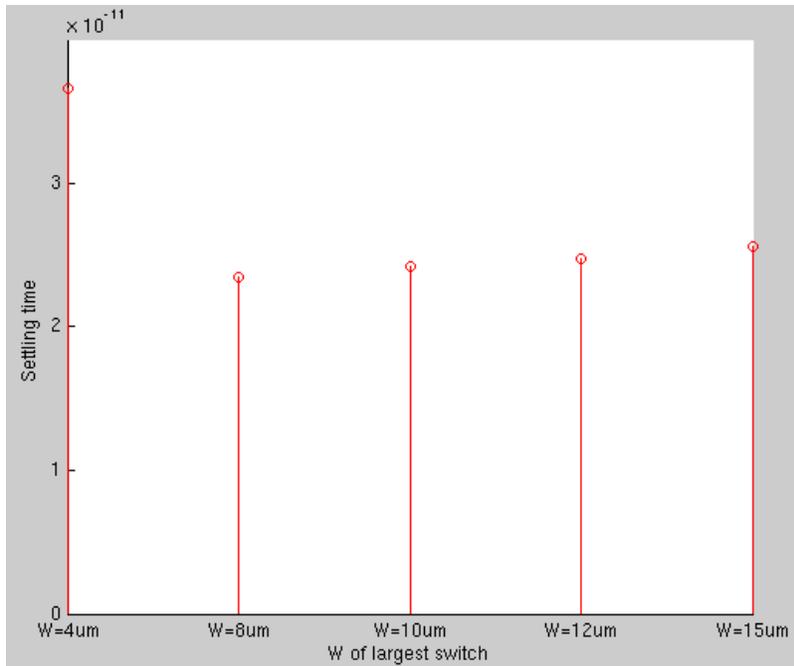


Figure 5.2.7.1 - Settling time for different switch sizes ($m=32$)

The common-mode voltage is equal to the input common-mode voltage and thus fixed to 700mV. The reasons for this choice will be explained in the section devoted to the design of the preamplifier and the comparator. This prompts the choice of a PMOS transistor for the reset switches, which can be operated with a larger overdrive voltage when the supply is 1.1V. An aspect ratio of $15\mu\text{m}/40\text{nm}$ for the in-line reset switch, together with a differential switch with a size of $18\mu\text{m}/40\text{nm}$ ensure the on-resistances considered in Table 5.2.5.1 are indeed achieved. The differential switch needs to be larger than the other ones in order to shorten the reset time by bringing the outputs to the same voltage faster.

The final reset time is of 50ps and the settling time is of 23.5ps.

5.3. Comparator and Preamplifier

The architecture study revealed that the comparator is the critical block of the design, as it can easily reduce the speed of the ADC. This section aims to present the design of a StrongArm comparator and the preamplifier preceding it and to expose the trade-offs that allow a maximization of its performance.

5.3.1. Comparator

The design of the comparator (Figure 5.3.1.1), preamplifier (Figure 5.3.5.1) and biasing block (Figure 5.3.5.2) that will be described in this section was performed in [26] using the same 40nm CMOS technology, but with a supply of 1.2V. For this design, $V_{DD} = 1.1V$. All three circuits were characterized with the new supply voltage, as it was expected that the speed will decrease slightly and adaptations were made where necessary.

The Strong-Arm comparator, first introduced by Kobayashi et al., [39], has gained popularity due to its high-speed, low-power operation. The schematic depicted in Figure 5.3.1.1 is useful for understanding its operating principle.

During the reset phase ($clk = 0$), PMOS transistors $M_{r1} - M_{r2}$ pull the output nodes to V_{DD} such that every comparison begins from a well-defined state. Transistors $M_{r3} - M_{r4}$ serve a similar function and are used to speed up the reset time by setting the intermediate nodes to V_{DD} by charging the parasitic capacitance associated to nodes A and B.

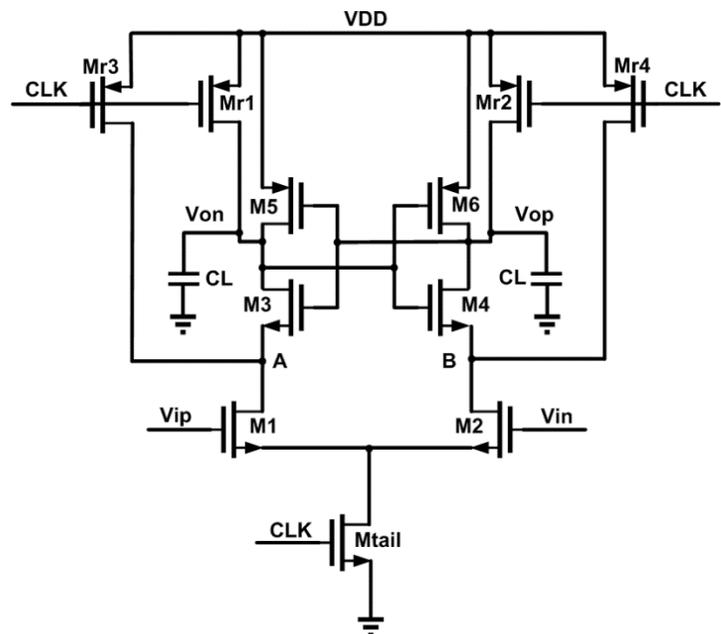


Figure 5.3.1.1 - Transistor-level implementation of the Strong-Arm comparator

The tail transistor (M_{tail}) is turned off and makes the output insensitive to the input signal.

During the comparison phase ($clk = 1$), the reset transistors ($M_{r1} - M_{r4}$) are off and the tail transistor is turned on. The differential input signal causes a difference in the currents flowing through the transistors in the input pair, which begin to discharge the

intermediate nodes A and B . As soon as these nodes reach a voltage that is one NMOS-threshold voltage (V_{Tn}) lower than the supply, transistors M_3 - M_4 turn on and begin to discharge the output nodes. As soon as the output nodes are one PMOS-threshold voltage (V_{Tp}) lower than V_{DD} , transistors M_5 – M_6 turn on and the latch begins to regenerate the signal, which causes either V_{op} or V_{on} to go to V_{DD} , while the other one will be discharged to ground. Neither of the phases features a direct path between the two supply rails, which ensures a low power consumption level.

5.3.2. Speed

The time that needs to elapse until one of the output nodes is discharged to $V_{DD} - V_{Tp}$ is called the *integration time* and is denoted by t_{int} . The time required to reach a decision from the moment the latch turns on is referred to as the *regeneration time* and is denoted by t_{reg} . The total comparator delay depends on the sum of these two times:

$$t_{comp} = t_{int} + t_{reg} \quad (5.22)$$

Wicht et al., [27] , have shown that these can be modelled as:

$$t_{int} = \frac{C_L V_{Tp}}{I_D} \quad (5.23)$$

$$t_{reg} = \frac{C_L}{g_{m,eff}} \ln\left(\frac{V_{out}}{G \cdot \Delta V_{in}}\right) \quad (5.24)$$

where C_L represents the load capacitor of either output of the comparator, I_D is the drain current flowing through one of the transistors in the input pair with zero differential input, $g_{m,eff}$ is the effective transconductance of the latch, G is the amplification of the comparator, V_{out} represents a valid logic level for the output and ΔV_{in} is the input differential voltage. The term $\frac{C_L}{g_{m,eff}}$ is usually referred to as the regeneration time constant (τ_{reg}).

During the architecture study it was assumed that the minimum input voltage acceptable is $\Delta V_{in} = \frac{V_{LSB}}{4}$ and that an inverter with a delay of 30ps is placed between the output of the comparator and the subsequent block, which leads to a minimum valid output voltage of $V_{out} = \frac{V_{DD}}{2}$.

Equations (5.23)-(5.24) show that a high-speed comparator design relies on having a small output capacitance, a large bias current, as well as a large amplification. The input common mode voltage influences the current in the comparator. A higher V_{CM} translates to

a larger current, which is beneficial for the comparator speed, as the research in [27] demonstrates. Figure 5.3.2.2 illustrates this for an input signal of $\Delta V_{in} = \frac{V_{LSB}}{4}$ and confirms the dependence. The same figure is useful for determining the regeneration time constant, the gain of the comparator and the integration time.

The integration time represents the time needed for the comparator to reach a valid output in case a large signal is applied at its input. Plotting the delay vs. the input signal (Figure 5.3.2.1) shows that for high input signals there is an asymptotic part of the curve that corresponds to this time. In the case of this design, $t_{int} = 23.5\text{ps}$. To determine the regeneration time constant, two different points on the curve need to be considered and substituted in Equation (5.25) together with the input voltages they correspond to. This particular design exhibits a $\tau_{reg} = 10.7\text{ps}$. Once these two parameters are known, the gain of the comparator is derived from Equation (5.22) for a particular input signal (in this case $\frac{V_{LSB}}{4}$) and is equal to $G = 4.3$.

$$t_{delay1} - t_{delay2} = \tau_{reg} \ln\left(\frac{V_{in2}}{V_{in1}}\right) \quad (5.25)$$

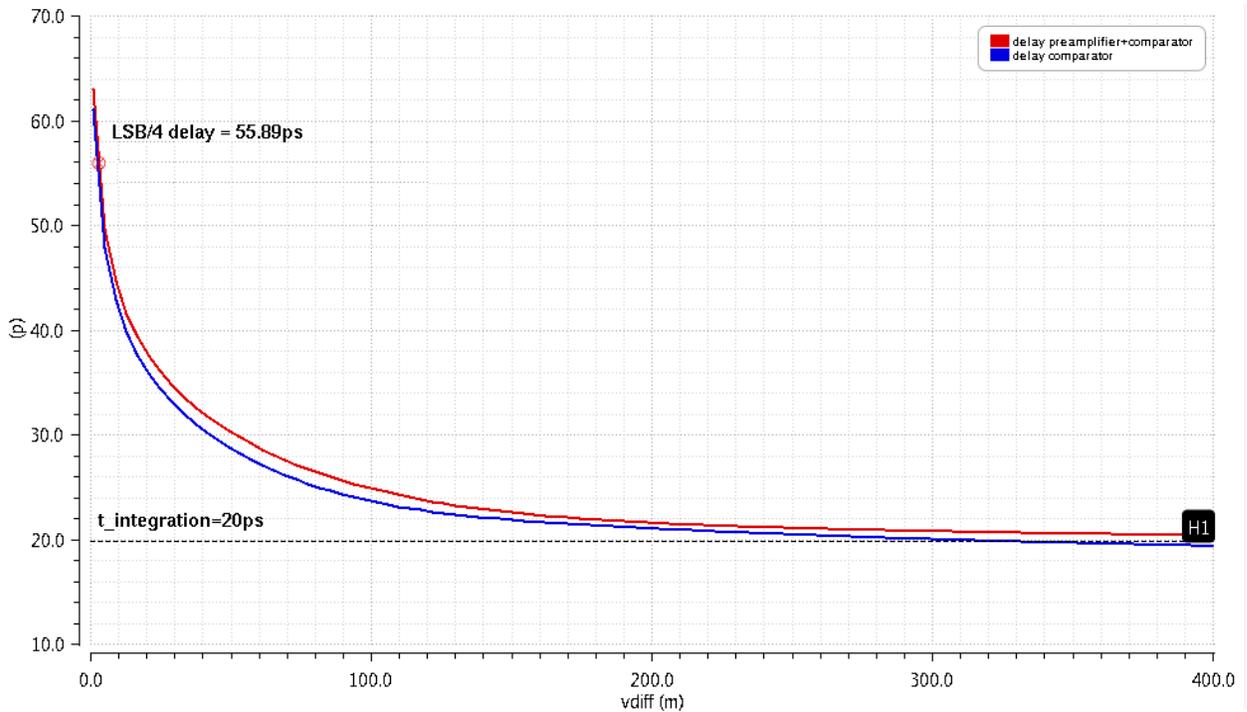


Figure 5.3.2.1 - Comparator delay vs. input signal

The same paper, [27], shows that the speed of the comparator is highly dependent on the input common mode level, which is confirmed by the simulation results (Figure 5.3.2.2). The adopted common-mode voltage is of 770mV, which coincides with the output common-mode of the preamplifier.

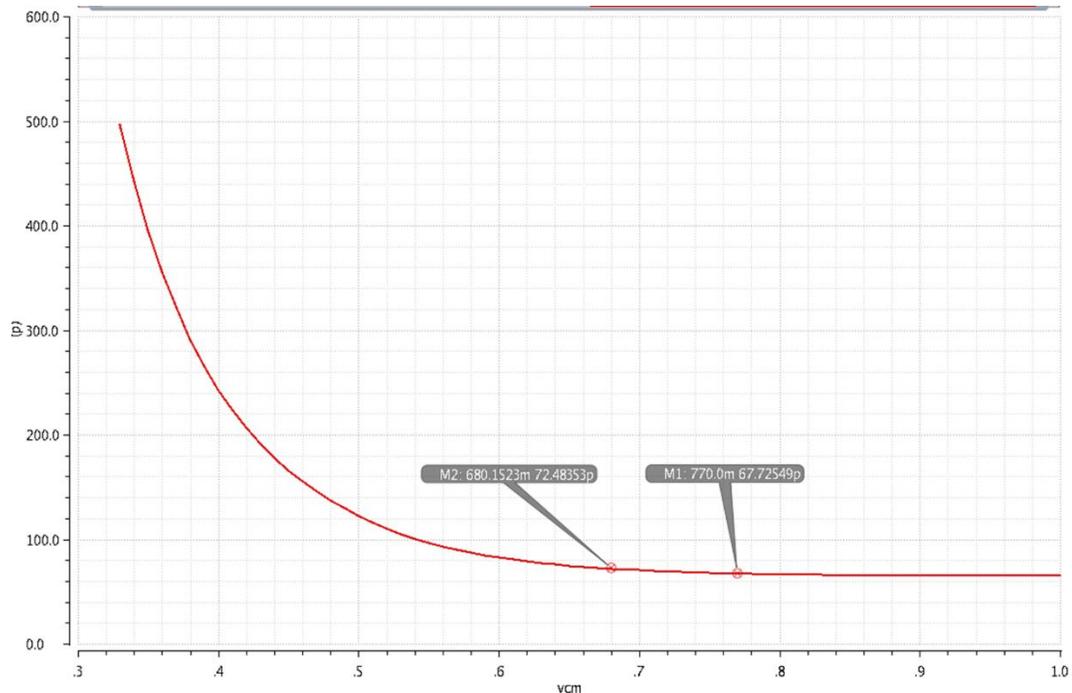


Figure 5.3.2.2 - Comparator delay vs. input common mode voltage ($V_{in} = \frac{V_{LSB}}{4}$)

5.3.3. Noise

The input-referred noise of the comparator needs to be kept as low as possible, seeing that a high noise level can introduce significant errors in the decision of the block. Nuzzo et al., [40], have established that the input-referred noise can be decreased by choosing the aspect ratio of the input pair such that its overdrive voltage is kept as low as possible, for example by lowering the input common-mode voltage. This requirement is in contradiction with the requirement for high-speed operation, so a trade-off exists.

Figure 5.3.3.1 shows the input-referred noise of the comparator for the minimum input voltage ($\Delta V_{in} = \frac{V_{LSB}}{4}$), simulated using a periodic noise analysis and leads to the same conclusion. An input common-mode voltage of $V_{CM,in} = 770\text{mV}$ was chosen as the best compromise between speed and noise. At this common-mode level, an input-referred noise of about 1mV is expected, which is about 3 times lower than one tenth of an LSB.

