A Compact Sub-1V Capacitively-Biased BJT-Based Temperature Sensor

By

Sanja Kastratović

Supervisors:

Prof. dr. K.A.A. Makinwa

Dr. Z. Tang

In partial fulfilment of the requirements for the degree of

Master of Science in Electrical Engineering

At the Department of Microelectronics

Faculty of Electrical Engineering, Mathematics and Computer Science

Delft University of Technology

To be defended on November 30th, 2023

Student number:	5506239
Thesis Committee:	Prof. Dr. K.A.A. Makinwa
	Dr. D.G. Muratore
	Dr. K. Souri

Abstract

This thesis presents the design of a low-power, sub-1V, BJT-based temperature sensor. It is based on a capacitively-biased (CB) BJT front-end, in which capacitors are discharged across diode-connected BJTs to obtain proportional-to-temperature (PTAT) and complementary-to-temperature (CTAT) voltages. These voltages are then digitised by a switched-capacitor Delta-Sigma Modulator (DSM), which employs energy-efficient inverter-based amplifiers that can operate from a sub-1V supply. To mitigate the effects of component mismatch and 1/f noise, dynamic error-cancellation techniques such as chopping, auto-zeroing and dynamic element matching are used in both the front-end and the DSM. After a 1-point temperature calibration, the sensor achieves a simulated inaccuracy of ±0.20°C (3σ) over temperatures ranging from -55°C to 125°C. From simulations, it does this while operating from a 0.9V supply and dissipating only 270nW. It occupies an estimated area of 0.066mm^2 in a 180nm CMOS process. Compared to previous CB temperature sensors, this design occupies 3.8x smaller area and dissipates 3x less power.

Acknowledgements

The last two years have been a journey of personal and professional growth. Here, I would like to thank those who were there to share the happiest moments as well as help me push through the many difficulties encountered. Without them, I wouldn't have been able to reach this point in my life.

First and foremost, I would like to thank my supervisor, Prof. Dr. Kofi Makinwa, for giving me the opportunity to join the EI group and work on such an interesting project. During the past year, he shared his wisdom, knowledge and guidance, providing me with enough support and also independence to build myself into the engineer I am today.

Secondly, I would like to thank my daily supervisor Dr. Zhong Tang who provided enormous support throughout the past year and without whom this thesis couldn't have been completed. Even though he left the lab a few months into my thesis, he was always willing to have hours-long technical discussions and encourage me whenever I felt falling behind.

I would like to extend my enormous gratitude to SiTime, specifically Dr. Kamran Souri, soon-to-be Dr. Jan Angevare and Dr. Saleh Shalmany, for stepping in when I felt completely lost about my thesis and giving me the helping hand to move forward. They invested invaluable time and I am grateful for all the knowledge I gained from them. It was an incredible experience having both input from an academic and industrial perspective and I hope to keep this in my mind moving forward in my career.

I want to thank the Microelectronics Department for the scholarship that allowed me to come to Delft in the first place and pursue this degree.

I want to thank the EI team, all the PhDs and other MSc students on the 15th floor, for the shared discussions, laughter, food, and late nights filled with work. Among them, I want to extend special thanks to Avish, Jida, Shenyang, Giorgos, Sundeep, Nandor and Miloš. I want to thank all the friends I met in Delft, for sharing all the good memories and making me feel at home in this city: Shruti, Nivedita, Chieh, Maciej, Ilker, Andrea, Deb, Sari, and the list goes on. Additionally, I want to thank my friends scattered across the world who provided me with emotional support throughout these two years despite the time zones: Nouf, Mariam, Mao, Ahmad, Hadi, Milena and Adrian.

Last but not least, I want to thank my family for always having my back, never judging me and always believing in me. I wouldn't have been here without their continuous love and support.

Table of Contents

1 Introduction 11
1.1 Motivation 11
1.2 Types of CMOS temperature sensors 11
1.2.1 Operating principle of BJT-based temperature sensors 12
1.3 Capacitive Biasing 13
1.4 CBD-based temperature sensors 15
1.4.1 State-of-the-art CBD-based temperature sensor 17
1.5 Project goals 20
1.6 Organisation of the thesis 21
2 System level design 23
2.1 Reducing the area of the CBD SC temperature sensor 23
2.2 Scaling the front-end 24
2.3 Charge balancing readout scheme 28
2.4 Scaling the autozero capacitors C_{AZ} 29
2.3.1 Amplifier choppers 32
2.5 Summary 35
3 Circuit implementation 36
3.1 CBD Front-end 36
3.1.1 Scaling down the PNP 36
3.1.2 Scaling down the front-end sampling capacitor, C_S 39
3.1.3. Front-end biasing voltage generator 43
3.2 Delta-sigma modulator readout 44
3.2.1 DSM topology 44
3.2.3 Amplifier design 44
3.2.3 Implementation of the 2nd integrator 51
3.3 Timing control 52
3.3.1 Clock boosters 53
3.3 Area estimation 55
3.4 Summary 56
4 Simulation results 57

4.1 Modulator behaviour 57
4.2 Accuracy 59
4.2.1 Power supply sensitivity of the front-end 62
4.3 Layout and area estimation 63
4.4 Power dissipation 64
4.5 Resolution 66
4.5.1 Figure of merit 67
4.6 Summary 68
5 Conclusion and future work 69
5.1 Conclusion 69
5.2 Future work 69
5.2.1 Reducing the conversion period 69
5.2.2 Reducing digital power 69
5.2.3 Reducing analog power 69
Bibliography 74

List of Figures

Figure 1.1 Working principle of a BJT-based temperature sensor. 12
Figure 1.2 CTAT and PTAT behaviour of VBE and ΔVBE [4]. 13
Figure 1.3 a) Static-biased diode; b) Capacitively-biased diode. 14
Figure 1.4 Working principle of a capacitively-biased diode [1]. 14
Figure 1.5 Switch-capacitor CBD-based sensor circuit [2]. 16
Figure 1.6 CBD-based temperature sensor with DTMOST as sensing element: a) VCTAT generator; b) DTMOST sensing code; c) relaxation oscillator [7]. 17
Figure 1.7 Block diagram of a CBD-based temperature sensor [13]. 17
Figure 1.8 Single-sided switched-capacitor (SC) integrator with embedded CB PNP front-end [1]. 18
Figure 1.9 Block level architecture of the CBD sensor proposed in [1]. 19
Figure 1.10 Measured bitstream average $3\Delta VBEVBE1$ over temperature of chip [1]. 19
Figure 1.11 Measured bitstream average $\mu = 3\Delta VBEVBE$ across temperature for a time ratio p=32 and p=64. 20
Figure 2.1 Die micrograph of chip in [1]. 23
Figure 2.2 Area distribution of the chip in [1] in mm^2. 24
Figure 2.3 Current through a diode (ID) versus discharge time (t) for different CS sizes from 500fF to 16pF. 25
Figure 2.4 BJT's β variation with respect to emitter current IE. 26
Figure 2.5 a) Temperature error in ΔVBE at different Ibias currents; b) Static biasing of two BJTs. 27
Figure 2.6 Simulated value of $\mu = 3\Delta VBEVBE$ and $\mu = 2\Delta VBEVBE$ across temperature, for p = 32. 28
Figure 2.7 Proposed charge balancing scheme. 28
Figure 2.8 a) Front-end and the inverter-based integrator and b) their timing diagram. 29
Figure 2.9 Single-sided inverter-based amplifier in a) autozero phase; b) integration phase. 30
Figure 2.10 Noise folding due to autozeroing. 31
Figure 2.11 Chopper added to the first integrator: a) autozero phase; b) integration phase. 32
Figure 2.12 Noise power spectral density (PSD) versus frequency with a) autozero; b) autozero and chopping for fCH < fAZ ; c) autozero and chopping for fCH ≥ fAZ. 32
Figure 2.13 Chopping sampled noise on CAZ in the time domain. 33
Figure 2.14 a) Integrator behaviour; b) Timing diagram. 35
Figure 3.1 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with ideal switches and C_S = 490fF, VDD=0.9V and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom). 36
Figure 3.2 a) I_E and b) V_{BE} versus time for PNPs of different areas with C_S = 490fF and VDD=0.9V across temperature. 37
Figure 3.3 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with real switches and C_S = 490fF, VDD=0.9V and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom) for a) 2x2 PNP; b) 10x10 PNP. 38
Figure 3.4 Decimated output over temperature for a 2x2, 5x5 and 10x10 PNP device with C_S = 490fF and p = 32 for a) VDD = 0.9V and b) VDD = 0.95V. 38
Figure 3.5 CBD Front-end with NPN device. 39