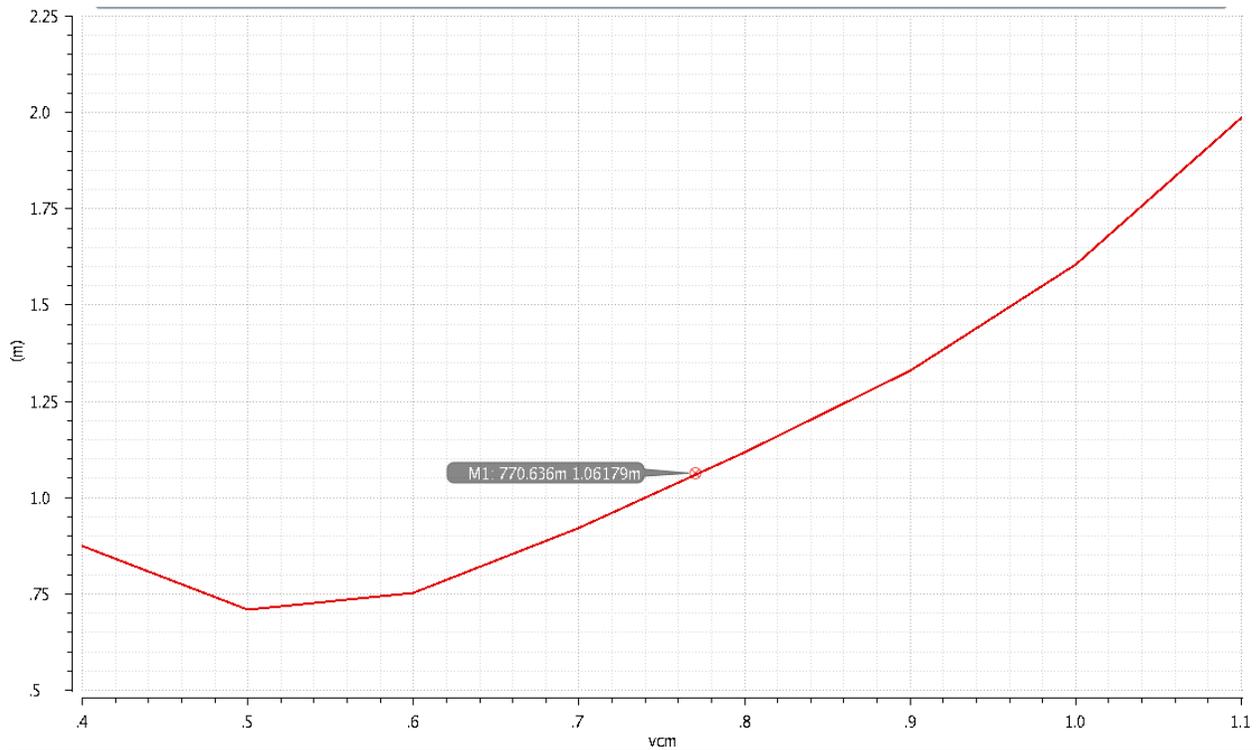


Figure 5.3.3.1 – Input-referred noise of the comparator vs. input common-mode voltage ($\Delta V_{in} = \frac{V_{LSB}}{4}$)

5.3.4. Reset time

The reset time of the comparator is not in the critical path of the ADC because of the 3-bit/cycle resolving scheme, which allows one set of comparators to stay in a long reset mode while the other set is making its decision. Failing to allocate enough time for the reset of the comparators may introduce errors in the output because of a residual voltage from the previous comparison. In order to maintain the desired accuracy level, it is important to reduce the input-referred residual voltage to well below $\frac{V_{LSB}}{4}$. If the reset time is defined as the time that needs to elapse from the negative edge of the comparator clock until the differential output of the comparator has reached 0 with an accuracy of less than $\frac{V_{LSB}}{4}$, then the reset time for this design is of $t_{reset} = 50\text{ps}$. The worst case residual error can occur when a very large input signal is followed by one comparable to the LSB level. This condition can be simulated using the overdrive-recovery test. By running this simulation for different reset times, the residual voltage at the output of the comparator can be determined for different reset times (Figure 5.3.4.1). The plot shows that if the reset time is increased over 70ps, the residual voltage no longer decreases. Owing to the fact that the two comparator sets operate in parallel, the reset time for each is longer than this, which confirms that no

errors will be introduced in the conversion due to the residual voltage from the previous comparison.

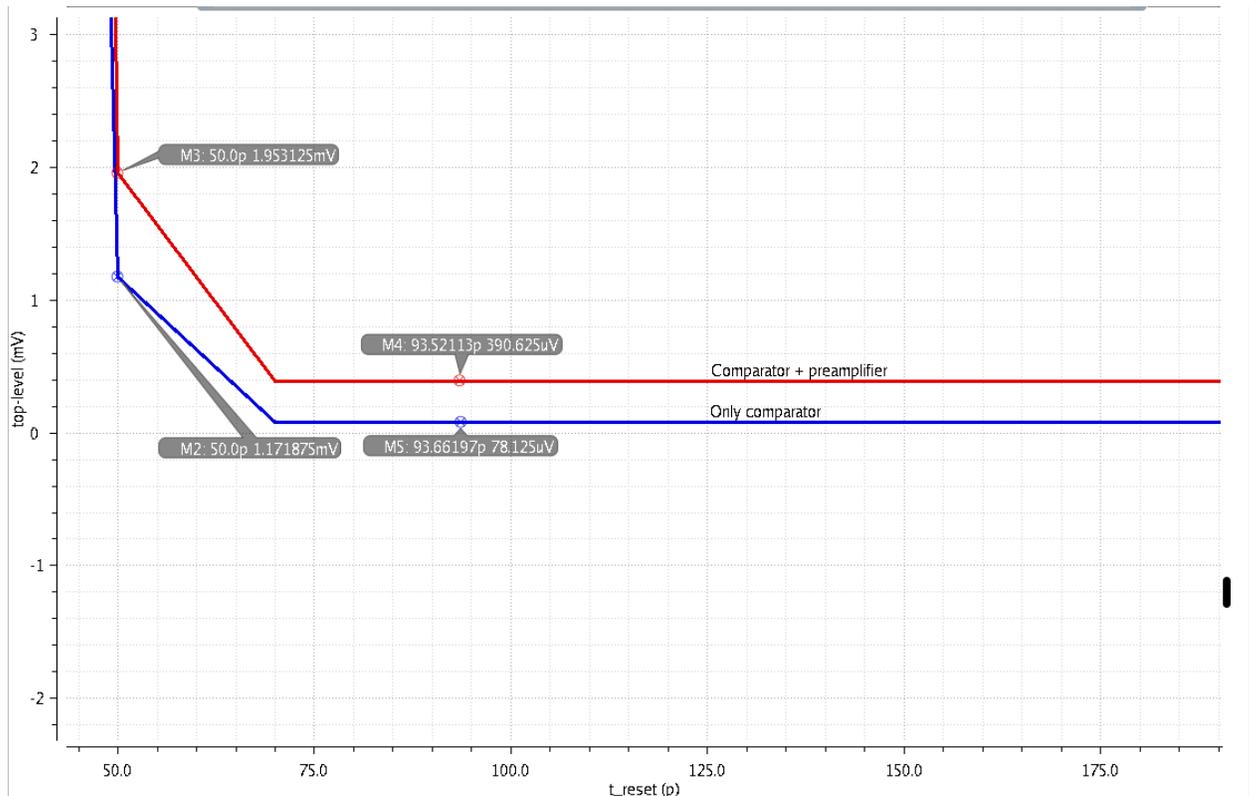


Figure 5.3.4.1 - Residual voltage at the output of the comparator for various reset times

5.3.5. Preamplifier

The main disadvantage of the comparator architecture chosen relates to kickback, an unwanted effect that appears due to the parasitic capacitances of the transistors in the circuit. As soon as one of the outputs has been regenerated to V_{DD} , part of the signal is propagated back to the input and can disturb the voltage levels, leading to a wrong decision. In order to mitigate this effect, as well as to reduce the offset voltage and the noise of the Strong-Arm comparator, a preamplifier was added in front of the comparator. This choice also has the benefit of reducing the offset of the comparator by a factor equal to the preamplifier gain, as well as reducing the input-referred noise of the comparator.

Figure 5.3.5.1 represents the schematic used to design the preamplifier. The circuit features two differential pairs ($M_1 - M_2$ and $M_3 - M_4$), which allows it to be connected directly to the track-and-hold (V_{ip}, V_{in}) and DAC (V_{refp}, V_{refn}) outputs. The two differential signals are subtracted and the result is available at the output of the preamplifier ($V_{op} - V_{on} = \Delta V_{in} - \Delta V_{ref}$). The distribution of the input voltages is important for ensuring a correct operation. Placing V_{ip} and V_{refp} on the same input pair allows the amplifier to operate

The negative resistance needs to be lower than the equivalent conductance $\frac{1}{R_x}$ seen at the output nodes, otherwise the positive feedback slows down the reset operation. This condition stems from Equation (5.28) and can be written in the form:

$$g_{m5,6}R_x < 1 \quad (5.29)$$

As shown before, R helps set the gain of the preamplifier, but it also plays a role in defining the output common mode of the circuit (Equation (5.30)), which needs to be set at 770mV to obtain the desired speed-noise trade-off for the comparator.

$$V_{CM,out} = V_{DD} - R\left(\frac{I_{tail1} + I_{tail2} + I_{tail3}}{2} - I_{d7}\right) \quad (5.30)$$

As the formula shows, transistors $M_7 - M_8$ are also used to set the common mode voltage. In order to generate the appropriate gate voltage for them (V_{biasp}), the replica biasing technique is used.

The biasing of the tail transistors (V_{bias}) is generated using the constant G_mR block depicted in Figure 5.3.5.2, [26], which ensures that the gain of the preamplifiers does not vary with temperature and process-induced uncertainties. The bias voltage fed to the preamplifiers is set at $V_{bias} = 640\text{mV}$.

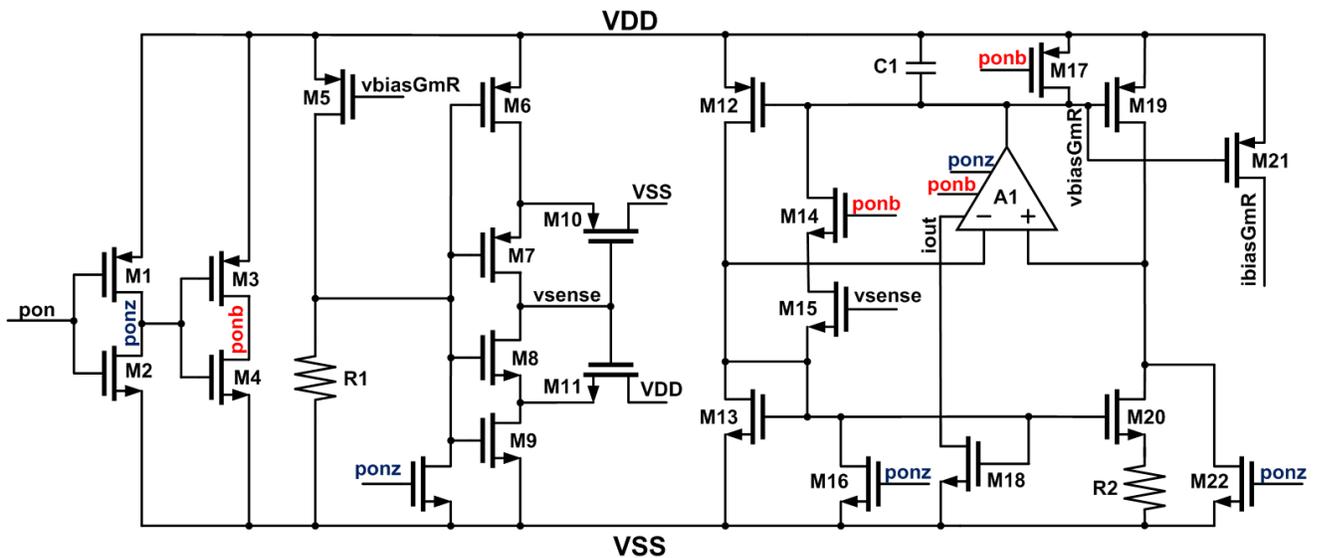


Figure 5.3.5.2 - Constant G_mR bias circuit used for the preamplifiers

The main functionality of the circuit in Figure 5.3.5.2 is given by the loop of the gate-source voltages of transistors M_{13} , M_{20} and the voltage drop across resistor R_2 . If M_{13} is sized

A transient simulation with an input signal amplitude of $\frac{V_{LSB}}{4}$ is used to determine the gain of the preamplifier (Figure 5.3.5.4) - $G_{preamp} = 2.83$. The waveforms also show that the reset switch performs its function correctly, as the differential output voltage is reset to zero when $clk = 0$. The input capacitance of the preamplifier is $C_{pre} = 20.3fF$, slightly higher than the 12fF assumed during the architecture study. The difference comes from the additional set of dummies ($M_{d5} - M_{d8}$) that was needed to prevent a signal-dependent non-linearity from occurring in the DAC references due to feed-through. The tail current of the differential input pair is of $9.8\mu A$ and the bias of the output stage is $4.8\mu A$.

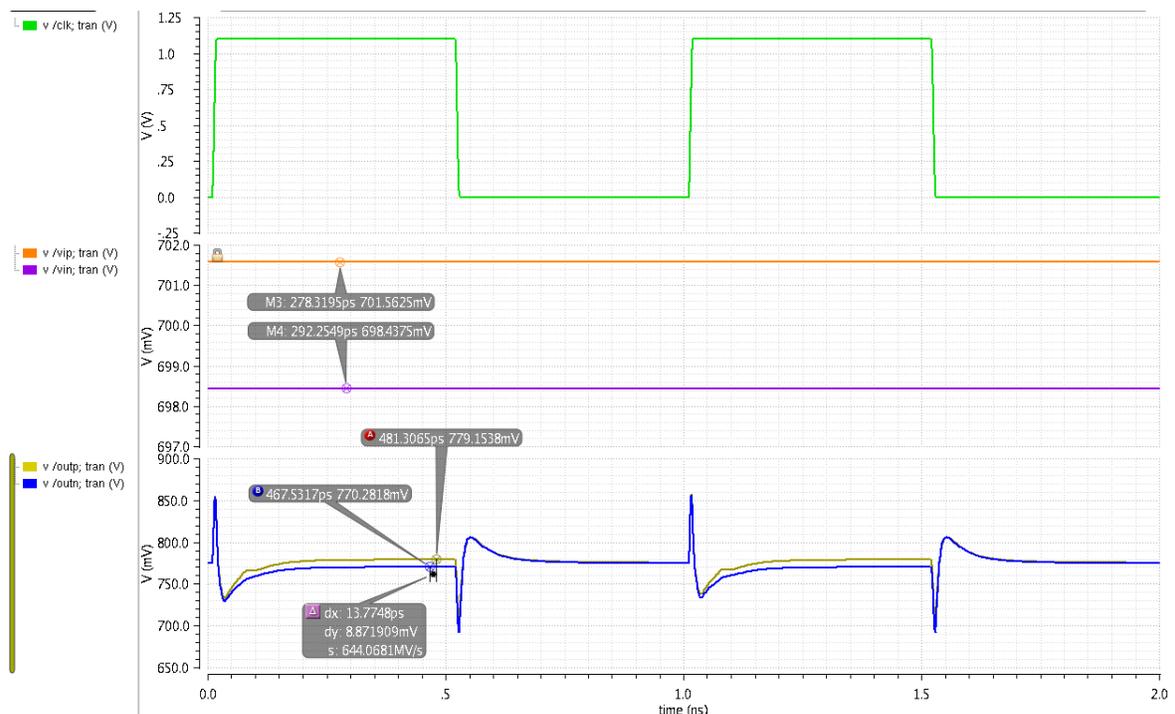


Figure 5.3.5.4 - Preamplifier gain ($\Delta V_{in} = \frac{V_{LSB}}{4}$)

5.3.6. Offset for both the comparator and the preamplifier

According to Pelgrom's formula, [41], the size of a pair of transistors that need to be matched predicts the process spread that can affect their threshold voltages. In the case of the comparator, transistor pairs $M_1 - M_2$, $M_3 - M_4$, $M_5 - M_6$ and $M_7 - M_8$ are critical in terms of matching and failing to manufacture them identically leads to the appearance of an offset voltage at the input of the comparator, which can lead to erroneous comparison results.

In order to estimate the input-referred offset voltage, the variance of the three transistor pairs should be considered:

$$\sigma_{\Delta V_T, total}^2 = \sigma_{in,12}^2 + \sigma_{in,34}^2 + \sigma_{in,56}^2 + \sigma_{in,78}^2 \quad (5.33)$$

The preamplifier helps reduce the input-referred offset:

$$\sigma_{offset,in}^2 = \sigma_{offset,preamp}^2 + \left(\frac{\sigma_{offset,comparator}}{G_{preamp}} \right)^2 \quad (5.34)$$

The 3-bit/cycle resolving scheme is unfortunate with respect to the offset-induced errors, as different comparators will be affected by different process variations, which means that the offset of each will not resemble that of the others. Similar to a flash ADC, this effect translates to a signal-dependent error, but this can be easily removed in the digital domain by using calibration schemes, [42].