Figure 3.6 Measured inaccuracy of a 1x1 NPN-based CBD front end without (top) and with a 1-point PTAT trim at 25°C (bottom). *39*
Figure 3.7 Measured inaccuracy of a 5x5 PNP-based CBD front end with Cs = 0.8pF without (top) and with a 1-point PTAT trim at 25oC (bottom). *40*
Figure 3.8 Single-sided front-end with switches a) representation; b) real implementation. *41*
Figure 3.9 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with real switches and C_S = 800fF and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom) for a) VDD = 0.9V and b) VDD = 0.95V. *42*
Figure 3.10 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with real switches and CS = 800fF and p = 32, for halved discharge time t1 and t2 without (top) and with a 1-point PTAT trim at 25oC (bottom) for VDD = 0.9V. *43*
Figure 3.11 Front-end biasing circuit and its timing diagram. *43*
Figure 3.12 DSM Block architecture. *44*
Figure 3.13 Switch-capacitor implementation of first integrator. *45*
Figure 3.14 Single-sided inverter-based amplifier in a) autozero phase; b) integration phase. *46*
Figure 3.15 a) Monte Carlo amplifier's DC gain; b) Amplifier's input-referred noise at room temperature. *47*
Figure 3.16 OTA Biasing circuit. *47*
Figure 3.17 First integrator with choppers. *49*
Figure 3.18 CMFB circuit during a) Φ1 and b) Φ2. ..... *50*
Figure 3.19 a) Front-end VCM shift circuit diagram; b) Voltage across CS and parasitic capacitors during sampling and integration phase. *51*
Figure 3.20 Second integrator. *52*
Figure 3.21 System-level chopping. *53*
Figure 3.22 Dynamic element matching (DEM) timing diagram. *53*
Figure 3.23 Circuit schematic of the implemented single-capacitor clock booster [16]. *54*
Figure 3.24 Circuit schematic of the clock boosters in [1]. *54*
Figure 3.25 Area distribution in mm^2 of each block in a) design in [1]; b) this design. *55*
Figure 4.1 Output swing of the first integrator across corner and temperature. *57*
Figure 4.2 Output swings of the a) 2nd integrator; and b) the SC adder across corners and temperature. *58*
Figure 4.3 Output VCM of the a) 1st integrator and b) 2nd integrator across temperature and corner. *59*
Figure 4.4 Input VCM of the front-end across temperature and corner. *59*
Figure 4.5 Monte Carlo (20 points, mismatch and process spread) inaccuracy of the front-end and first integrator (without OTA choppers) with C_S = 800fF, V_{DD}=0.9V and p = 32 without trimming (top) and with a 1-point PTAT trim at 25°C (bottom). *60*
Figure 4.6 Monte Carlo (20 points, mismatch and process spread) inaccuracy of the front-end and first integrator with OTA choppers with C_S = 800fF, V_{DD}=0.95V and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom). *61*
Figure 4.7 Monte Carlo (20 points, mismatch and process spread) inaccuracy of the front-end and first integrator with OTA choppers with CS = 800fF, VDD=0.95V and p = 32 without (top) and with a 1-point PTAT trim at 25oC (bottom). *61*
Figure 4.8 Monte Carlo (20 points, process spread) simulation of power-supply sensitivity (PSS) at a) SS - 55C; b) TT 27oC; c) FF 125oC. *62*

Figure 4.9 Monte Carlo (20 points, process spread) simulation of power-supply sensitivity (PSS) with halved discharge time at a) SS -55C; b) TT 27oC; c) FF 125oC. *63*
Figure 4.10 Area comparison of design in [1] (top) and estimated layout of this design (bottom). *64*
Figure 4.11 Simulated analogue power consumption breakdown per block at 0.9V supply. *65*
Figure 4.12 Simulated total power consumption breakdown per block at 0.9V supply. *65*
Figure 4.13 Modulator's frequency response with and without chopping. *66*
Figure 4.14 Resolution versus conversion time. *67*
Figure 4.15 Figure of Merit versus conversion time. *67*
Figure 5.1 a) Schematic of the first amplifier and b) choppers timing signal. *70*
Figure 5.2 Opreation of the proposed ping-pong front-end with its timing diagram. *72*

List of Tables

Table 1.1 Comparison of the state-of-the-art low-supply temperature sensors. *21*
Table 3.1 Coefficients for the implemented DSM. *44*
Table 3.2 Dominant capacitor values in the original [1] and this design. *56*
Table 4.1 Output swings of the a) 1st integrator; b) 2nd integrator; c) switched-capacitor adder across temperature and corner. *58*
Table 4.2 Comparison of power consumption between the original [1] and this design at VDD = 1V. *66*
Table 4.3 Performance summary and comparison to the state-of-the-art. *68*

List of Abbreviations

ADC	analog-to-digital converter
BJT	bipolar junction transistor
BS	bitstream
CB	capacitively-biased
CBD	capacitively-biased diode
CMFB	common mode feedback
CTAT	complementary to absolute temperature
DEM	dynamic element matching
DSM	delta-sigma modulator
FoM	figure of merit
GBW	gain bandwidth
HVT	high threshold voltage
IoT	internet of things
OSR	oversampling ratio
OTA	operational transconductance amplifier
PSD	power spectral density
PSS	power supply sensitivity
PTAT	proportional to absolute temperature
RIA	relative inaccuracy
SC	switched-capacitor
S(Q)NR	signal to (quantisation) noise ratio
TD	thermal diffusivity

1 Introduction

1.1 Motivation

Temperature is one of the most extensively measured environmental parameters due to its influence on the behaviour of nearly all physical, chemical, mechanical, and biological systems. Therefore, temperature sensors find broad application in various fields, including healthcare and the automotive industry, particularly for the purpose of correcting temperature-induced inaccuracies in analog circuits. Given the accuracy, low cost and digital output capabilities of temperature sensors in modern CMOS technologies, significant research effort has been dedicated to their design. Recently, the advent of the Internet of Things (IoT) has heightened the interest in ultra-low power sensors that can maintain good accuracy across a wide temperature range. Hence, the primary focus of this thesis is the design of such sensors in CMOS technology.

1.2 Types of CMOS temperature sensors

Within CMOS technology, many circuit elements are available for temperature sensing. The selection of a sensing element then depends on the specific requirements dictated by the intended application, encompassing factors such as accuracy, resolution, power consumption and cost.

For decades, bipolar junction transistors (BJTs) have been extensively used as temperature sensors and bandgap voltage references. When a BJT is biased at a fixed current, its base-emitter voltage (V_{BE}) demonstrates complementary-to-absolute temperature (CTAT) behaviour. By biasing two BJTs at different current densities, the difference between their base-emitter voltages, denoted as $\Delta V_{BE} = V_{BE1} - V_{BE2}$, exhibits proportional-to-absolute temperature (PTAT) behaviour. Due to the exponential relationship between the BJT's collector current I_C and V_{BE}, this approach also cancels the effect of current variations, rendering ΔV_{BE} relatively insensitive to process variations. As it turns out, the spread in V_{BE} can be effectively corrected through a low-cost, one-point PTAT trim [4], resulting in inaccuracies as low as 0.06°C (3σ) from -55°C to 125°C. Using the box method, i.e. by dividing the peak-to-peak inaccuracy (3σ) by the temperature range, this translates into a relative inaccuracy (RIA) of 0.06% [9].

When biased in the sub-threshold region, a MOSFET's drain current (I_D) is a temperature-dependent exponential function of its gate-source voltage (V_{GS}). Typically, $V_{GS} < V_{BE}$, enabling MOSFET-based temperature sensors to function at lower supply voltages than their BJT-based counterparts. However, I_D is determined by the oxide capacitance (C_{OX}) and the threshold voltage (V_{TH}), both of which are spread with process. Consequently, MOSFET-based sensors usually require a two-point temperature calibration to achieve low inaccuracies, such as 0.1% [12].

In recent studies [10], it has been demonstrated that resistor-based temperature sensors can achieve higher resolution and energy efficiency than either BJT or MOSFET-based sensors. They can also operate from sub-1V supplies. Nevertheless, to achieve competitive accuracy, they require at least two-point

trimming due to the sensitivity of resistors to process spread. After a two-point trim, relative inaccuracies as low as 0.03% can be achieved [10], whereas a single-point trim results in only 0.56% [10].

Thermal diffusivity (TD) sensors measure the delay of heat pulses propagating through a silicon substrate. This delay is a function of the thermal diffusivity of silicon, a well-defined parameter that exhibits a significant dependence on temperature. TD sensors use heating elements to generate heat pulses, which then diffuse through the silicon substrate and are detected by a temperature sensor, usually a thermopile. A major drawback of TD sensors is their power consumption, typically a few milliwatts, which renders them unsuitable for low-power applications [4]. Nevertheless, they can achieve relative inaccuracies of 0.2% without trimming [11].

Of the four types of sensors, BJT-based temperature sensors are the most widely used because they can achieve excellent accuracy after a low-cost, one-point calibration with a relatively low power consumption.

1.2.1 Operating principle of BJT-based temperature sensors

The base-emitter voltage (V_{BE}) of a BJT, where $I_C >> I_S$, can be expressed as:

$$V_{BE} = \frac{\eta kT}{q} \ln\left(\frac{I_C}{I_S}\right) \tag{1.1}$$

where η is the diode non-ideality factor (for an ideal diode η=1), k is the Boltzmann constant (1.38×10^{-23} J/K), T is the absolute temperature in Kelvin, q is the electron charge (1.6×10^{-19} C), I_S is the saturation current, and I_C is the collector current of the BJT. Due to the strong temperature-dependence of I_S (~ T^4), V_{BE} exhibits a temperature coefficient of approximately -2 mV/°C, leading to its CTAT characteristic.

When two BJTs are biased at different current densities (*Figure 1.1*), the difference in their base-emitter voltages, ΔV_{BE}, demonstrates PTAT behaviour as:

$$\Delta V_{BE} = \frac{\eta kT}{q} \ln\left(\frac{I_{C1}}{I_{C2}}\right) \tag{1.2}$$

It can be seen that ΔV_{BE} only depends on the ratio of the biasing currents, $p = \frac{I_{C1}}{I_{C2}}$, which can be accurately defined by good layout, as well as circuit techniques such as dynamic-element matching.

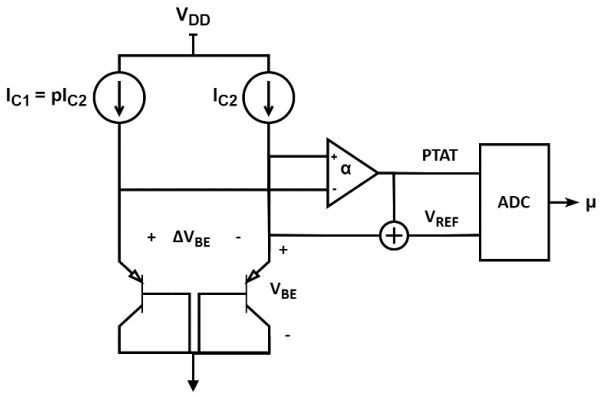


Figure 1.1 Working principle of a BJT-based temperature sensor.

In general, the front-end of a BJT-based temperature sensor employs two diode-connected BJTs, as shown in *Figure 1.1*. With V_{BE} and ΔV_{BE} established, a temperature-independent reference voltage (V_{REF}) (*Figure 1.2*) can be generated as:

$$V_{REF} = V_{BE} + \alpha \Delta V_{BE}$$

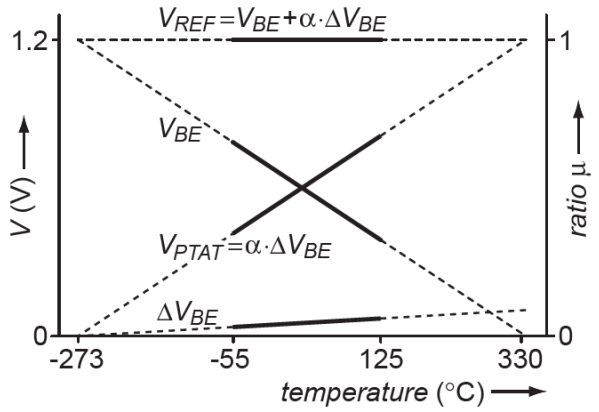


Figure 1.2 CTAT and PTAT behaviour of VBE and ΔVBE [4].

Using the PTAT voltage ΔV_{BE} as the input and V_{REF} as the reference for an analog-to-digital converter (ADC), its digital output can be expressed as:

$$\mu = \frac{V_{PTAT}}{V_{REF}} = \frac{\alpha \Delta V_{BE}}{V_{BE} + \alpha \Delta V_{BE}} \tag{1.3}$$

A final digital output, D_{out}, in degrees Celsius can be acquired through linear scaling:

$$D_{out} = A\mu + B \tag{1.4}$$

Where A~600K, B~-273K and the value of α typically ranges between 8 and 20 [4].

1.3 Capacitive Biasing

In conventional BJT- or diode-based temperature sensors, the BJTs (or diodes) are biased by current sources (*Figure 1.3a*). In addition to V_{BE}, the headroom they require increases the sensor's minimum supply voltage to approximately 1.2V. Alternatively, the BJT (or diode) can be biased by a capacitor pre-charged to V_{DD}, as depicted in *Figure 1.3b*. In this setup, the capacitor functions as a dynamic current source devoid of 1/*f* noise and requires little headroom.

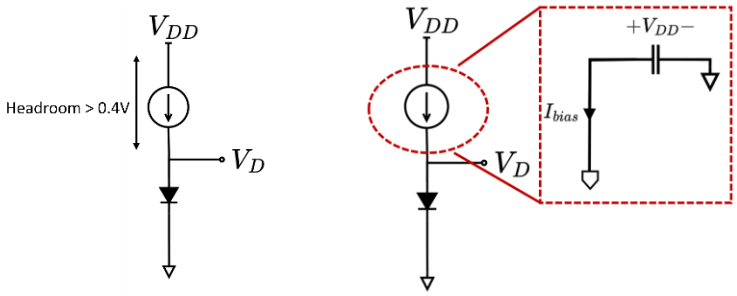


Figure 1.3 a) Static-biased diode; b) Capacitively-biased diode.

The working principle of a temperature sensor based on the capacitively-biased diode (CBD) technique is illustrated in *Figure 1.4*. In an initial phase (rst), a capacitor (C) is pre-charged to V_{DD}. In a second phase, the capacitor is connected across the diode. After a brief period (tens of nanoseconds), the voltage across the capacitor (V_D) will be a logarithmic function of time. This relationship can be expressed as [2]:

$$V_D(t) \simeq -\eta V_T \ln\left(\frac{I_S}{C\eta V_T}t\right) \tag{1.5}$$

where η is the diode ideality factor (1 for an ideal diode, ~2 for silicon diodes with small currents), V_T is the thermal voltage and I_S is the diode saturation current. Likewise, the diode current (I_D) is well-defined and proportional to $1/t$ [2].