A Monte Carlo simulation using 200 runs is used to predict the input-referred offset of the comparator (Figure 5.3.6.1) and preamplifier (Figure 5.3.6.2) as separate elements, as well as that of their combination (Figure 5.3.6.3). It can be concluded that the preamplifier lowers the input-referred offset of the comparator. The remainder, mostly coming from the preamplifier itself, needs to be calibrated, because it is higher than the LSB level.

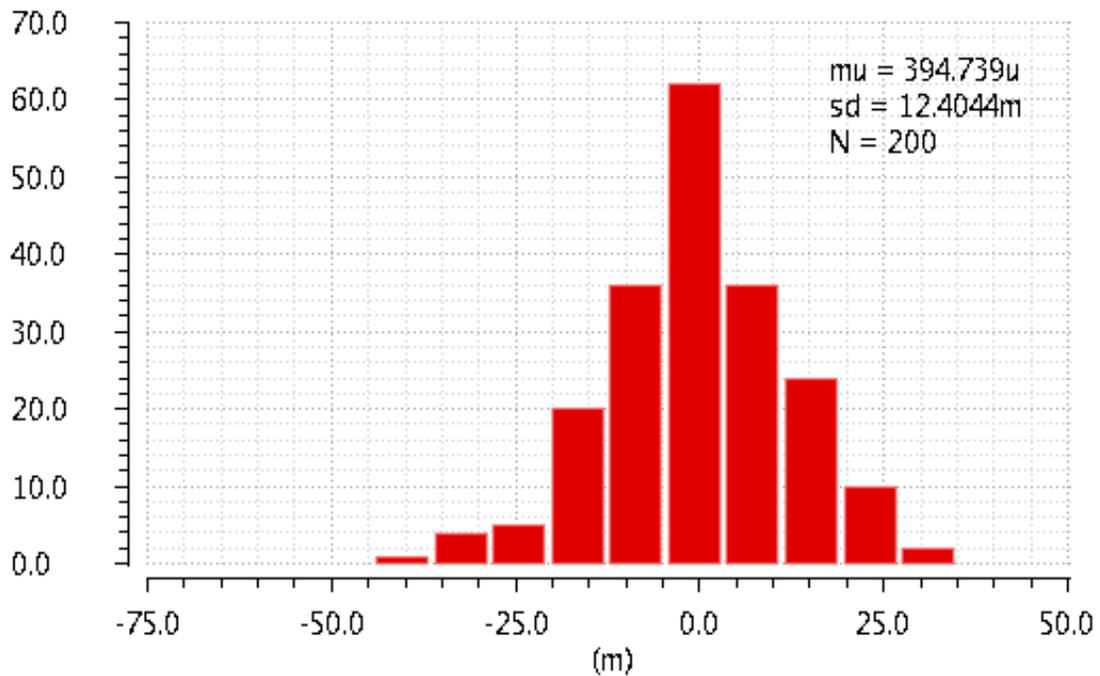


Figure 5.3.6.1 - Offset of the comparator only

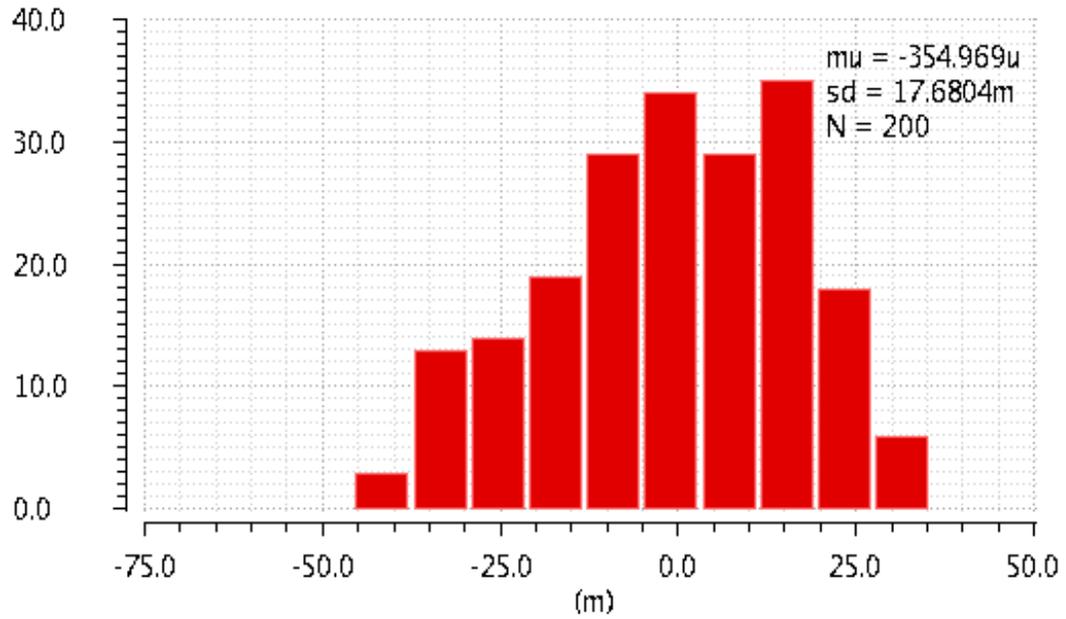


Figure 5.3.6.2 - Offset of the preamplifier only

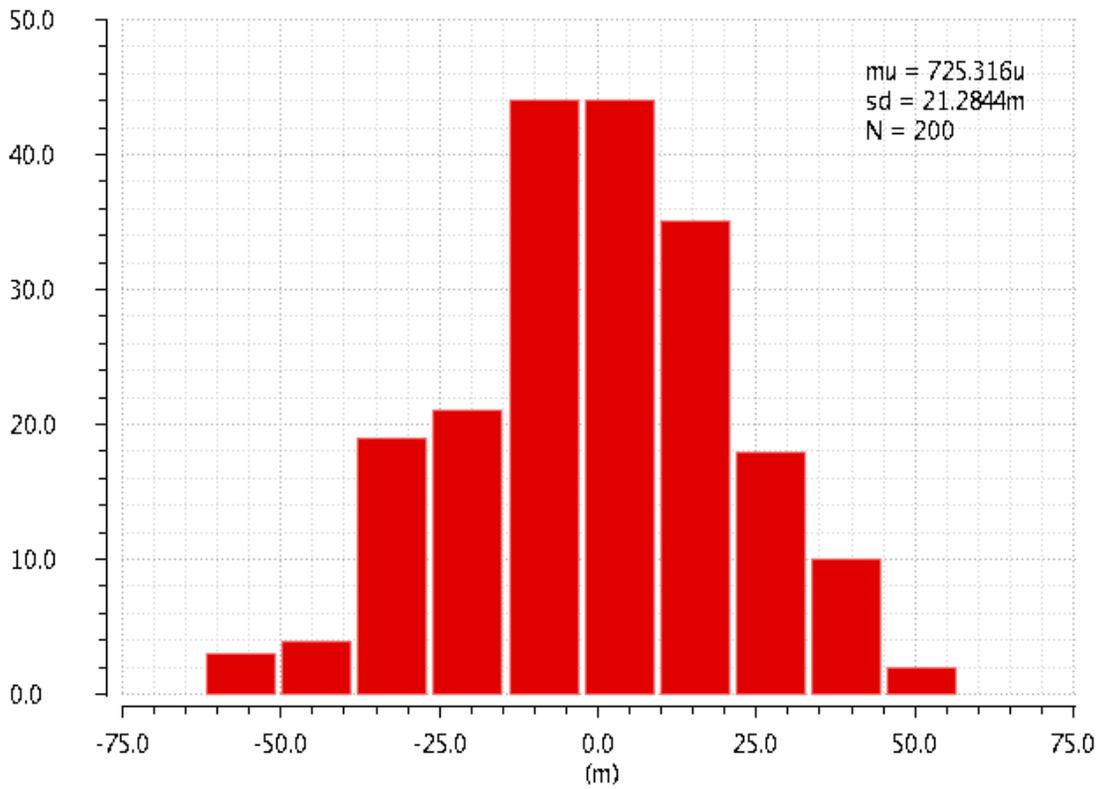


Figure 5.3.6.3 - Offset of the comparator preceded by the preamplifier

5.3.7. Metastability

This section aims to assess whether the output forcing circuit is the best way to treat metastability issues or if simply extending the time allotted to the comparator to make a decision is a more power-efficient solution, albeit one that reduces the maximum sampling frequency of the ADC.

A metastable event occurs when the input of the comparator is very small, which translates to a long decision time, potentially an infinite one if $\Delta V_{in} = 0$, as is predicted by Equations (5.23-5.24). A mechanism to reduce the incidence of metastable events is crucial for maintaining the desired ADC accuracy and several approaches can be employed for this purpose. Chapter 4 already addresses this problem at system level by proposing to force the decision of the comparator to V_{DD} if it has failed to produce a valid output when the time assigned for this operation has been exceeded. The risk of this approach is that the SR-latch that is acted upon may become metastable itself and thus the problem is merely moved to the next block.

A different course of action implies sizing the comparator such that the minimum input signal is smaller, which tends to decrease the probability of encountering a metastable event. If V_{in_min} is the minimum valid input accepted and it is assumed that the signal can take uniformly distributed values in the interval $[0, V_{REF}]$, then Equation (5.35), taken from [43], represents this probability for an N -bit flash ADC.

$$P_{metastability} = \frac{2(2^N - 1)V_{in_min}}{V_{REF}} \quad (5.35)$$

Assuming that all the transistors in the comparator can be scaled simultaneously by the same factor (hereon: $mfactor$), the trade-off between the minimum input signal allowable, the comparator delay and the power dissipation emerges.

Figure 5.3.7.1 shows the dependence of the comparator delay on the input signal if the $mfactor$ is varied. There is little to be gained in terms of speed, but the power penalty is quite high, as Figure 5.3.7.2 demonstrates.

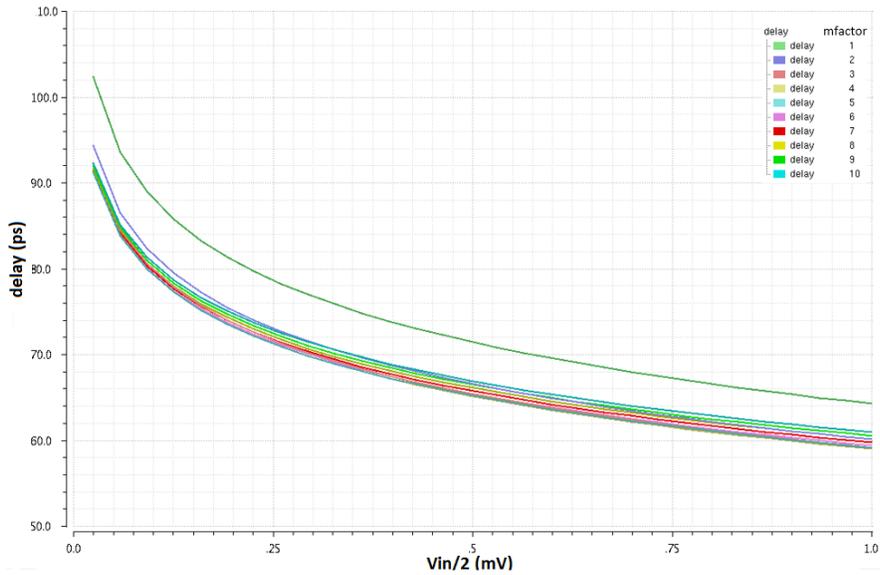


Figure 5.3.7.1 - Comparator delay vs. V_{in} ($V_{in} = 2half_diff$) for different $mfactors$

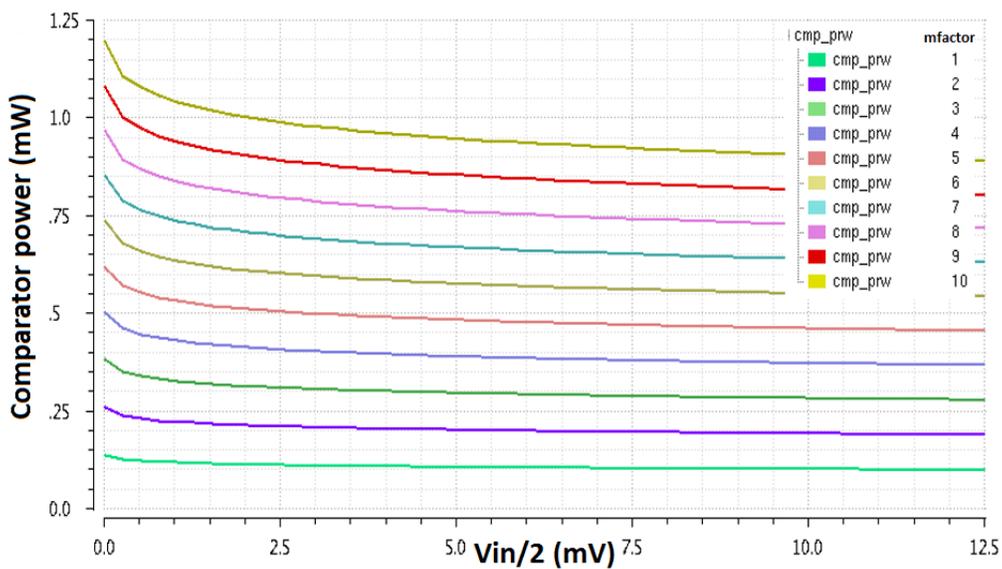


Figure 5.3.7.2 - Comparator power vs. V_{in} ($V_{in} = 2half_diff$) for different $mfactors$

However, spending more power allows the comparator to process a smaller input signal in the 94ps allotted (Figure 5.3.7.3). A shallow optimum exists for $mfactor = 3$, but the power consumption corresponding to it is 3 times higher than that for $mfactor = 1$. This design uses $mfactor = 5$, which means speed was prioritized over power consumption.

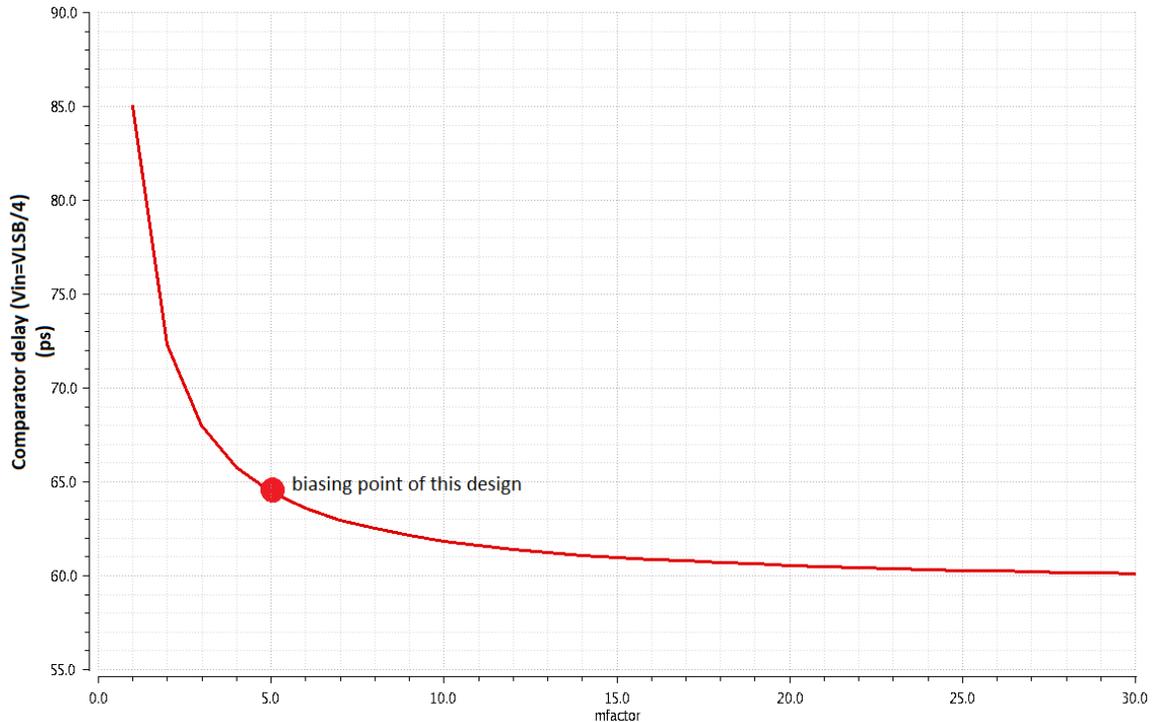


Figure 5.3.7.3 - Minimum input voltage for 94ps (54ps comparator only)

The sizing of the comparator is found to be a good trade-off between power consumption and speed, as it allows the specification (a comparison time of 94ps) to be achieved with a low enough input signal without dissipating an excessive amount of power.