$$I_D(t) \simeq \frac{\eta C V_T}{t} \tag{1.6}$$

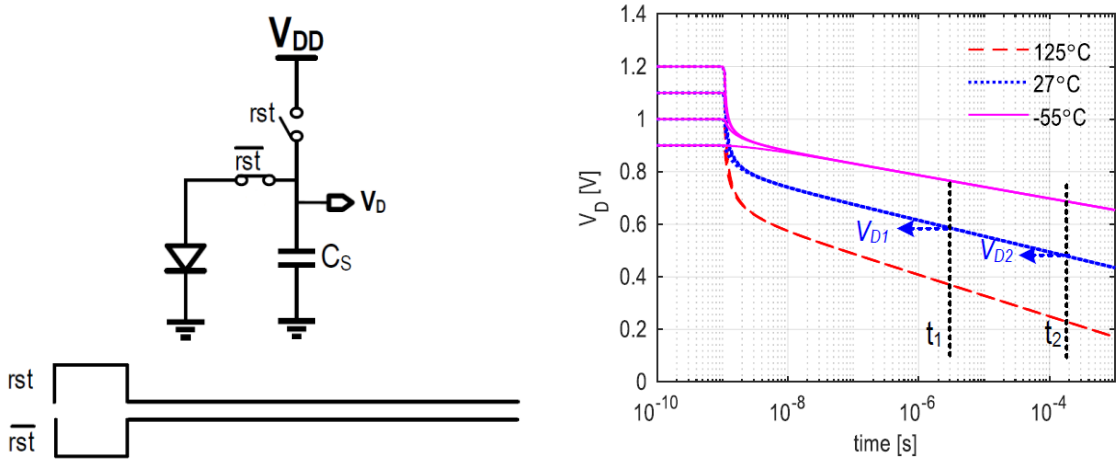


Figure 1.4 Working principle of a capacitively-biased diode [1].

A notable feature of the CBD technique is that, following the initial settling period, V_D becomes independent of the initial voltage (V_{DD}). The exact expression for the voltage across the diode is [3]:

$$V_D(t) = -\eta V_T \ln\left[1 - \left(1 - exp\left(-\frac{V_{DD}}{\eta V_T}\right)\right) exp\left(-\frac{I_S}{C\eta V_T} t\right)\right] \quad [2] \tag{1.7}$$

As depicted in *Figure 1.4*, we can identify two primary regions in the discharging process. Firstly, when

$$-\frac{I_S}{C\eta V_T} t \ll -\frac{V_{DD}}{\eta V_T} \quad [2] \tag{1.8}$$

the voltage on the diode is approximately V_{DD}. Secondly, when

$$-\frac{I_S}{C\eta V_T} t > -\frac{V_{DD}}{\eta V_T} \quad [2] \tag{1.9}$$

the exponential term containing V_{DD} becomes insignificant, and V_D undergoes a logarithmic decay independent of the initial voltage. This implies that sampling V_D in this region yields a voltage that is insensitive to supply voltage noise.

Additionally, the minimum supply voltage is only slightly higher than the diode's forward voltage drop, which is approximately 0.8V at cold temperatures. No additional headroom is necessary for other components, such as current sources. Moreover, in comparison to current source biasing, a CBD-based front-end requires fewer components, potentially enhancing accuracy.

To extract temperature information, the temperature-dependent voltage is sampled on the capacitor. At a specific time, t, the diode is disconnected from the capacitor, sampling V_D on C. When the diode is replaced by a diode-connected BJT, the voltage $V_D = V_{BE}$ is shown by Equation *(1.1)*. By sampling the voltage at two different time instances, t_1 and t_2, and taking their difference, a PTAT voltage, ΔV_{BE} can be obtained.

$$\Delta V_{BE} = \frac{kT}{q} \ln\left(\frac{I_{C1}}{I_{C2}}\right) \simeq \eta \frac{kT}{q} \ln\left(\frac{t_2}{t_1}\right) \tag{1.10}$$

Here, it is evident that the PTAT voltage (ΔV_{BE}) is dependent on the time ratio, t_2/t_1, rather than on the absolute discharge time. As explained in the previous subsection, the sampled voltages V_{BE} and ΔV_{BE} can then be used to obtain the temperature in degrees Celsius.

A drawback of the CBD technique is the kT/C noise associated with sampling V_{BE} on a capacitor. Additionally, it is sensitive to variations in the master clock period, since the bias current in the diode is set by the discharge time of the capacitor. However, it is worth noting that such errors will only impact V_{BE} and not ΔV_{BE}, which is defined by a time ratio. The clock sensitivity of CBD-based temperature sensors is ~0.13 °C/% of clock variation, which is not a major source of error considering that on-chip clocks with inaccuracies below 0.1% can be achieved.

1.4 CBD-based temperature sensors

The first CBD-based temperature sensor was presented in [2]. As shown in *Figure 1.5*, its temperature sensing element is a bulk diode. During a reset phase, switches M1, M2, M5 and M6 are closed, leading to the charging of capacitors C_1 and C_2. Afterwards, C_1 and C_2 are discharged by closing switches M3 and M4 during the time intervals ts_1 and ts_2, respectively. At the end of ts_1, the sampled voltage $V(C_1)$ is scaled by a capacitive DAC and transferred to C_{az}, while the comparator is auto-zeroed. At the end of ts2, the sampled voltage $V(C_2)$ is compared with the result, during the comparison step (tc). Using an 8-step SAR

algorithm, the setting of the capacitor DAC is adjusted until $V(C_1) \sim V(C_2)$. The required DAC setting is then a digital representation of temperature.

Operating from supply voltages as low as 0.85V, the sensor achieves a relative inaccuracy of 2.8%. It is also quite compact, occupying an area of only 0.0025 mm^2, partially due to its implementation in a 16-nm FinFET technology. However, its power supply sensitivity (PSS) is on the high side at 1.5°C/V and it dissipates 18µW [2].

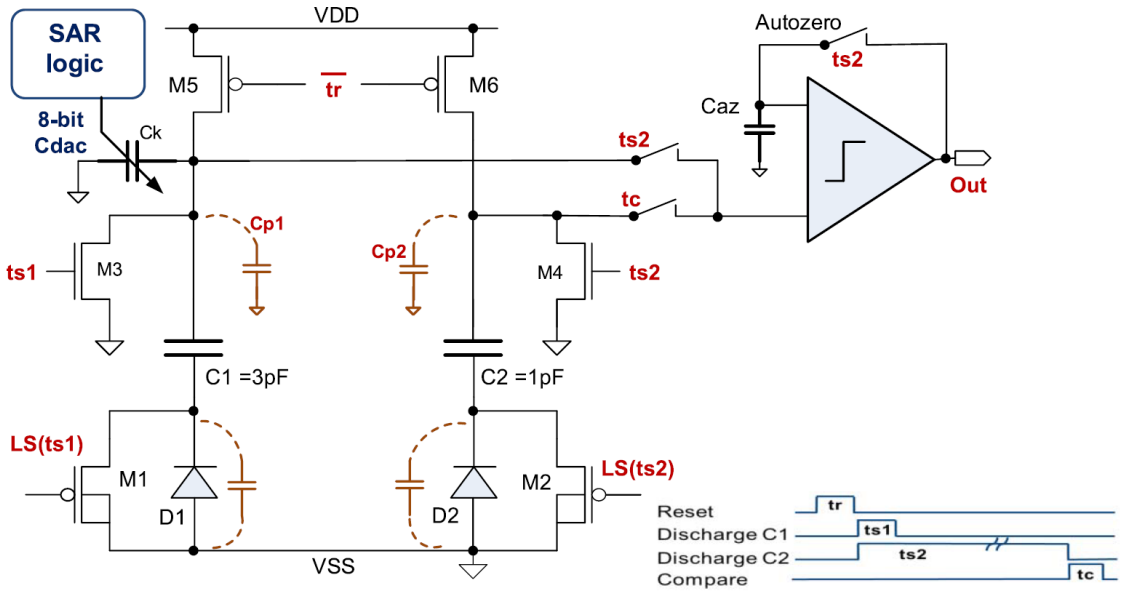


Figure 1.5 Switch-capacitor CBD-based sensor circuit [2].

Capacitive biasing can be applied to diodes, BJTs and MOSFETs. The design illustrated in *Figure 1.6* is based on a DTMOST, enabling operation at a supply voltage of 0.85V and achieving a supply sensitivity of 0.27°C/V. It operates by discharging the capacitor C across the diode-connected DTMOSTs and sampling the temperature-dependent voltage V_C (*Figure 1.6b*). This voltage is then compared to a reference voltage V_{CTAT} provided by a conventional CTAT generator circuit (*Figure 1.6a*). Lastly, the comparator detects when V_C reaches V_{CTAT} and generates the output clock signals (*Figure 1.6c*), that control the charge and discharge times of the sensing core.

However, the design achieves a poor relative inaccuracy of 1.8% with one-point trim, mainly due to the spread in the DTMOSTs present in both the sensing core and in the V_{CTAT} generator. Furthermore, at cold temperatures, the slowly-decaying CBD-voltage fails to reach the reference voltage. This limits its temperature range to -10°C to 90°C. It is compact, occupying 0.017 mm^2 in a 28nm process, but it dissipates a relatively large power of 33.75 µW [7].

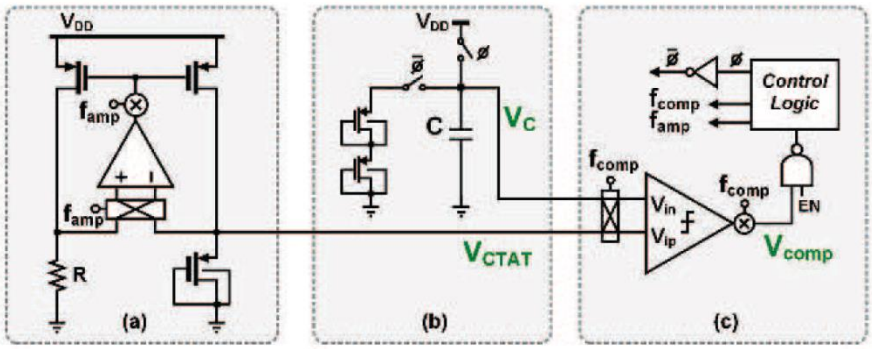


Figure 1.6 CBD-based temperature sensor with DTMOST as sensing element: a) VCTAT generator; b) DTMOST sensing code; c) relaxation oscillator [7].

A sensor that achieves better relative inaccuracy (0.67%) is depicted in *Figure 1.7*. In this design, the sampling capacitors of a delta-sigma modulator are also used to generate two temperature-dependent voltages, V_{D1} and V_{D2}, across two P+/Nwell diodes. A CB pair consists of two such circuits that differentially transfer the charge sampled on two C_S to the integrator. The second C_S can sample V_{D2} by discharging across an identical diode for a duration t_2 that sets the current ratio $p = \frac{t_2}{t_1}$. This way, ΔV_D is differentially transferred to the integrator. Alternatively, the second C_S can remain discharged, and the CBD pair in the front-end transfers only one V_D to the integrator.

The sensor operates from a 1V supply voltage, which is primarily limited by the amplifiers used in the modulator. Additionally, it covers the military temperature range: from -55°C to 125°C. With a power consumption of 2.2µW, it occupies a chip area of 0.021mm^2 [13].

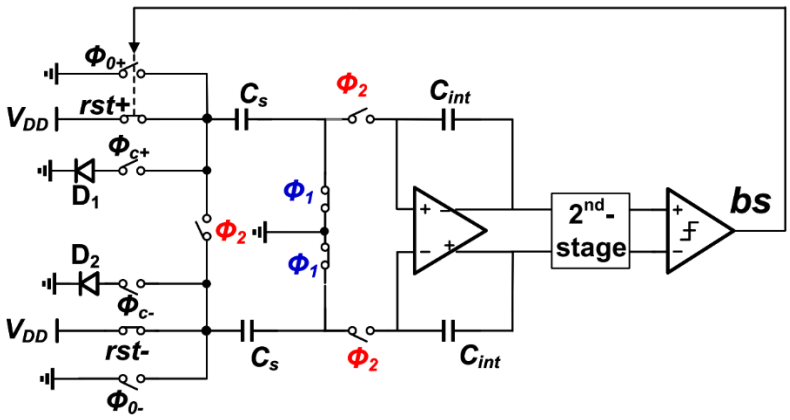


Figure 1.7 Block diagram of a CBD-based temperature sensor [13].

1.4.1 State-of-the-art CBD-based temperature sensor

A state-of-the-art CBD-based temperature sensor, which is an improved version of the design in [13], is shown in *Figure 1.8*. Based on diode-connected PNPs, it attains an impressive relative inaccuracy of 0.17%

after a one-point trim. The discharging currents through the PNPs are turned off by switching their bases, thus avoiding errors in the sampled V_{BE} due to the PVT-dependent voltage drop across a switch in series with the BJT. Furthermore, the modulator employs inverter-based pseudo-differential amplifiers, allowing the sensor to operate from a 0.95V supply voltage [1].

Figure 1.8 shows the detailed single-sided operation of the CBD front-end along with the first integrator. The sampling phase (Φ_1) can be divided into two parts: the precharge phase (rst) and the discharge phase (Φ_3). During the rst phase, the sampling capacitor C_S is precharged to V_{DD} (*Figure 1.8.1*). When Φ_3 is high (*Figure 1.8.2*), C_S is disconnected from the supply voltage and the PNP is turned on by connecting its base to the ground. C_S discharges through the diode-connected PNP for a duration of t, controlled by the Φ_1 and Φ_3 switches. The BJT is turned on through its base and not through other connections (e.g. the emitter side) to minimise the voltage drop across the switch. Additionally, switch Φ_3 opens shortly after Φ_1 to prevent charge injection from sampling onto C_S. The voltage V_{BE} is sampled on the capacitor C_S during this process (*Figure 1.8.3*). During the same phase (Φ_1), the integrator is autozeroed to mitigate its offset and 1/f noise and is biased in the correct operating region.

In Φ_2, the PNP is turned off in a supply-independent manner by connecting its base to a voltage $V_B \sim V_{BE}$. This V_B is generated by an auxiliary CB PNP. Finally, the front-end is connected to the integrator, facilitating the transfer of charge from C_S to C_{INT} for integration.

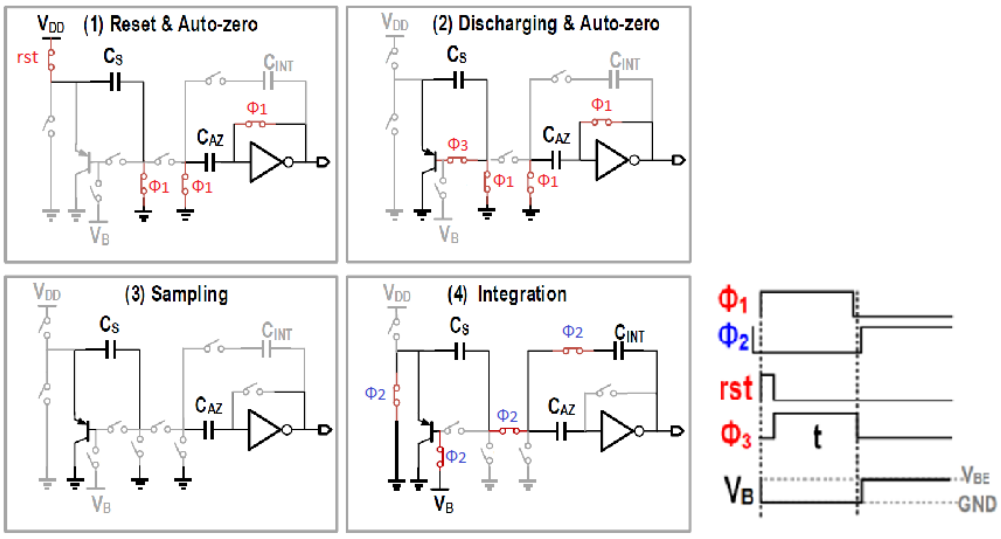


Figure 1.8 Single-sided switched-capacitor (SC) integrator with embedded CB PNP front-end [1].