5.3.8. Layout

The comparator is one of the most critical blocks of the ADC, as its delay greatly affects the maximum sampling speed of the ADC. As such, laying it out and extracting the parasitic capacitances associated provides a good indication of the maximum sampling speed of the entire circuit post-layout. The inverters placed at the outputs of the comparator was also laid out in order to measure the delay of the comparator with the load it will see in the taped-out version of the circuit.

Figure 5.3.8.1 shows the layout, while Figure 5.3.8.2 shows the extracted parasitic capacitances associated to the circuit after the post-layout simulation. The output of the comparator sees a parasitic capacitance of about 2.7fF in addition to the capacitance of the inverter it needs to drive.

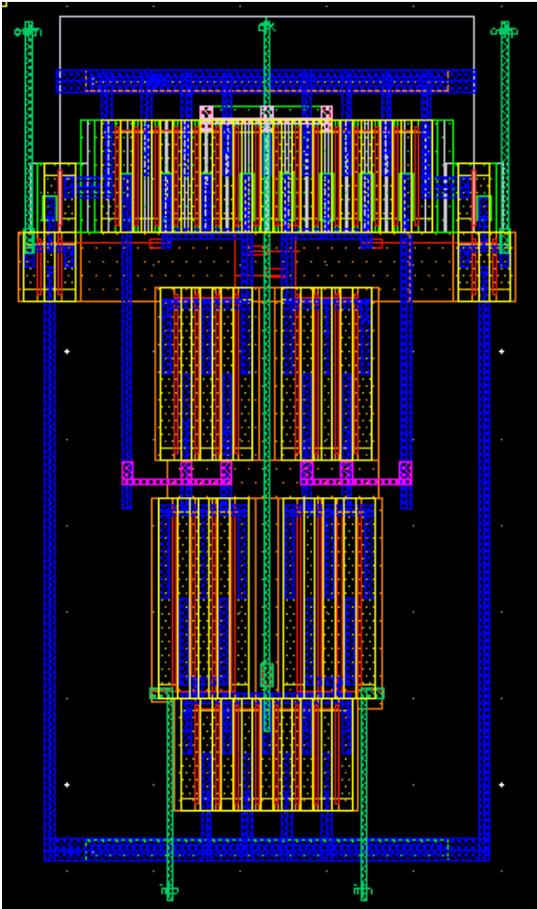


Figure 5.3.8.1 – Layout of the comparator

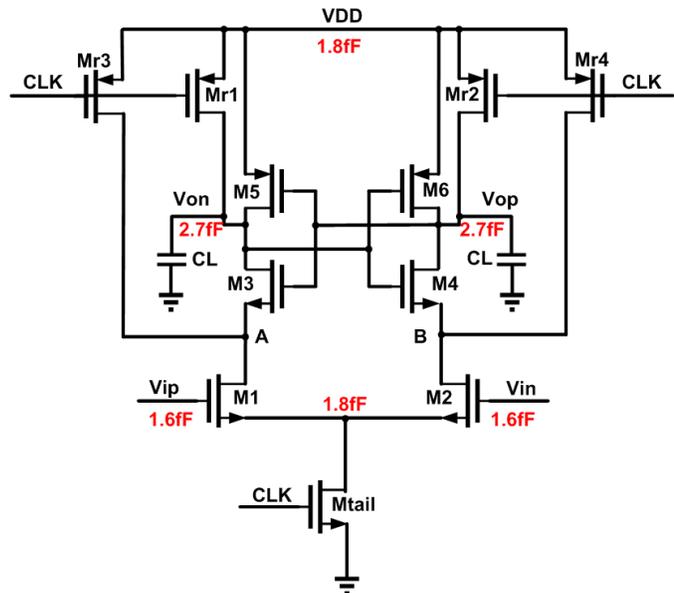


Figure 5.3.8.2 - Post-layout parasitic capacitances associated to the layout

The delay of the comparator vs. the differential input signal was simulated using the schematic and the extracted layout (Figure 5.3.8.3).

| Model | τ_{reg} | t_{int} | G | Delay for an input of $\frac{V_{LSB}}{4}$ | % degradation compared to the schematic |
|-----------|--------------|-----------|-----|---|---|
| Schematic | 10.7ps | 23.5ps | 4.3 | 67.1ps | - |
| Layout | 12.6ps | 22.5ps | 2.7 | 75.3ps | 12.2% |

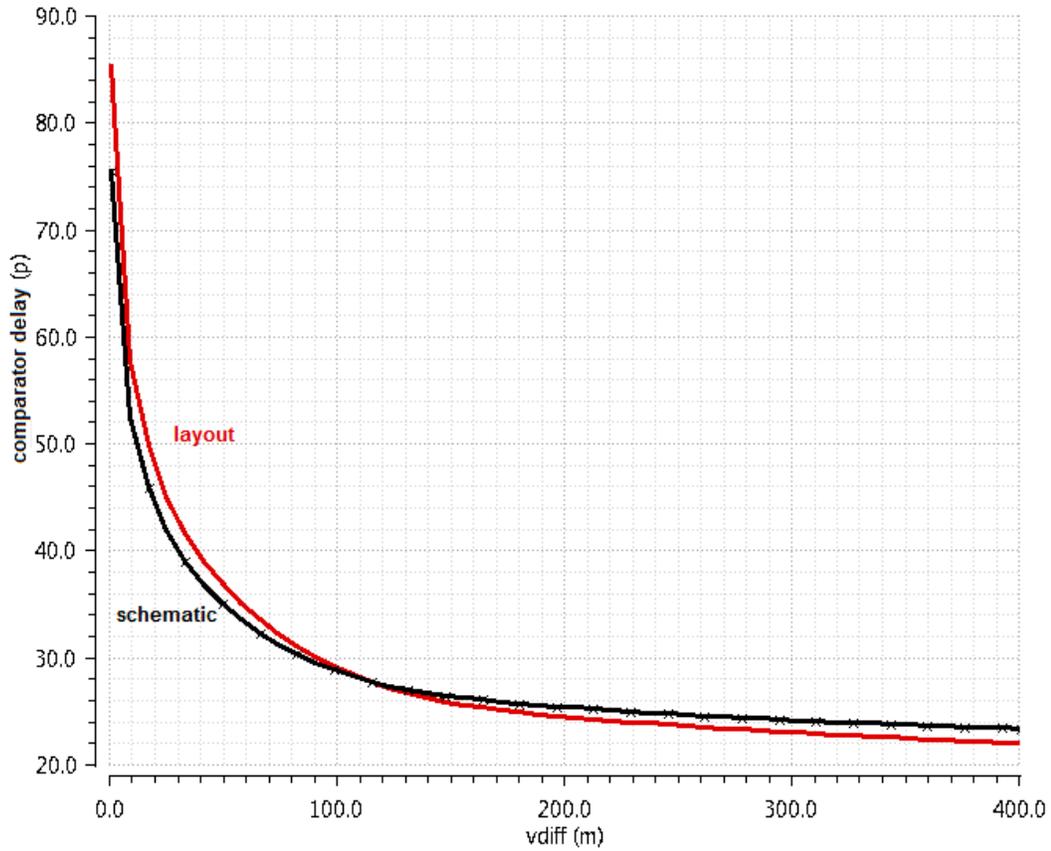


Figure 5.3.8.3 - Delay of the comparator simulated using the schematic and the extracted layout

5.3.9. Conclusion

The final comparison delay is within the 94ps that were assumed for an input of $\frac{V_{LSB}}{4}$ during the architecture study. The noise level is below $\frac{V_{LSB}}{4}$ and the offset is reduced by the preamplifier, while its residue can be removed through calibration. By laying out the comparator, it is shown that a 12.2% degradation in speed is expected for the manufactured block.

5.4. SAR Logic

This section describes the transistor-level implementation of the asynchronous logic of the SAR loop, whose behavioral model has been explained in Chapter 4. The signal names from Figure 2.1 are kept for clarity. Because of the finite time available for the completion of the project, the output-forcing circuit was not implemented. The simplified block diagram, showing only the parts of the digital logic that were implemented and in which several blocks were numbered in order to better structure the explanations of the implementation is shown in Figure 5.4.1.

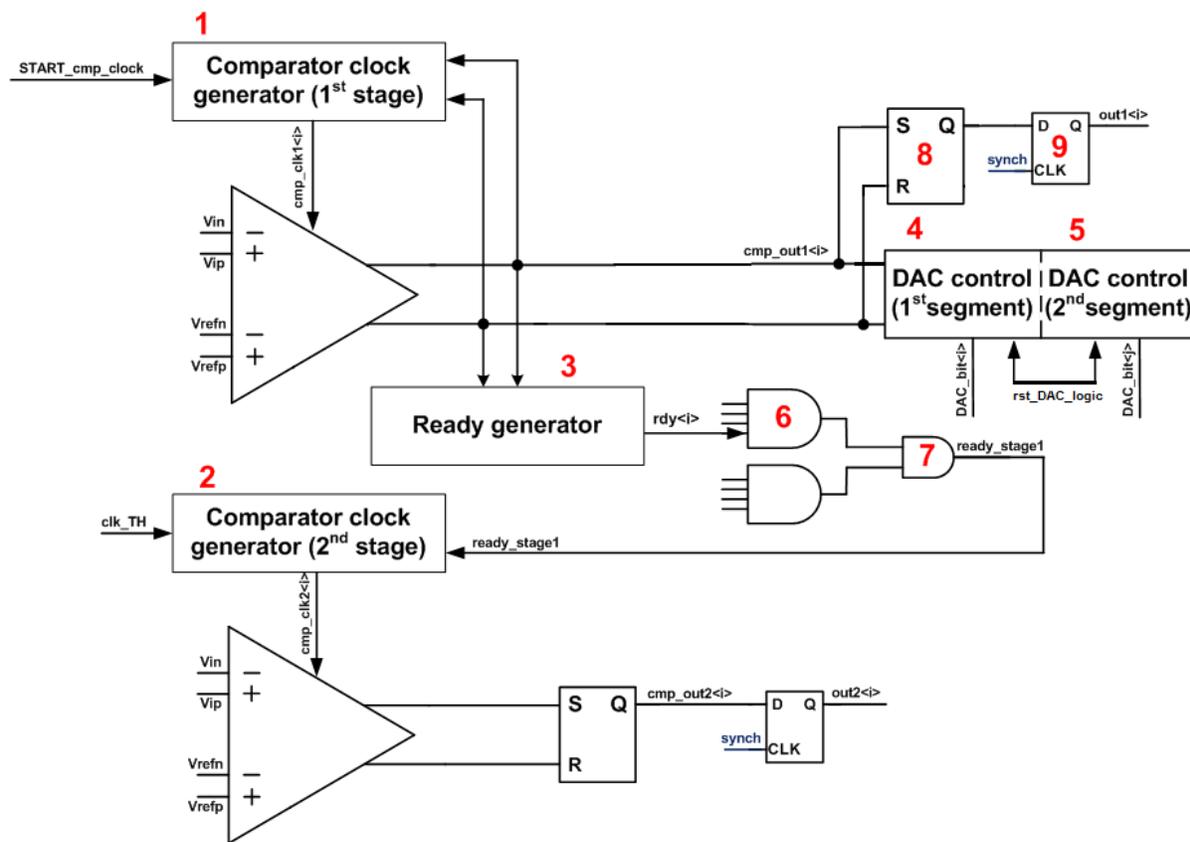


Figure 5.4.1 - Block diagram of the digital logic blocks involved in the SAR loop

5.4.1. Comparator clock generator for the first set of comparators

The clock of each comparator in the first set is generated individually based on when the decision of that particular comparator has been made. The positive edge of the comparator clock marks the beginning of a new comparison, while its negative edge is responsible for resetting the device. The *START_cmp_clk* signal is common to all clock

generators from the first set and ensures that the all comparators are triggered at the same time.

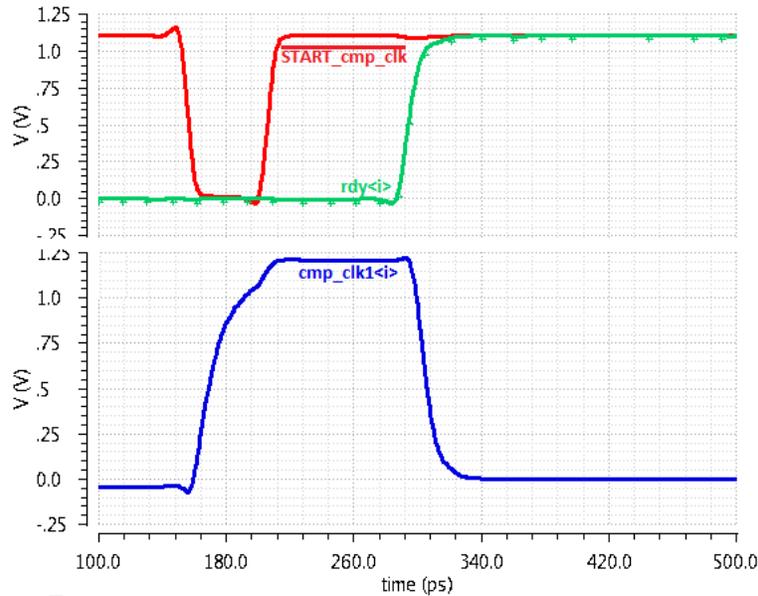


Figure 5.4.1.1 – Waveforms corresponding to generation of the clock of the first set of comparators

This signal is shaped like a pulse and appears slightly earlier than the track-and-hold clock such that the resulting positive edge of the comparator clock coincides with the end of the tracking phase. The signal denoted by $rdy < i >$ represents the ready signal associated to comparator i and is equal to 1 whenever the two outputs of the comparator are of different polarity.

Before each new comparison, the ready signal is low and the output of the comparator clock generator is pulled high as soon as $START_cmp_clk$ turns high. After the comparator has reached a decision, $rdy < i >$ will go high, indicating that a valid output has been processed, which means that the comparator can be safely reset (Figure 5.4.1.1).

5.4.2. Comparator clock generator for the second set of comparators

The generation of the clock signal for the second set of comparators uses a similar circuit. In this case, the $ready_stage1$ signal, generated using NAND gates, signals when the first stage has finished giving a result and starts the clocking phase of the second set of comparators (Figure 5.4.2.1). In case one of the comparators in the second set becomes metastable, its ready signal would never be generated, so a reset scheme based on the ready signal would not ensure that the comparators are reset in every case.

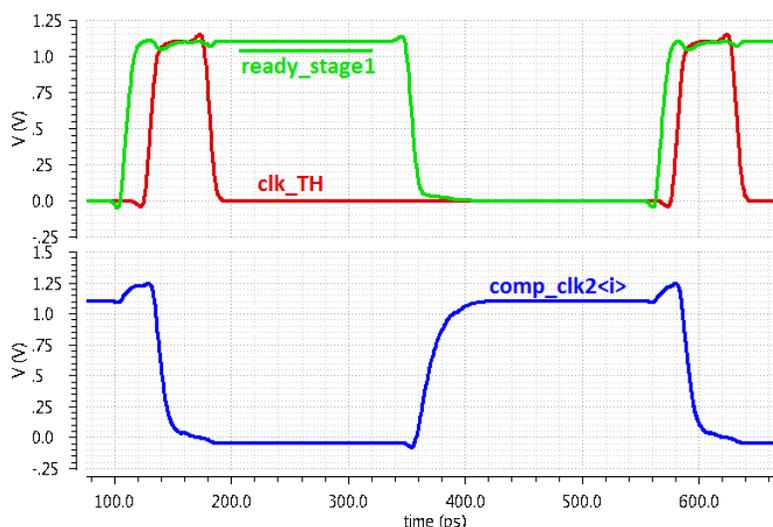


Figure 5.4.2.1 – Waveforms corresponding to the generation of the clock of the second set of comparators

A more robust solution was adopted by using the track and hold clock for resetting the second set of comparators. In this way, the second stage also has sufficient time to dispense of the previous outputs while the first stage is processing a new input signal. For this implementation, each comparator is controlled by its own clock generator, but seeing that the clocks of the second stage are identical, a single block with 7-times larger transistors could also be used to generate these clocks.

5.4.3. Ready generator

To save time in the SAR loop, an asynchronous SAR logic was adopted in order to trigger the second set of comparisons as soon as the first are completed. At the core of this mechanism lies the ready-signal generator, a block that accompanies each comparator in the first stage and is used to signal the end of the decision-making process.

The circuit has three different inputs – the two outputs of the comparator it serves ($outp_cmp < i >$ and $outn_cmp < i >$), as well as the inverted-version of the rst_DAC_logic signal. When the comparators are reset, its outputs are charged to V_{DD} . When the DAC logic is reset, the ready signal is discharged to ground in preparation for a new comparison. When the logic reset signal is low, the circuit waits for the decision of the comparator and its output remains unmodified. As soon as one of the comparator outputs goes low, the ready signal is charged to V_{DD} and kept high until the next conversion, when the DAC reset will be high again. The circuit has to drive the DAC logic for both segments of the DAC, as well as an input of a NAND4-gate that collects the local ready signals and combines them into a global one. Seeing that the DAC logic is sized to drive switches with a

low on-resistance, enough driving capability needs to be ensured for the ready generators as well.

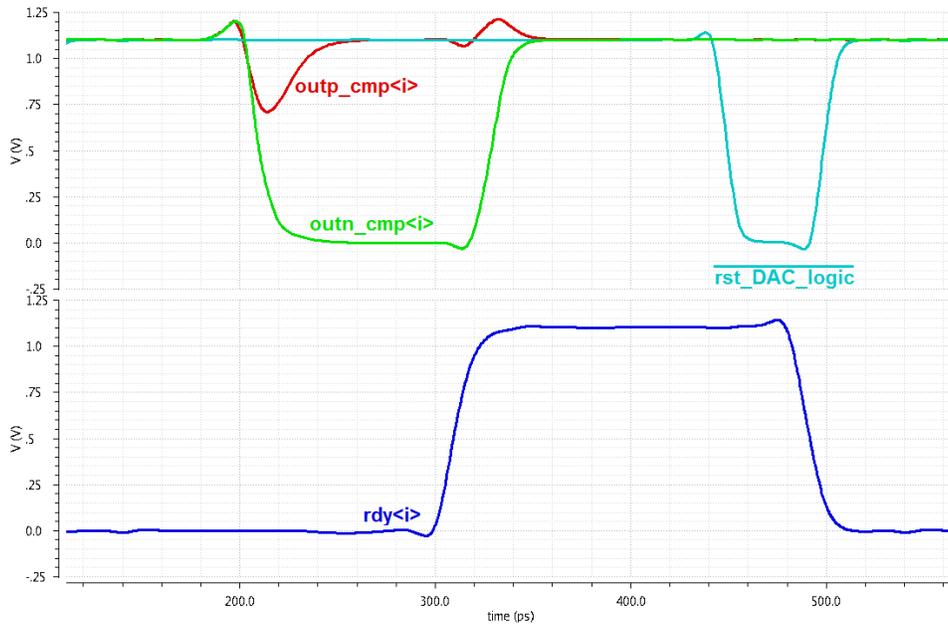


Figure 5.4.3.1 – Waveforms corresponding to the generation of the ready signal of the first set of comparators

5.4.4. DAC control for the first segment

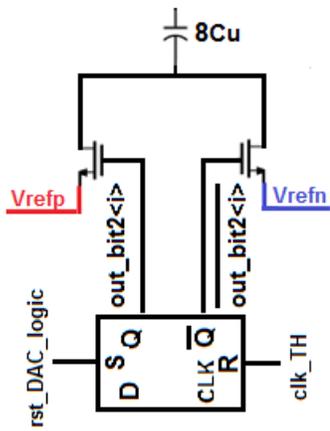


Figure 5.4.4.1 - Example of control signal generation for the first segment of the DAC

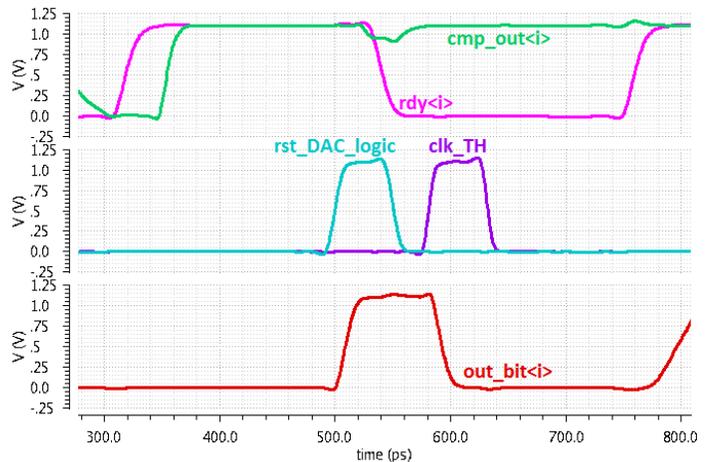


Figure 5.4.4.2 – Example of the control signals of one bit of the first segment of the DAC

The control circuit for the first segment of the DAC is a D flip-flop (Figure 5.4.1.1). When the clock signal goes high ($Clk = 1$) the D-input will be copied at the output of the circuit. The set and reset signals act directly on the output to force it to the value desired. The D-input is connected to the output of the comparator and the circuit is clocked using the ready signal of the same comparator. The set and reset inputs are connected either to the

DAC logic reset, or to the track-and-hold clock, depending on which bit needs to be switched. Details on the matter were provided in Figure 5.2.1.5.

The waveforms shown in Figure 5.4.4.2 show the control signal for the first bit of the first segment of the DAC. The circuit discussed is connected to only one of the bottom plate switches. In order to generate the opposite signal to control the other switch, the same schematic is used, but the reset and set signals are reversed, while the D-input is connected to the comparator output of reversed polarity.

5.4.5. DAC control for the second segment

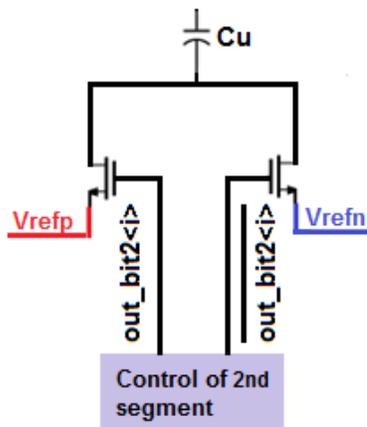


Figure 5.4.5.1 - Example of control signal generation for the second segment of the DAC

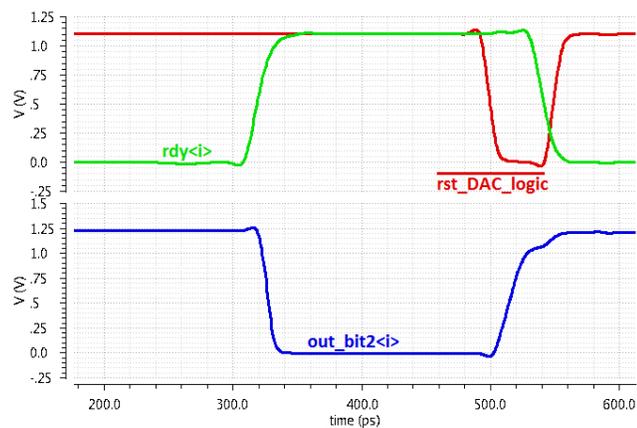


Figure 5.4.5.2 - Example of the control signals of one bit of the second segment of the DAC

Depending on which bit needs to be controlled (as described in Figure 5.2.1.5), the control signals for the bottom plate switches might differ. An example is shown for the control of bit 0 of the positive side of the DAC (Figure 5.4.5.1 and Figure 5.4.5.2). The output signals of the two circuits are inverted versions of each other, in order to ensure that only one switch connected to the bottom plate of the capacitor is switched at one particular moment. The DAC logic reset signal ensures that half of the capacitors in the second segment are connected to V_{refp} , while the other half are connected to V_{refn} during the reset period and while the references for the first set of comparators are generated. As soon as the first three bits are determined, the ready signals turn high and connect the bottom plate of the capacitor such that the correct reference for that particular DAC is generated. In the case of the other switch, the output is pulled low.

5.4.6. NAND4 and NAND2 gates for global ready generation

The speed of each comparator in the first stage depends on the input signal it sees, so different comparators will finish their task at different times. As a consequence, not all ready signals will turn “1” at the same moment. A mechanism to detect when the first stage has

made all decisions is needed such that the second set of comparators is only clocked after the first three bits have been determined. An AND-gate can be used to collect the local ready signals and combine them into a single ready signal for the entire stage (*ready_stage1*).

However, an AND-gate with 7 inputs is not trivial to implement as it requires the stacking of 8 transistors. In order to avoid this, this implementation uses two NAND-gates with 4 inputs and a NAND gate with 2 inputs to perform the same task.

5.4.7. Synchronization blocks

The reset of the comparators, which brings their outputs back to V_{DD} , takes place before all bits have been determined. For this reason, the bits must be stored until the conversion of each sample is done such that all of them are fed to the output at the same time. Failing to do so will result in loss of information.

The outputs of the comparators are followed by 3 inverters whose outputs are low during the reset time, and have an opposite polarity at the end of a comparison. The particularities of these two signals, together with the asynchronous nature of the circuit, invite an implementation of a storage element in the form of an SR latch. When both inputs are low during the reset time of the comparator, the latch keeps its previous value. When one of the outputs goes high, the output of the SR-latch will coincide with the value that is found at the positive output of the comparator.

The output of each SR-latch is connected to a D-flip flop clocked with the same signal (*synch*) in order to align the output bits once every clock signal. The SR-latch and the D-flip flop are not in the critical path of the ADC, so their delay does not play a role in choosing the maximum sampling frequency. However, their delay influences the latency, so a short one is still desirable.

5.4.8. Conclusion

The implementation of the circuits described above reveals the delay and power consumption of each block used for the digital logic (Table 5.4.8.1). The total delay is larger than the one assumed during the architecture study. The main difference appears because only the DAC logic delay was considered during that phase of the project, but in reality, the critical path contains the circuits needed to generate the clocks of the comparators as well as the global ready signal. This is not a problem, however, because the assumption for the speed of the comparator was too pessimistic during the architecture study, which means that the maximum sampling frequency can still be achieved despite the increase in the logic delay.

Furthermore, the power and delay estimations in [26], which constituted the foundation of the architecture study, are found to be much too optimistic. It is expected that the Verilog-A models for the logic are replaced with their transistor-level implementation,

the delay will increase significantly and that the maximum sampling speed will be reduced accordingly. Because the design in [26] has a 3-2-3-bit resolving scheme, two global ready signals need to be generated using NAND-gates. Only the time required to generate these signals exceeds the assumption made for the entire logic, and these are not the only digital circuits in the critical path, which further demonstrates that the assumption was far from reality.

Table 5.4.8.1 - Delay and power consumption of the logic blocks in Figure 5.4.1

| Number in Figure 5.4.1 | Name of the block | Delay | Power |
|------------------------|---|------------------------------------|---------------|
| 1 | Clock generator for the first set of comparators | 40ps (including tapered inverters) | 19.7 μ W |
| 2 | Clock generator for the second set of comparators | 85ps (including tapered inverters) | 20 μ W |
| 3 | Ready signal generator | 50ps (including tapered inverters) | 21.1 μ W |
| 4 | DAC control – first segment | 40ps | 48.6 μ W |
| 5 | DAC control – second segment | 10ps | 12.9 μ W |
| 6 | NAND4 | 20ps | 202.4 μ W |
| 7 | NAND2 | 10ps | 157.9 μ W |
| 8 | SR-latch | 10ps | 8.8 μ W |
| 9 | D-flip flop used for synchronization | 10ps | 3 μ W |

5.5. Reference Generation

As shown previously, the DAC settling time is strongly influenced by the output resistance of the reference buffer, so a reference generation with a very low output resistance is desired. Kull, [44], [8], demonstrates that such a requirement can be met if the circuit shown in Figure 5.5.1 is used. Although the design of this circuit lies outside the scope of this project, the performance assumed for this block is confirmed by the results in [44], [8].

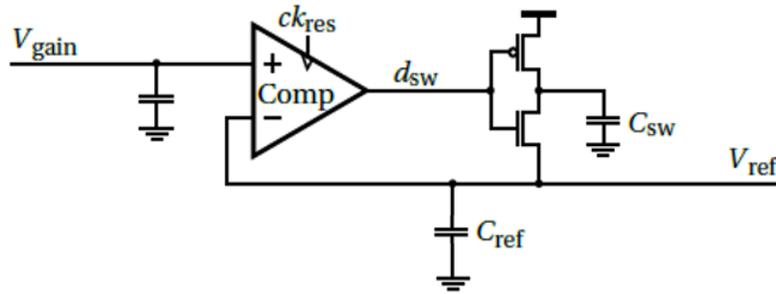


Figure 5.5.1 - Reference buffer with low output resistance [44], [8]

The buffer uses an external voltage reference (V_{gain}) and features a clocked comparator and a reservoir capacitor (C_{sw}). If the comparator is clocked when the DAC is reset, V_{gain} is compared with V_{ref} before the DAC is switched and its value is adjusted using the inverter by adding charge from the reservoir capacitor to C_{ref} . C_{sw} should be large enough such that any charge sharing errors due to the transfer between the reference buffer and the total DAC capacitance remains below half an LSB.

Kull et al. demonstrate that an output resistance of about 2Ω is possible if a large enough C_{sw} is chosen [8]. This value is lower than the 5Ω resistance that was assumed for the DAC settling time calculations, which confirms that the delay expected is pessimistic.

5.6. Calibration

As is the case of flash ADCs, the multi-bit per cycle approach adopted in the design of this SAR can suffer from significant offset problems. The two sets of comparators operating in parallel may be affected by transistor mismatch differently, which means that their offset will no longer be a DC effect that can be corrected simply by adding the corresponding voltage with the opposite polarity. Signal-dependent errors can occur and contribute to the overall non-linearity of the ADC. Several calibration methods of dealing with this problem in the digital domain have been developed and have been shown to reduce the offset by [45], [46], [47], [42]. Implementing an efficient calibration scheme lies outside the scope of this work, but a solution that could be applied to the comparator used in this ADC, taken from [42], is presented for the sake of completeness.

The calibration process goes through two steps:

1. A zero-differential input voltage is applied to the preamplifier connected at the input of the comparator, for example by connecting the V_{ip} and V_{refp} (and V_{in} and V_{refn}) inputs to V_{CM} .
2. The charge pump controls a circuit that injects differential current into the latch present in the comparator, which compensates for the offset before the next comparison takes place.

The resolving scheme chosen for this ADC allows one set of comparators to be calibrated while the other one is resolving the input, which means that both comparators can be calibrated in the background, ensuring an offset-free operation for each sample processed.

Alternatively, an extra comparator could be added, such that one comparator is calibrated in each cycle, while the others resolve the input signal.

5.7. Noise performance

In order to guarantee a 6-bit accuracy for the SAR ADC, the noise contribution of the circuit blocks needs to be lower than the LSB level, [33]. In what follows, noise contribution of every circuit element is surveyed and summarized in Table 5.7.1.

During the conversion of every sample, there are at most two comparators that has to deal with a differential input signal close to LSB level (one for each stage), so the noise contribution of only two comparators and preamplifiers is important for the accuracy. Similarly, the $\frac{kT}{C}$ noise of a single DAC is considered, keeping in mind that each DAC is connected to two different preamplifiers.

The clock jitter is also an important noise contributor, and it is assumed that $\sigma_{jitter} = 400\text{fs}$, which according to data existent at NXP is a realistic specification.