The block-level architecture of the design proposed in [1] is depicted in *Figure 1.9*. This design incorporates three differential CB pairs to drive a second-order single-bit delta-sigma modulator (DSM). The clock signals of each CB pair are controlled by the bitstream (BS) output. In this way, a differential charge proportional to either V_{BE} or ΔV_{BE} can be transferred to the modulator, as follows:

BS = 0: $e = k\Delta V_{BE}$

BS = 1: $e = (k-1)\Delta V_{BE} - V_{BE2}$

Where k is the number of CB pairs in the front-end, which is 3 here. This gives a bitstream average, μ:

$$(1-\mu)k\Delta V_{BE} + \mu\big((k-1)\Delta V_{BE} - V_{BE2}\big) = 0 \tag{1.11}$$

$$\mu = \frac{k\Delta V_{BE}}{V_{BE1}} = \frac{3\Delta V_{BE}}{V_{BE1}} \tag{1.12}$$

The resulting μ makes good use of the modulator's dynamic range, as shown in *Figure 1.10*. The decimated output can be linearised as follows:

$$\mu_{lin} = \frac{\alpha\mu}{\alpha\mu + 1} = \frac{3\alpha\Delta V_{BE}}{3\alpha\Delta V_{BE} + V_{BE1}} \tag{1.13}$$

which can be further digitally processed and converted into temperature:

$$T = A\mu_{lin} + B \tag{1.14}$$

where α, A and B are fitting coefficients.

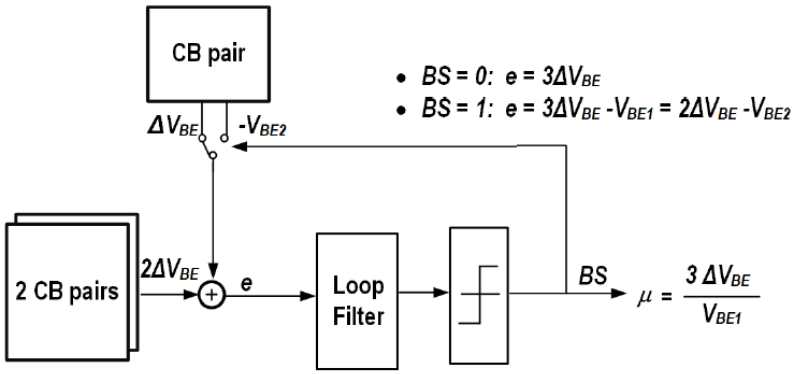


Figure 1.9 Block level architecture of the CBD sensor proposed in [1].

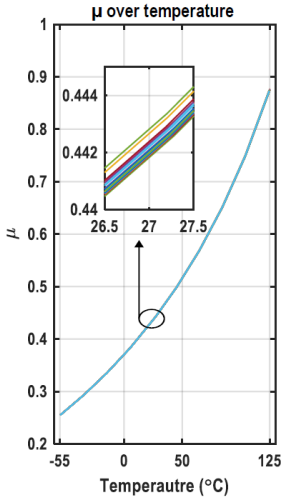


Figure 1.10 Measured bitstream average $\frac{3\Delta V_{BE}}{V_{BE1}}$ over temperature of chip [1].

As an experiment, a version of the sensor with smaller sampling capacitors C_S = 800fF was taped out to investigate whether the design could be scaled down further. However, measurements on 10 chips showed that this version only achieves an inaccuracy of ±0.31°C (3σ), after a PTAT trim at room temperature.

The measured bitstream average $\mu = \frac{3\Delta V_{BE}}{V_{BE}}$ for two different sampling time ratios (t_1/t_2 = 32 and 64) is shown in *Figure 1.11*. It can be seen that for large time ratios, the ADC clips at high temperatures. This is because from Equation *(1.6)*, reducing C reduces the bias current, which, in turn, reduces V_{BE}, while ΔV_{BE} increases with larger time ratios. To avoid this, the charge-balancing scheme must be modified, as will be discussed in the following chapter. However, these results show that the area and power dissipation of this design can be further optimised, while maintaining good accuracy over the military temperature range.

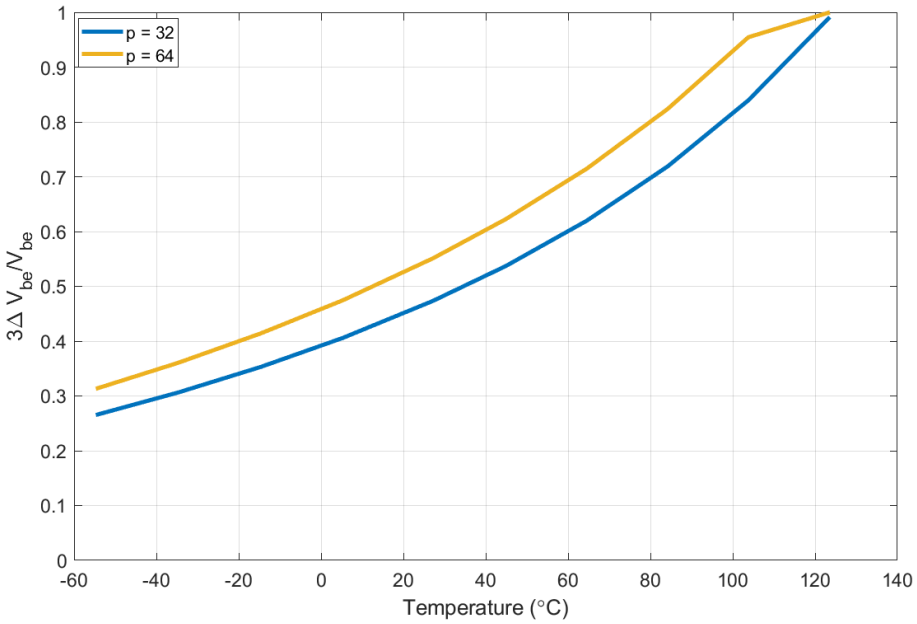


Figure 1.11 Measured bitstream average $\mu = \frac{3\Delta V_{BE}}{V_{BE}}$ across temperature for a time ratio p=32 and p=64.

1.5 Project goals

The performance of state-of-the-art CBD-based temperature sensors with low supply voltages and low power characteristics are summarised in *Table 1.1*.

Table 1.1 Comparison of the state-of-the-art low-supply temperature sensors.

	2019 [2]	2020 [7]	2021 [13]	2023 [1]	Design goals
Technology	16nm FinFET	28nm	55nm	180nm	180nm
Type	Bulk diode SAR	DTMOST OSC	P+/Nwell DT ΣΔ	PNP DT ΣΔ	PNP DT ΣΔ
Area [mm^2]	0.0025	0.017	0.021	0.25	< 0.125
Supply [V]	0.85 – 1	0.85 – 1.15	1 – 1.2	0.95 – 1.4	< 0.9
T. Range [°C]	-15 – 105	-10 – 90	-55 – 125	-55 – 125	-55 – 125
3σ Error [°C] (Trim point)	+1.5/-2.0 (0)	±2.0 (0) ±0.9 (1)	±1.4 (0) ±0.6** (1)	±0.45 (0) ±0.15 (1)	< ±0.20 (1)
R.I.A. [%] (Trim point)	2.9 (0)	4 (0) 1.8 (1)	1.6 (0) 0.67 (1)	0.5 (0) 0.17 (1)	< 0.22 (1)
Power [µW]	18	33.75	2.2	0.81	< 0.30
Tconv [ms]	0.013	0.1	6.4	128	128
Res. [mK]	300	10.2	15	1.8	< 20
PSS [°C/V]	1.5	0.27	3.7	0.2	~0.2
Res. FoM [pJK^2]	21	0.36*	3.1	0.34	< 0.34

* Needs an additional frequency-to-digital converter ** With systematic error correction

Note that the resolution FoM, a metric of their energy efficiency, is calculated as:

$$Res.FoM = \frac{Energy}{Conversion} \times Resolution^2 \tag{1.15}$$

From this table, it is evident that the design in [1] achieves the best relative inaccuracy of 0.17%, after a one-point trim. However, despite its low power dissipation (810nW) and excellent performance in other respects, it occupies more area (0.25 mm^2 in 180nm CMOS) than the other designs. As we will see in the following chapter, most of this area is occupied by capacitors. By reducing their size, both the area and the power consumption of the sensor could be further minimised.

With this objective in mind, the goal of this project is to design a sub-1V CBD-based temperature sensor in the same 180nm technology with less area (<0.125 mm2) and lower power dissipation (<300 nW). The sensor should also maintain a competitive inaccuracy of <0.2°C over the military temperature range, -55°C to 125°C, after a one-point trim.