Table 5.7.1 - Noise contribution of the different circuit blocks

| Contributor | Notation and Formula | Assumptions | Observations |
|-----------------------------|--|--|---|
| Differential DAC | $V_{nDAC}^2 = 4 \frac{kT}{C_{DAC} + 2C_{pre}}$ | $C_{DAC} = 108.5\text{fF}$ $C_{pre} = 12\text{fF}$ (values used in the circuit) | The factor “4” accounts for differential operation and for the fact that the DAC is used twice per conversion cycle |
| Differential track-and-hold | $V_{nTH}^2 = 2 \frac{kT}{C_{TH}}$ | $C_{TH} = 14C_{pre}$ (value used in the circuit) | The factor “2” accounts for differential operation |
| Comparator and preamplifier | $V_{cmp+pre}^2 = 2V_{noise_{cmp+pre}}^2$ | $V_{noise_{cmp+pre}} = 1.2\text{mV}$, as measured during the simulation | The factor “2” accounts for the fact that two comparators and preamplifiers are used per conversion cycle (with an LSB-level input) |
| Clock jitter | $V_{njitter}^2 = V_{FS} \cdot 2\pi \cdot f_{in} \cdot \sigma_{jitter}$ | $\sigma_{jitter} = 400\text{fs}$ | - |

The total noise contribution is:

$$V_{noise_{total}} = \sqrt{V_{n_{DAC}}^2 + V_{n_{TH}}^2 + V_{n_{cmp+pre}}^2 + V_{n_{jitter}}^2} \cong 1.442\text{mV} \quad (5.36)$$

If $V_{LSB} = 12.5\text{mV}$, then $\frac{V_{LSB}}{\sqrt{12}} = 1.693\text{mV}$ and $V_{noise_{total}} \ll \frac{V_{LSB}}{\sqrt{12}}$, which is true for the assumptions considered.

5.8. Considerations on the total loop delay

The post-layout simulations show the degradation that is expected for the comparator delay after the circuit is manufactured. The same percentage of degradation will be seen for the logic delay, so increasing the delay of the schematic with the appropriate factor provides a good estimation of the total loop delay expected for the whole circuit.

The loop delay, using the same notations adopted throughout this chapter, is given by:

$$t_{loop} = t_{track} + t_{rst_{DAC}} + t_{comp1} + t_{logic} + t_{DAC} + t_{comp2} \quad (5.37)$$

If a metastable event happens in the first set of comparisons ($V_{in_{comp1}} < \frac{V_{LSB}}{4}$), t_{comp1} will be higher than was assumed during the architecture study. However, due to the way the references are generated, the longest comparison in the next set will not cause a metastable event because the differential input voltage will be at least $\frac{3V_{LSB}}{4}$ (Figure 4.4.3). The interdependence of the two comparator delays, the input voltage corresponding to the longest comparison in the second set is given by:

$$V_{in2} = V_{LSB} - V_{in1} \quad (5.38)$$

The logic delay is expected to degrade by the same factor as the comparator delay after the entire circuit is laid out, while the DAC reset time, the DAC settling time and the tracking time will remain unchanged. The DAC reset and settling times depend only on the DAC capacitance, which will remain unchanged after the layout thus maintaining the same values for these delays. Furthermore, the total DAC capacitance was sized such that it is much larger than the input capacitance of the preamplifiers, so a 12% increase in the latter will not alter the settling time significantly. The linearity of the track-and-hold is much higher than that of the ADC (more details about the SNDR of the ADC are given in Chapter 6), which means that it was designed with enough margin to maintain the 50ps tracking time even

after the layout. This means that the maximum sampling frequency of the signal will degrade by a smaller factor than the comparator delay (Table 5.8.1).

Table 5.8.1 - Contributors to the loop delay

| | t_{track} | $t_{rst_{DAC}}$ | t_{comp1} | t_{logic} | t_{DAC} | t_{comp2} | t_{loop} |
|------------------|-------------|-----------------|-----------------|-------------|-----------|-----------------|---------------------------------------|
| Schematic | 50ps | 50ps | t_{comp1} | 135ps | 25ps | t_{comp2} | $260ps + t_{comp1} + t_{comp2}$ |
| Layout | 50ps | 50ps | $1.12t_{comp1}$ | 151ps | 25ps | $1.12t_{comp2}$ | $276ps + 1.12(t_{comp1} + t_{comp2})$ |

Based on the considerations mentioned previously, the loop delay was calculated based on the minimum input signal that represents the limit between a metastable event and a one that does not cause this issue. The results corresponding to the schematic and to the layout are plotted in Figure 5.8.1.

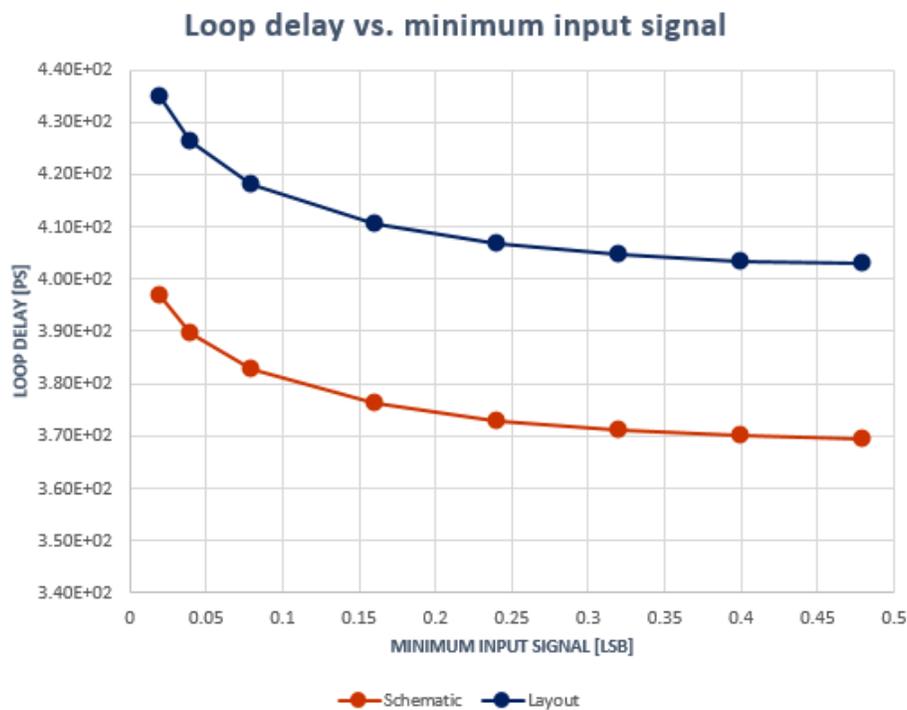


Figure 5.8.1 - Minimum loop delay vs. input signal

When the minimum input signal is $0.02V_{LSB}$, the parasitic elements associated to the custom layout only cause an increase of 40ps compared to the loop delay predicted by the schematic simulations.

Based on the minimum signal allowable, the relation between the sampling speed and the probability that a metastable event will occur during the conversion can be determined (Figure 5.8.2). The plots reveal that a trade-off between accuracy and speed always exists.

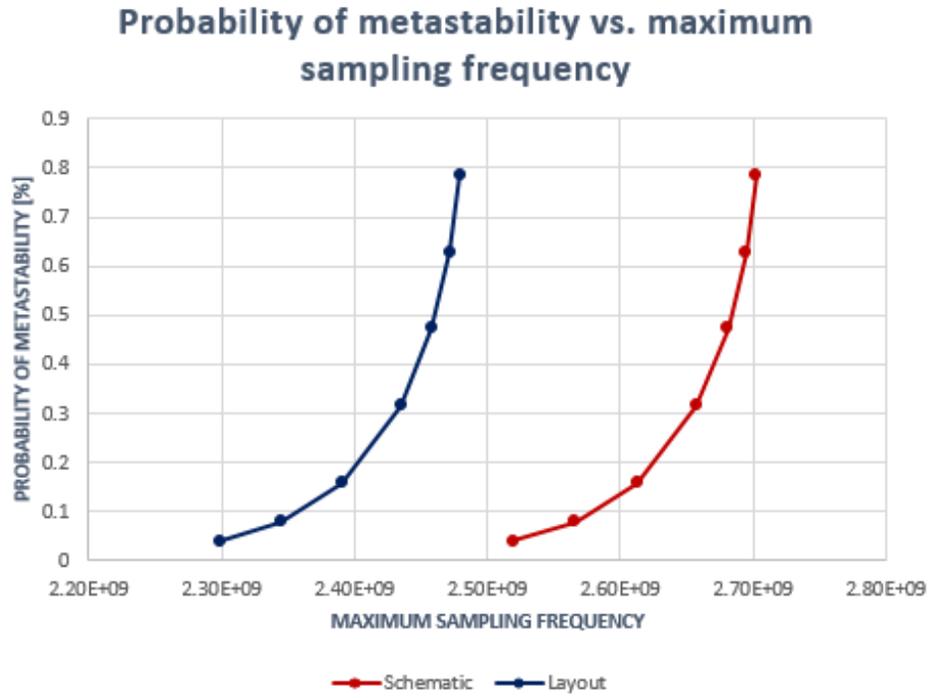


Figure 5.8.2 - Probability of metastability vs. maximum sampling frequency

6. Top Level Simulations

6.1. Summary of Performance

The 6-bit ADC designed features a 3-bit/cycle resolving scheme and uses a reference voltage of 515mV to quantize an 800mV_{pp} full-scale signal. The converter is expected to run at a nominal speed of 2.3GS/s after it is manufactured and consumes a total of 41 mW when used at this operating frequency. The effective number of bits (ENOB) for a Nyquist input at this sampling speed is $ENOB = 5.67$ bits and accompanies a Walden Figure-of-Merit of 350fJ/conv-step.

The power consumption of each block in the ADC is summarized in Table 6.1.1.

Table 6.1.1 – The simulated power consumption of different ADC components

| Block | Power consumption @2.3GS/s |
|---|----------------------------|
| Preamplifiers | 11.6mW |
| Comparators | 1.5mW |
| Digital logic for the control of the DAC | 6mW |
| Other digital blocks needed for asynchronous operation | 5.2mW |
| Clock buffers | 14.6mW |
| Preamplifier biasing | 1.3mW |
| Output buffers (SR-latches and D-flip flops used for synchronization) | 896μW |
| Total without clock buffers | 26.5mW |
| Total | 41mW |

*The power consumed by the reference buffers is not included in this summary because the design of the reference buffers was not part of the scope of this thesis. The same is true for the reference buffer of the boosting voltage needed in the track-and-hold.

Figure 6.1.1 shows the distribution of power consumption among the various circuit elements. The main contributors are the clock buffers and the preamplifiers. The former refer to the control signals of the ADC (*clk_TH*, *resetDAC*, *rst_DAC_logic*, *START_cmp_clk* and *synch*), which were simulated using tapered inverters based on the loads that they needed to drive. The fact that some of these signals were connected to 98 flip flops (14 for each of the 7 DAC control blocks in the case of *rst_DAC_logic*) means that they need to drive a large capacitance, which explains the high power consumption of the buffers. The estimation used during the architecture study was based on the assumption made in [26], and it was difficult to predict this power then, as it is highly dependent on the implementation chosen for the DAC control logic.

The preamplifiers need a static current to operate, and in this sense, they are not the most energy-efficient solution, accounting for 28% of the ADC's power consumption. Replacing them with dynamic preamplifiers or adopting an interpolation scheme to reduce their number is a good option for reducing the current needed to power them. Despite the fact that a lower supply voltage is used compared to the 8-bit design in [26], a larger unit capacitance was needed for the DAC in order to ensure the total DAC capacitance was 10 times larger than the input capacitance of the preamplifiers, which explains the extra power consumption.

The digital logic that controls the DAC and governs the asynchronous operation requires about 11mW, which is much higher than the 2mW estimation provided by [26]. The difference stems mainly from the fact that Ramkaj assumed that a couple of inverters would be sufficient to perform these operations, but more careful consideration quickly revealed that asynchronous operation is far from being this trivial.

The layout of the comparator revealed a 12.2% increase in parasitic capacitance compared to the schematics, so it is expected that the dynamic power of the comparator, clock generators and digital logic will increase by the same factor. The other blocks are expected to maintain their power consumption, as they mainly consume static power. The overall increase in power dissipation of the ADC is expected to be less than 12.2%.

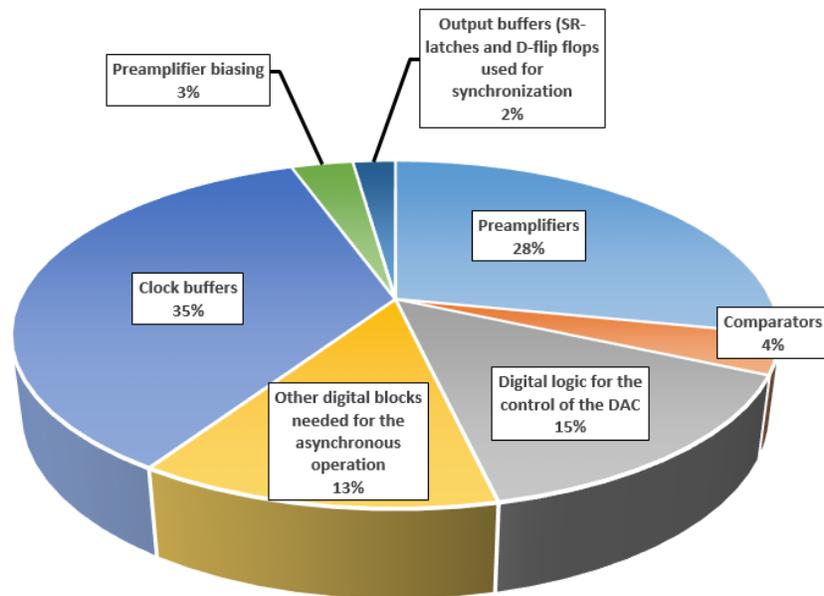


Figure 6.1.1 - Distribution of simulated power consumption among circuit blocks

Figure 6.1.2 shows the spectrum of the reconstructed output signal for a full-scale input at Nyquist frequency and a sampling rate of 2.5GS/s, obtained when simulating using the schematic. The distortion and noise components account for a loss of 0.33 bits (an ENOB of 5.67 bits) of the full resolution of the ADC, which is an excellent result, bearing in mind that the probability of metastability was not reduced using the output-forcing circuit. If a

signal is close to one of the reference voltages in the first set of comparators, then it is at least 1 LSB away from the references in the next, which means that a metastable error can only happen once per conversion. The asynchronous generation scheme proves beneficial by distributing the available conversion time more judiciously than in the synchronous SAR: a long comparison can “borrow” some time from the next one, which will surely take less time because of the larger differential input signal.

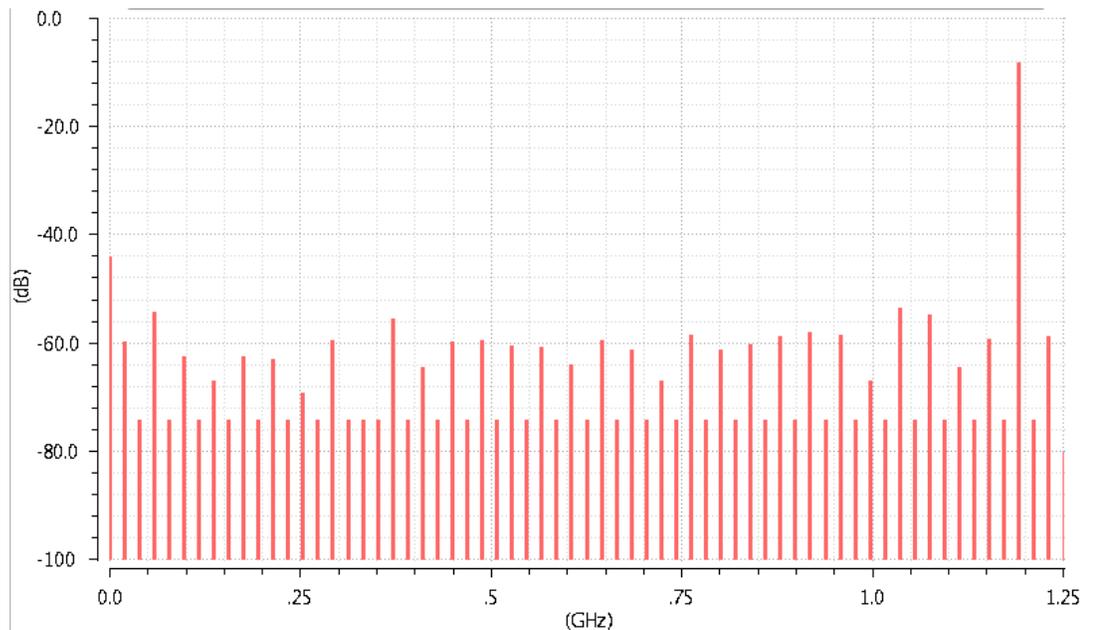


Figure 6.1.2 - FFT of the reconstructed output signal simulated at schematic level (full-scale input, f_{in} @ Nyquist, $f_s=2.5GS/s$)

Subtracting the reconstructed output signal from the track-and-hold output signal shows the error of the ADC (Figure 6.2.5). The points where a larger error appears correspond to the situations in which the input signal was close to one of the references.

Another error not shown on any of the plots, but which will affect the behavior of the circuit once it is taped-out is the offset of the comparator and preamplifier. While the input-referred offset of the comparator is less significant because it is divided by the amplification of the preamplifier, the total input-referred offset can be bigger than the LSB voltage level in some cases. It was shown in Chapter 5 that effective solutions to correct this error in the digital domain exist in the form of calibration.

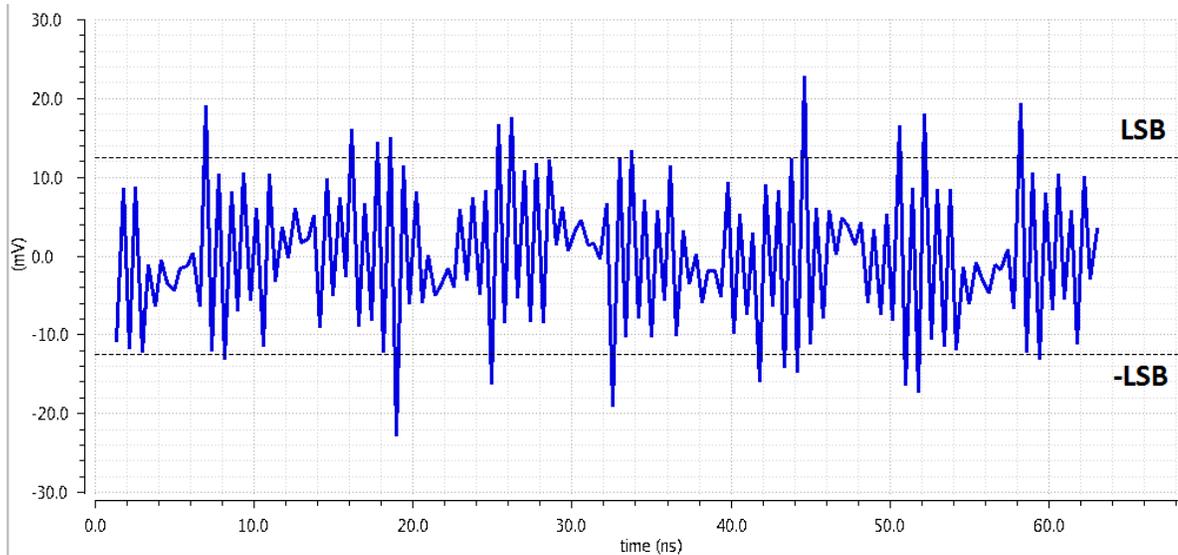


Figure 6.1.3 - Error signal for a full-scale input with f_{in} @Nyquist, $f_s=2.5GS/s$, based on schematic simulations

6.2. Linearity

The linearity of the ADC was assessed using the 6 traditional metrics explained below.

1. The total-harmonic-distortion (THD, Equation (6.1)) is usually expressed in dB and represents the ratio between the power of the input signal and the distortion components found in the spectrum of the reconstructed output of the ADC. V_{FS} is used to denote the full scale of the ADC.

$$THD = 10 \log_{10} \frac{\left(\frac{V_{FS}}{2\sqrt{2}}\right)^2}{V_{rms_{harmonics}}^2} \quad (6.1)$$

2. The spurious-free dynamic (SFDR) range is the ratio between the amplitude of the fundamental and the highest undesired component, no matter its origin. This metric gives a better idea about the smallest input signal that can be tolerated by the ADC without significant error.
3. The signal-to-noise ratio (SNR, Equation (6.2)), expressed in dB, shows how much larger the signal RMS value is compared to the main noise sources of the circuit (in this case, thermal and quantization noise).

$$SNR = 10 \log_{10} \frac{P_{signal}}{P_{noise}} \quad (6.2)$$

4. The signal-to-noise-and-distortion ratio (SNDR, Equation (6.3)), also expressed in dB, provides information on the ratio between the signal power and the total power of the distortion components, together with the noise.

$$SNDR = 10 \log_{10} \frac{P_{signal}}{P_{noise} + P_{distortion}} \quad (6.3)$$

5. The effective number of bits (ENOB) is related to the SNDR through Equation (6.4) and represents the resolution of an ideal ADC that would render the same performance as the circuit considered.

$$ENOB = \frac{SNDR - 1.76}{6.02} \quad (6.4)$$

6. The Walden Figure-of-Merit (FoM), combines the performance of an ADC into a single number (Equation (6.5)).

$$FoM = \frac{Power}{2^{ENOB} \cdot \min(2BW, f_s)} \quad (6.5)$$

The first 5 performance metrics are simulated for different sampling speeds and the results are plotted in Figure 6.2.1 for the schematic simulations. The expected performance after layout is summarized in Figure 6.2.2, based on the degradation of the sampling speed that was explained in Chapter 5. It can be easily seen that for sampling speeds equal or lower than 2.3GS/s, the ENOB stays around 5.6, while the SNDR is higher than 35dB. The performance tends to worsen in all aspects if the nominal sampling rate is exceeded. This is explained by the fact that shortening the sampling period leaves insufficient time for the second set of comparators to make a decision, so the effective resolution drops to about 3 bits.

The effective resolution was also plotted versus the sampling frequency (Figure 6.2.3 for the schematic simulations and Figure 6.2.4 for the expected post-layout performance) and the input frequency (Figure 6.2.5) and reveals that up to Nyquist the ADC achieves a resolution better than 5.5 bits. As the input signal reaches the maximum bandwidth allowed for the application (10GHz), the resolution tends to decrease slightly. This happens because the track-and-hold linearity is worse at that frequency, owing to the variation of the on-resistance of the switch and the non-linearity of the input capacitance of the preamplifiers.

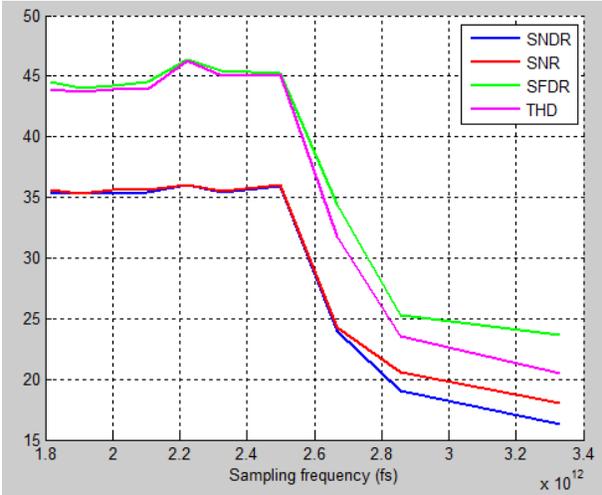


Figure 6.2.1 – SNDR, SNR, SFDR, THD vs. sampling frequency from schematic simulations

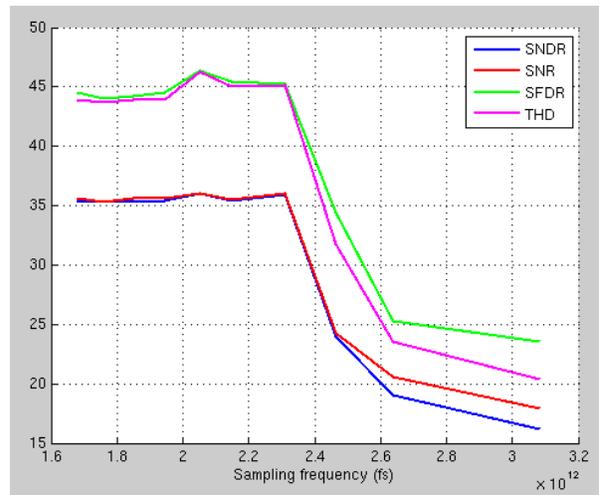


Figure 6.2.2 - SNDR, SNR, SFDR, THD vs. sampling frequency expected after layout

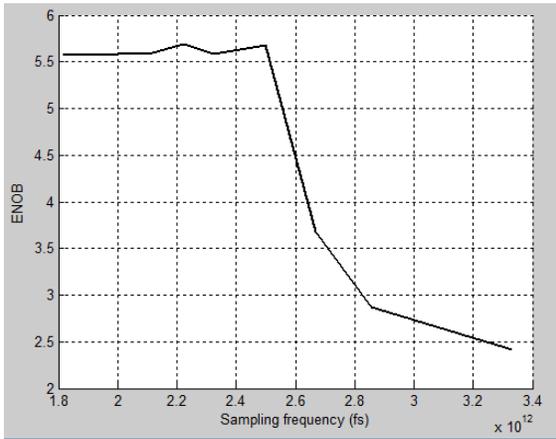


Figure 6.2.3 - ENOB vs. sampling frequency from schematic simulations

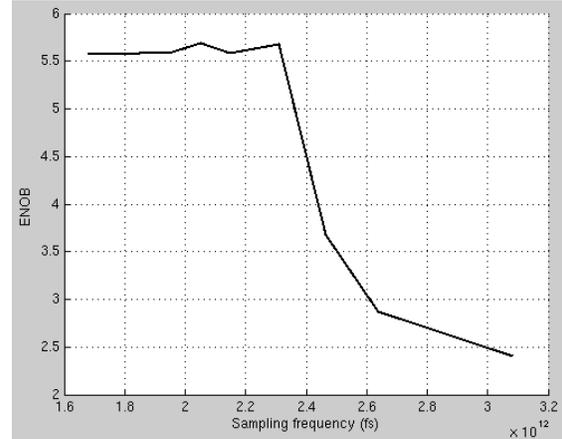


Figure 6.2.4 - ENOB vs. sampling frequency expected after layout

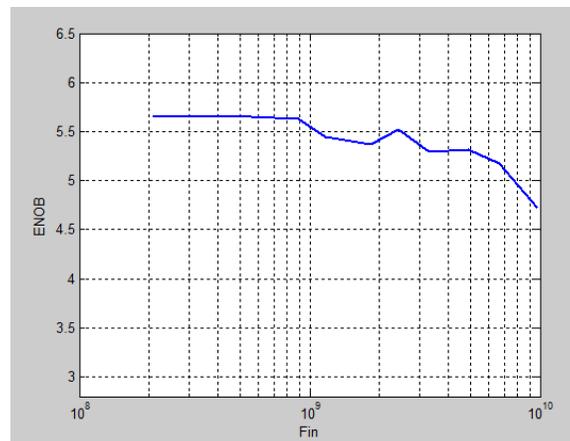


Figure 6.2.5 - ENOB vs. input frequency from schematic simulations

Table 6.2.1 summarizes the performance and main specifications of the ADC for a Nyquist input signal. The maximum sampling frequency corresponds to the post-layout estimations, while the other performance metrics correspond to the schematic simulations.

Table 6.2.1 – Simulated Performance Summary of the ADC

| | |
|------------------------------|----------------------------------|
| Technology node | 40nm CMOS |
| Resolution | 6 bits |
| Supply Voltage | 1.1V |
| Sampling rate | 2.3GS/s (post-layout estimation) |
| SNDR @ Nyquist | 35.9 dB |
| SFDR @ Nyquist | 45.2 dB |
| THD @ Nyquist | -45.1 dB |
| SNR @ Nyquist | 36.02 dB |
| ENOB @ Nyquist | 5.67 bits |
| Total Power Consumption | 41mW |
| Walden Figure-of-Merit (FoM) | 350fJ/conv-step |

6.3. Comparison with similar state-of-the art ADCs

In what follows, the ADC will be compared with a few other state-of-the art designs in order to analyze the efficiency of the techniques adopted. Table 6.3.1 offers a bird’s-eye view of the performance of the designs surveyed, while Figure 6.3.1 shows the positioning of this design on the Murmann plot with respect to the papers surveyed in the introduction of this thesis.

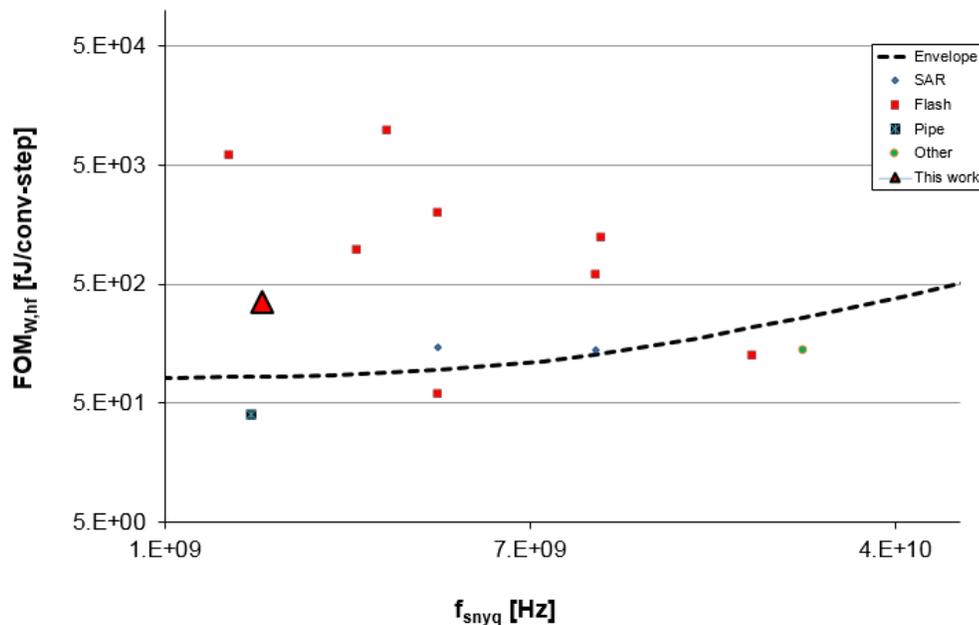


Figure 6.3.1 - Positioning with respect to the Murmann Plot

Table 6.3.1 - Comparison with state-of-the art ADCs

| | [14] | [8] | [12] | This work |
|---------------------------------|-------------------|---|---------------------------------|------------------------------|
| Technology | 65nm GP CMOS | 32nm SOI CMOS | 65 nm CMOS | 40nm LP CMOS |
| Architecture | Binary search, TI | Asynchronous SAR with alternate comparators | Synchronous 3-bit/cycle SAR, TI | Asynchronous 3-bit/cycle SAR |
| Power supply (V) | 1 | 1 | 1.2 | 1.1 |
| Sampling rate (GS/S) | 25 | 1.2 | 5 | 2.3 |
| Sampling rate of sub-ADC (GS/s) | 3.125 | 1.2 | 1.25 | 2.3 |
| Resolution (bits) | 6 | 8 | 6 | 6 |
| Power consumption (mW) | 88 | 3.1 | 10.6 | 41 |
| SNDR @ Nyquist (dB) | 29.57 | 39.3 | 30.13 | 35.9 |
| ENOB @ Nyquist (bits) | 4.62 | 6.235 | 4.71 | 5.67 |
| FoM @ Nyquist (fj/conv-step) | 143 | 34 | 67.37 | 350 |

The design in [12] offers the fairest comparison as it features a similar architecture: both the work of Chan et al. and this ADC adopt a 3-bit/cycle resolving scheme in order to increase the maximum sampling frequency. However, this design uses 14 comparators and preamplifiers to determine the output bits, whereas the one in [12] opts for a more power-efficient solution and interpolates the outputs of the comparators and preamplifiers. A drawback of this approach is that interpolation tends to be a slow operation, [32], and that it risks running into problems with the kickback coming from the comparators. The latter effect appears because the kickback from more than one comparator will be seen at the input of every preamplifier, which translates to a larger error. The differences in linearity further support this observation.

In order to maximize the speed of the single-channel ADC, the design in this thesis uses asynchronous processing in order to prevent a sizing of the SAR loop based on the slowest comparison. The work in [12] uses a synchronous approach, so less digital blocks are needed to control the loop, because intermediate clock signals do not need to be generated, as is the case of this ADC, which needs to wait until all local ready signals have been generated in order to start clocking the second set of comparators.

The design of Kull et al., [8], presents an 8-bit high-speed single-channel ADC running at 1.2GS/s and is a good example of a design aggressively sized to maximize speed. The authors opted for an asynchronous SAR operation, together with the use of alternate

comparators to save time in the critical path by resetting one comparator while the other is making a decision. In this respect, the design is similar to the one proposed in this thesis, but resolves only 1 bit every cycle. In order to achieve 8-bit accuracy, redundancy was used in the DAC. In this design, the settling time of the DAC was not an issue, so redundancy was not needed.