1.6 Organisation of the thesis

The remaining chapters of the thesis are organised as follows.

Chapter 2 introduces the proposed architecture. System-level techniques are described to reduce capacitor size and obtain a compact, low-power CBD temperature sensor.

Chapter 3 explores the circuit-level implementation of the sensor. Two subsections are discussed here: the implementation of the CBD front-end and the realisation of the delta-sigma modulator (DSM).

Chapter 4 presents the simulated results of the sensor design, encompassing accuracy, resolution, estimated area, and power consumption.

Lastly, Chapter 5 concludes the project with a discussion on future improvements.

2 System level design

To design a compact, low-power, yet high-accuracy sub-1V temperature sensor, this project builds upon the existing state-of-the-art CBD sensor described in [1]. This chapter elaborates on the chosen architecture and discusses the design choices made to improve the previous work.

2.1 Reducing the area of the CBD SC temperature sensor

Figure 2.1 shows the die micrograph of the existing CBD temperature sensor [1], while *Figure 2.2* shows how much area is occupied by each circuit block. As can be seen, the area is dominated by the CBD front-end and the autozero capacitors of the first integrator. The area of the digital part is dominated by the clock booster capacitors. In the following, different ways of reducing the size of these capacitors will be discussed. As will be seen, this also improves other aspects, such as power consumption, while maintaining the good accuracy and energy efficiency of the original design.

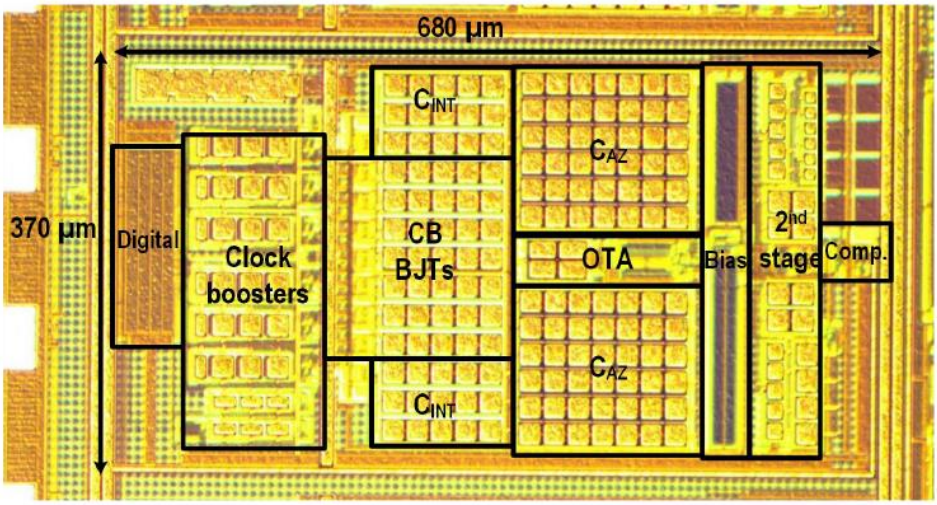


Figure 2.1 Die micrograph of chip in [1].

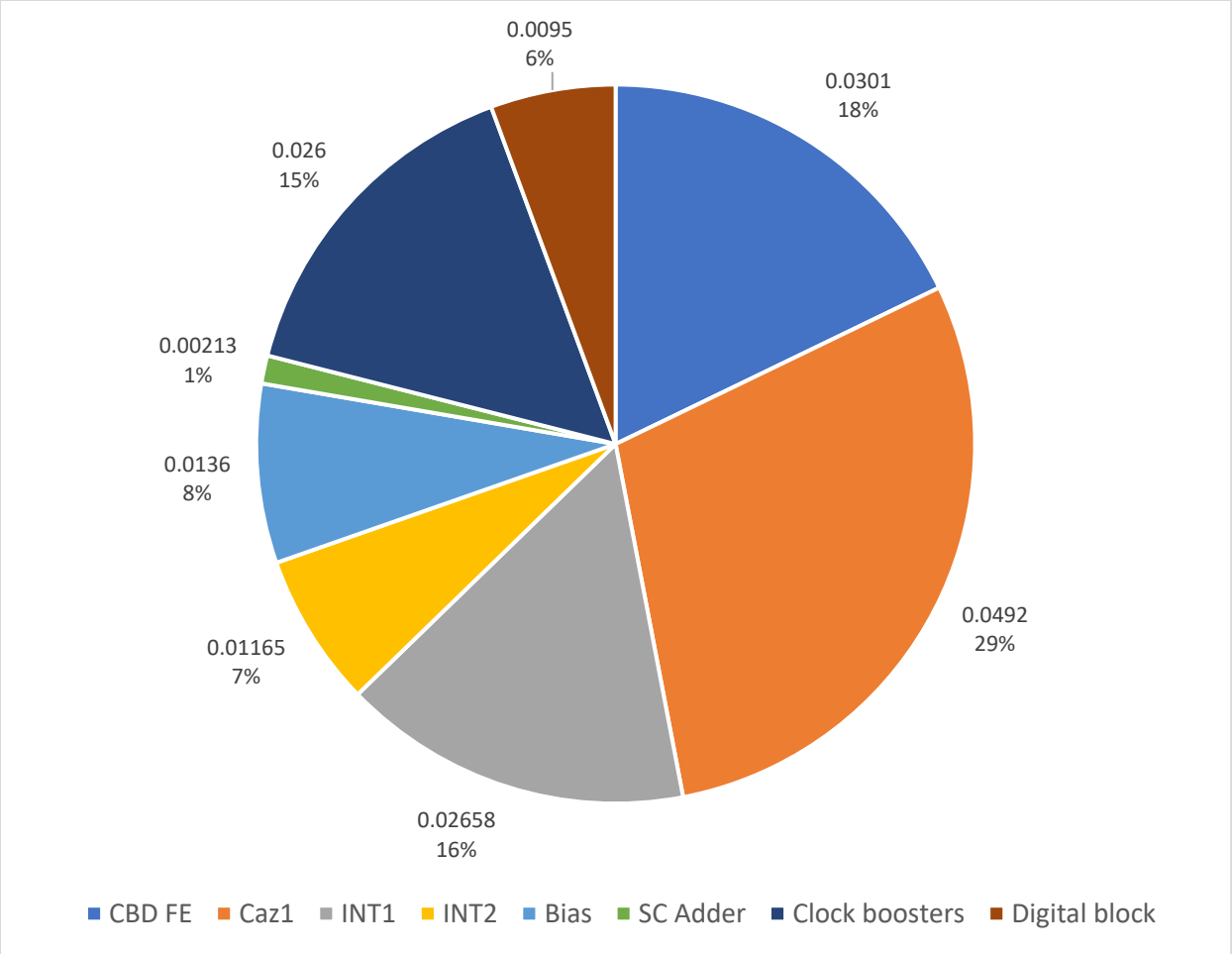


Figure 2.2 Area distribution of the chip in [1] in mm^2.

2.2 Scaling the front-end

The sampling capacitors (C_S) dominate the front-end area. Decreasing it, both the front-end area and dynamic power will be reduced, as given by:

$$P_{FE} = C_S \times f_S \times V_{DD}^2 \tag{2.1}$$

where f_s is the sampling frequency of the front-end (and DSM in this case).

By reducing their size, the current through the BJT and the sampled voltage are also affected. For a given discharge time, t, the BJT's emitter current as a function of C_S is given by:

$$I_E(t) \cong \frac{C_S V_T}{t} \tag{2.2}$$

Figure 2.3 shows this relationship for different values of C_S. It can be seen that I_E decreases as C_S decreases. For the design in [1], I_E is as low as 3.7nA at the end of the discharge. Reducing this further reduces front-end's accuracy due to the following reasons.