Another important aspect is that the work in [8] is implemented in a 32nm CMOS SOI technology, which means that the parasitic capacitances of the transistors are much lower than in bulk CMOS, generally leading to faster and more power-efficient designs. While the architectural advantages shown in [8] remain, implementing the same design in a less advanced technology node would surely render a slower design. Conversely, implementing the design described in this research in 32nm SOI is expected to increase its speed significantly due to the reduction in parasitic capacitances which currently pose a limit for the maximum speed. The efficiency of the two designs cannot be easily compared because of this fundamental difference in technology nodes used.

Last, but not least, Cai et al., [14], prove that binary search is also a good option when it comes to implementing high-speed designs. Their paper proposes an 8x interleaved 6-bit ADC running at 25GS/s (the sub-ADC runs at 3.125GS/s) in which 2 bits are determined first in a process the authors have named “soft-selection” by using a set of 5 comparators with references across the full scale of the ADC. The design is based on the observation that in a flash ADC, only a limited number of comparators are critical in terms of decision time, while the others generate the output bits with little effort. With this in mind, the second stage is comprised of a 6-bit flash ADC with 9 comparator banks. Based on the results of the soft-selection, one of the 9 banks is selected because the result of all the others can be determined later, knowing that the output of a flash ADC is thermometer-coded.

The linearity of the designs of [12] and [14] is specified for the time-interleaved ADC only and thus corresponds to the Nyquist frequency of the interleaved converter. This makes a direct comparison non-trivial, but some conclusions can still be drawn with respect to the linearity. In the case of [12], the synchronous algorithm does not allow the loop to be sized for a very small input signal of the comparator, otherwise the speed will have to suffer tremendously. The same is true for [8], which applies asynchronous processing, but has a single-bit/cycle approach, which means that in one sampling period 8 bits need to be resolved. The speed targeted by the sub-ADC of Cai et al., [14], does not allow a long decision time for the comparators either, which means that the ADC is more likely to run into metastability issues.

It can be concluded that a 2.3GS/s sampling speed has been achieved while maintaining excellent linearity compared to other state-of-the-art designs. Owing to the input bandwidth of over 10GHz, the ADC can indeed be used in a time-interleaved design running at 20GS/s if an interleaving factor of 10x is considered, which makes it suitable for the application envisioned.

7. Conclusion

The thesis demonstrated that designing a single-channel 2.3GS/s ADC is possible and that a 3-bit/cycle asynchronous SAR is the best-suited architecture to do so. By opting for asynchronous processing in the SAR loop, the critical path could be sized in a more efficient manner than in the conventional SAR ADC because the unused time from shorter comparisons could be allotted to the slower ones.

The efficiency obtained lies within the 350fJ/conv-step range and the design described is predominantly digital, so it can be ported to more advanced technology nodes with little effort and is, in this sense, “future-proof”. The speed achieved pushes the limits of the 40nm CMOS technology available, as every block used was optimized for speed. The operating frequency expected from the architecture study was optimistic and it was shown that even if minimum-delay dynamic digital blocks are used, and the maximum sampling frequency obtained during the schematic simulations is of 2.5GS/s. The comparator used throughout the design was laid out and the parasitic elements associated with it were extracted using post-layout simulations, so the estimation of the speed degradation places the final operating frequency of the circuit at 2.3GS/s after the circuit is manufactured.

Key Contributions

The main contributions of this thesis are as follows:

- The ADC implemented proved that the SAR topology, although traditionally associated with low-to-medium-speed applications, is a viable and more efficient alternative to the flash ADC even when high speed is required.
- The architecture study compared the speed and power performance of various flash, SAR and pipelined ADC topologies and showed that a 3-bit/cycle asynchronous SAR ADC is the best option to achieve a sampling frequency of 2.3GS/s when assuming that the same building blocks are available.
- The behavioral modelling of the chosen architecture revealed the system-level limitations on the speed imposed by the event-driven control logic, which required a significant amount of delay to generate the internal clock signals.
- The mathematical analysis of the DAC settling time elaborated on the influence of the parasitic capacitances of the transistors in limiting the speed and demonstrated that unless significant power is spent in controlling large switches, a high enough sampling rate cannot be achieved.
- A complete transistor-level implementation of the digital logic led to the understanding that despite the fact that few blocks are needed to realize the functionality of the SAR loop, a lot of tapering is required to be able to drive the DAC switches or the comparators, which further limits the speed of the converter.

- Asynchronous processing was shown to be beneficial for dealing with metastable events because it allowed slow comparisons to be performed in a longer time than a synchronous SAR ADC would permit, thus increasing the accuracy of the converter.
- The technological limitations of the maximum sampling frequency of a 6-bit, 3-bit/cycle SAR ADC were studied and explained thoroughly.

Future Work

After finalizing the implementation of the circuit, a few ideas for improving it sprung to mind. However, due to the time limitations, they could not be explored further but offer interesting avenues for further research and are thus listed below:

1. Interpolation could be used to reduce the number of comparators and preamplifiers in the circuit, which would lead to a lower power consumption because the circuit will be able to operate with a lower DAC capacitance and will feature a smaller number of transistors.
2. Reusing the first stage of comparators might be possible seeing that only 50ps are needed to reset them. This solution would trade sampling speed (the comparator reset time will now be part of the critical path) to gain a lower power consumption.
3. Last, but not least, completing the layout of the circuit would bring the design one step closer to implementation in the 20GS/s ADC.

8. Bibliography

- [1] Y. Chiu, B. Nicolic and P. Gray, "Scaling of Analog-to-Digital Converters into Ultra-Deep-Submicron CMOS," in *Custom Integrated Circuits Conference*, San Jose, 2005.
- [2] A.-J. Annema, B. Nauta, R. van Langevelde and H. Tuinhout, "Analog Circuits in Ultra-Deep-Submicron CMOS," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 1, pp. 132-143, 2005.
- [3] B. Jonsson, "On CMOS Scaling and A/D-Converter Performance," in *NORCHIP*, Tampere, 2010.
- [4] C. Enz and E. Vittoz, "CMOS Low-Power Analog Circuit Design," in *Designing Low Power Digital Systems, Emerging Technologies*, Atlanta, 1996.
- [5] M. Verhelst and B. Murmann, "Area scaling analysis of CMOS ADCs," *Electronics Letters*, vol. 48, no. 6, 2012.
- [6] B. E. Jonsson, "Using Figures-of-Merit to Evaluate Measured A/D-Converter Performance," in *International Workshop on ADC Modelling, Testing and Data Converter Analysis and Design and IEEE ADC Forum*, Orvieto, 2011.
- [7] L. Kull, "Challenges in implementing high-speed, low-power ADCs in CMOS," in *Optical Fiber Communications Conference and Exhibition (OFC)*, Los Angeles, 2015.
- [8] L. Kull, T. Toifl, M. Schmatz, P. A. Francese, C. Menolfi, M. Braendli, M. Kossel, T. Morf, T. M. Andersen and Y. Leblebici, "A 3.1 mW 8b 1.2 GS/s Single-Channel Asynchronous SAR ADC With Alternate Comparators for Enhanced Speed in 32 nm Digital SOI CMOS," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 12, pp. 3049-3058, 2013.
- [9] L. Kull, J. Pliva, T. Toifl, M. Schmatz, P. A. Francese, C. Menolfi, M. Braendli, M. Kossel, T. Morf, T. M. Andersen and Y. Leblebici, "A 110mW 6 bit 36 GS/s interleaved SAR ADC for 100 GBE occupying 0.048mm² in 32 nm SOI CMOS," in *IEEE Asian Solid-State Circuits Conference*, Kaohsiung, Taiwan, November 10 - 12, 2014.
- [10] B. Razavi, "Design Considerations for Interleaved ADCs," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 8, pp. 1806-1817, 2013.
- [11] L. Sumanen, M. Waltari and K. Halonen, "A 10-bit 200-MS/s CMOS Parallel Pipeline A/D Converter," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 7, pp. 1048-1055, 2001.
- [12] C.-H. Chan, Y. Zhu, S.-W. Sin, S.-P. U and R. Martinis, "A 5.5mW 6b 5GS/s 4x-Interleaved 3b/cycle SAR ADC in 65nm CMOS," in *IEEE International Solid-State Circuits Conference*, San Francisco, 2015.

Bibliography

- [13] Shafik and et al., "A 10Gb/s Hybrid ADC-Based Receiver with Embedded 3-Tap Analog FFE and Dynamically-Enabled Digital Equalization in 65nm CMOS," in *IEEE Solid-State Circuits Conference*, San Francisco, 2015.
- [14] S. Cai, E. Z. Tabasy, A. Shafik, S. Kiran, S. Hoyos and S. Palermo, "A 25GS/s 6b TI Binary Search ADC with Soft-Decision Selection in 65nm CMOS," in *IEEE Symposium on VLSI Circuits*, Kyoto, 2015.
- [15] V. Chen and L. Pilleggi, "A 69.5mW 20GS/s 6b Time-Interleaved ADC with Embedded Time-to-Digital Calibration in 32nm CMOS SOI," in *IEEE International Solid-State Circuits Conference*, San Francisco, 2014.
- [16] V. Chen and L. Pileggi, "An 8.5mW 5GS/s 6b flash ADC with dynamic offset calibration in 32nm CMOS SOI," in *IEEE Symposium on VLSI Circuits*, Kyoto, 2013.
- [17] B. Zhang, A. Nazemi, A. Garg, N. Kocaman, M. Ahmadi, M. Khanpour, H. Zhang, J. Cao and A. Momtaz, "A 195mW/55mW Dual-Path Receiver AFE For Multistandard 8.5-11.5Gb/s Serial Links in 40nm CMOS," in *IEEE International Solid-States Circuits Conference*, San Francisco, 2013.
- [18] Verbruggen and et al., "A 2.6mW 6b 2.2GS/s 4-times Interleaved Fully Dynamic Pipelined ADC in 40nm Digital CMOS," in *IEEE International Solid-State Circuits Conference*, San Francisco, 2010.
- [19] J. Cao, B. Zhang, S. Ullas, D. Cui, A. Vasani, A. Garg, W. Zhang, N. Kocaman, D. Pi, B. Raghavan, H. Pan, I. Fujimori and A. Momtaz, "A 500mW digitally calibrated AFE in 65nm CMOS for 10Gb/s Serial links over backplane and multimode fiber," in *IEEE International Solid-State Circuits Conference*, San Francisco, 2009.
- [20] M. Choi, J. Lee, J. Lee and H. Son, "A 6-bit 5-GSample/s Nyquist A/D Converter in 65nm CMOS," in *IEEE Symposium on VLSI Circuits*, Honolulu, 2008.
- [21] K. Deguchi, N. Suwa, M. Ito, T. Kumamoto and T. Miki, "A 6-bit 3.5-GS/s 0.9-V 98-mW Flash ADC in 90nm CMOS," in *IEEE Symposium on VLSI Circuits*, Kyoto, 2007.
- [22] C. Paulus, H.-M. Bluthgen, M. Low, E. Sicheneder, N. Bruls, A. Courtois, M. Tiebout and R. Thewes, "A 4GS/s 6b flash ADC in 0.13 μm CMOS," in *IEEE Symposium on VLSI Circuits*, 2004.
- [23] X. Jiang, Z. Wang and M.-C. F. Chang, "A 2 GS/s 6 b ADC in 0.18 μm CMOS," in *IEEE International Solid-State Circuits Conference*, San Francisco, 2003.
- [24] C.-C. Huang, C.-Y. Wang and J.-T. Wu, "A CMOS 6-Bit 16-GS/s Time-Interleaved ADC Using Digital Background Calibration Techniques," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 4, pp. 848-858, 2011.
- [25] A. Spagnolo, B. Verbruggen, P. Wambacq and S. D'Amico, "A 4.1-mW 3.5-GS/s 6-Bit Time-Interleaved ADC in 40-nm CMOS," *IEEE Transactions on Circuits and Systems - II: Express Briefs*, vol. 61, no. 7, pp. 466-470, 2014.

- [26] A. Ramkaj, "Analysis and Design of High-Speed Successive Approximation Register ADCs, MSc Thesis," Delft University of Technology, Delft, The Netherlands, 2014.
- [27] B. Wicht, T. Nirschl and D. Schmitt-Landsiedel, "Yield and Speed Optimization of a Latch-Type Voltage Sense Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp. 1148-1158, 2004.
- [28] W. Kester, "Analog Devices - MT-021 Tutorial - ADC Architectures II: Successive Approximation ADCs," 2009. [Online]. Available: <http://www.analog.com/media/en/training-seminars/tutorials/MT-021.pd>. [Accessed 5 May 2016].
- [29] S.-W. M. Chen and R. Brodersen, "A 6-bit 600-MS/s 5.3-mW Asynchronous ADC in 0.13-um CMOS," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12, pp. 2669 - 2680, 2006.
- [30] J. Yang, T. L. Naing and R. Brodersen, "1 GS/s 6 bit 6.7 mW Successive Approximation ADC Using Asynchronous Processing," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 8, pp. 1469 - 1478 , 2010.
- [31] C.-H. Chan, Y. Zhu, S.-W. Sin, S.-P. Ben U and R. P. Martinis, "A 6 b 5 GS/s 4 Interleaved 3 b/Cycle SAR ADC," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 2, pp. 365-377, 2016.
- [32] M. Pelgrom, *Analog-to-Digital Conversion*, Dordrecht: Springer, 2013.
- [33] F. Maloberti, *Data Converters*, Dordrecht: Springer, 2007.
- [34] S. P. Astgimath, "A low noise, low power dynamic amplifier with a common mode detect and a low power, low noise comparator for pipelined SAR-ADC, MSc Thesis," Delft University of Technology, Delft, The Netherlands, 2012.
- [35] S. Louwsma, E. van Tuijl and B. Nauta, *Time-interleaved Analog-to-Digital Converters*, Springer: Dordrecht, 2011.
- [36] S. Limotyraakis , S. Kulchyski, D. Su and B. Wooley, "A 150MS/s 8b 71mW time-interleaved ADC in 0.18 μ m CMOS," in *IEEE International Solid-State Circuits Conference*, San Francisco, 2004.
- [37] B. Razavi, *Design of Analog CMOS Integrated Circuits*, New York: Mc-Graw Hill, 2008.
- [38] A. Oppenheim, A. Willsky and H. Nawab, *Signals and Systems*, New Jersey: Prentice-Hall, 1996.
- [39] T. Kobayashi, K. Nogami, T. Shirotori and Y. Fujimoto, "A Current-Controlled Latch Sense Amplifier and a Static Power-Saving Input Buffer for Low-Power Architecture," *IEEE Journal of Solid-State Circuits*, vol. 28, no. 4, pp. 523-527, 1993.
- [40] P. Nuzzo, F. De Bernardinis, P. Terreni and G. Van der Plas, "Noise Analysis of Regenerative Comparators for Reconfigurable ADC Architectures," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 6, pp. 1441-1454, 2008.

Bibliography

- [41] M. Pelgrom, A. Duimnajer and A. Welbers, "Matching Properties of MOS Transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433-1440, 1989.
- [42] M. Ganzerli, "CMOS Integrated Circuits for High-Speed Serial Links," University of Modena and Reggio Emilia, Modena.
- [43] S. Hashemi and B. Razavi, "Analysis of Metastability in Pipelined ADCs," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 5, pp. 1198-1209, 2014.
- [44] L. Kull, "High-Speed CMOS ADC Design for 100 Gb/s Communication Systems, PhD Thesis," Ecole Polytechnique Federale de Lausanne, Lausanne, 2014.
- [45] F. Radice, M. Bruccoleri, M. Ganzerli, G. Spelgatti, D. Sanzogni, M. Pozzoni and A. Mazzanti, "A 6-bit 6-GS/s 95mW background calibrated flash ADC with integrating preamplifiers and half-rate comparators in 32nm LP CMOS," in *ESSCIRC*, Bucharest, Romania, 2013.
- [46] Y.-Z. Lin, C.-W. Lin and S.-J. Chang, "A 2-GS/s 6-bit Flash ADC with Offset Calibration," in *IEEE Asian Solid-State Circuits Conference*, Fukuoka, Japan, 2008.
- [47] A. Varzaghani, A. Kasapi, D. Loizos, S.-H. Paik, S. Verma, S. Zogopoulos and S. Sidiropoulos, "A 10.3-GS/s, 6-Bit Flash ADC for 10G Ethernet Applications," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 12, pp. 3038-3048, 2013.