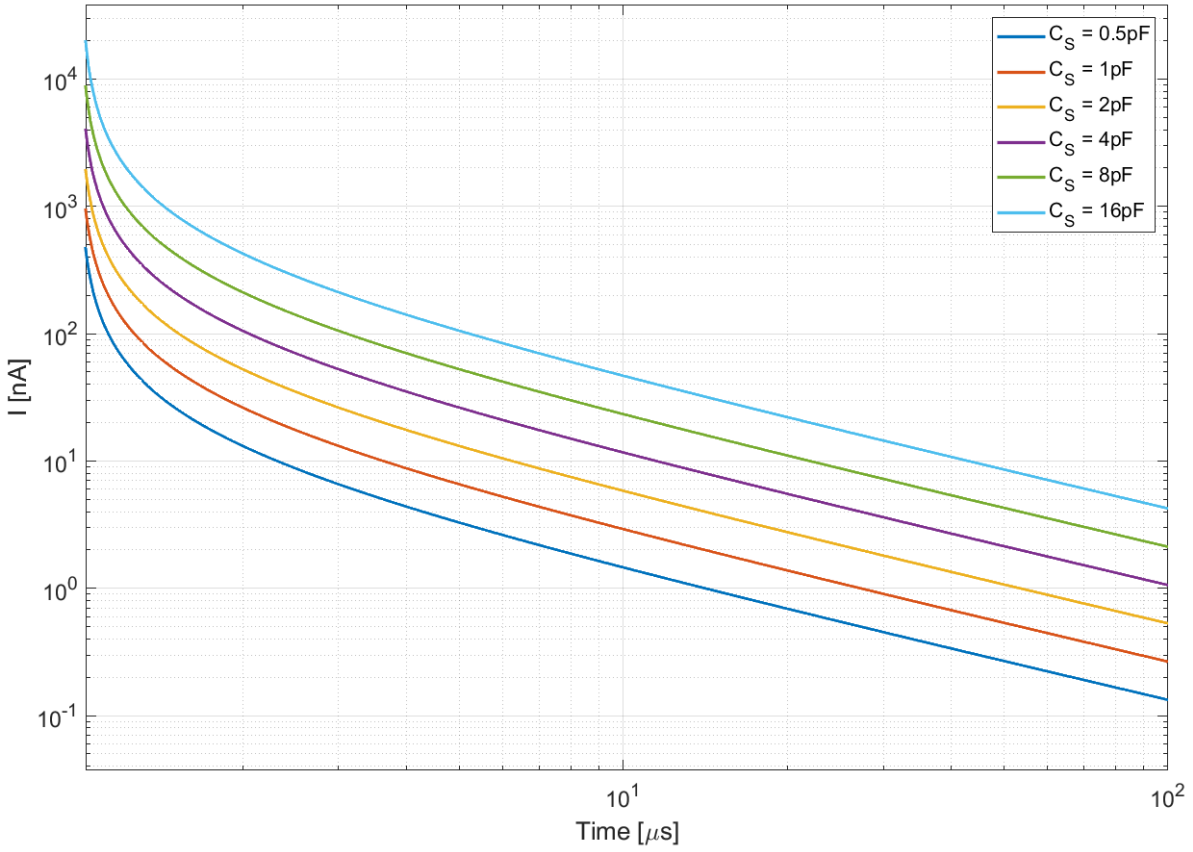


Figure 2.3 Current through a diode (ID) versus discharge time (t) for different CS sizes from 500fF to 16pF.

First, the BJT's current gain, β, decreases rapidly for small values of I_E, as shown in *Figure 2.4*. Since the PNPs are biased via their emitters, variations in β will affect their collector currents and thus, their ratios. The resulting the error in ΔV_{BE} is given by:

$$\Delta V_{BE,error} = \frac{kT}{q}\frac{\Delta\beta}{(\beta_1+1)(\beta_2+1)} \tag{2.3}$$

where β_1 and β_2 are the current gains of the two BJTs, and $\Delta\beta$ is their difference.

To reduce $\Delta\beta$, I_E should be chosen such that the BJTs are operated in a region where β is relatively constant. In *Figure 2.4*, we see that this occurs when I_E is greater than ~100 pA.

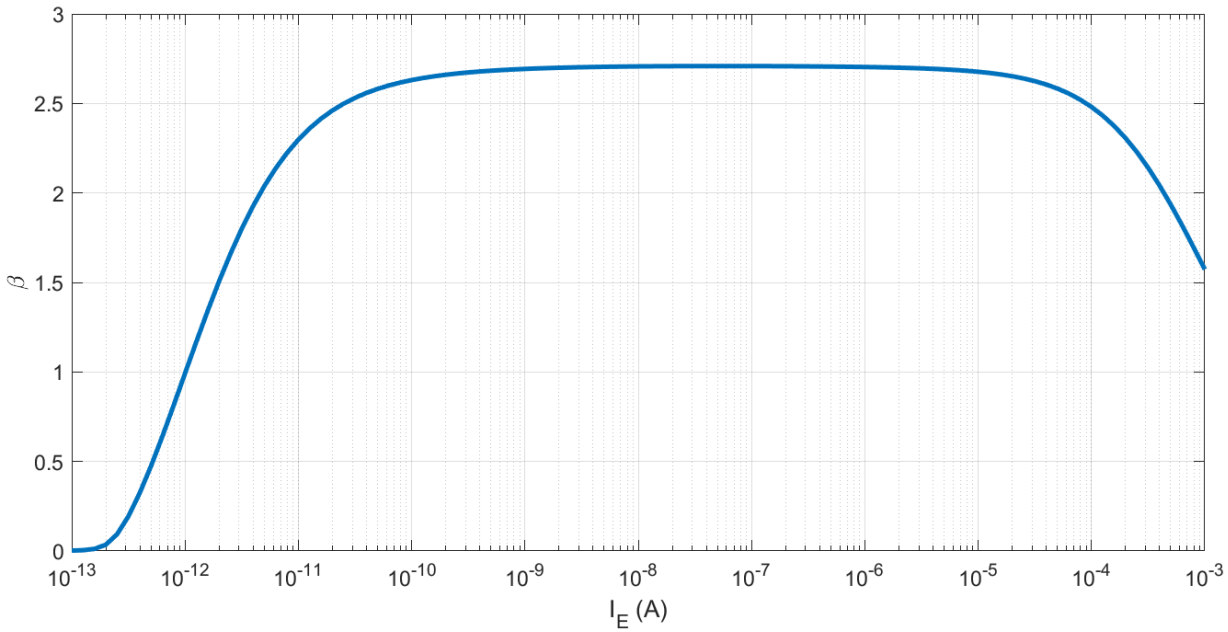


Figure 2.4 BJT's β variation with respect to emitter current IE.

Second, as the collector current reduces, the non-linear error due to the BJT's finite saturation current (I_S) increases, affecting the accuracy as follows:

$$\Delta V_{BE,error} = \frac{kT}{q}\left(\frac{pI_C + I_S}{I_C + I_S}\right) \tag{2.4}$$

To reduce the effect of I_S non-linearity, I_E should be chosen such that $I_C >> I_S$.

Considering these BJT non-idealities, an estimation of the minimum I_E and with that smallest C_S can be made, using the temperature sensitivity of the readout:

$$D_{out} = A\mu_{lin} + B \tag{2.5}$$

to ΔV_{BE} and V_{BE}, as well as the gain α, as derived in [4], where $\mu_{lin} = \frac{\alpha \Delta V_{BE}}{V_{BE} + \alpha \Delta V_{BE}}$.

$$S_{V_{BE}}^{D_{out}}(T) \simeq -\frac{T}{V_{REF}} \tag{2.6}$$

$$S_{\Delta V_{BE}}^{D_{out}}(T) \simeq \frac{A - T}{V_{REF}}\alpha \tag{2.7}$$

$$S_{\alpha}^{D_{out}}(T) \simeq \frac{T}{\alpha}\left(1 - \frac{T}{A}\right) \tag{2.8}$$

From these equations, it can be seen that the sensitivity of Dout to errors in ΔV_{BE} is the most significant. *Figure 2.5a* shows the temperature error in ΔV_{BE} for 25µm^2 PNPs biased at different I_E with a current ratio of p = 32, using the testbench shown in *Figure 2.5b*. This is the same as the current ratio used in [1], which achieved state-of-the-art accuracy. As I_E decreases, the temperature errors due to I_S and β increase, setting 400pA as a lower limit for achieving the target inaccuracy of 0.18°C over the military temperature range.

This leaves enough margin for additional sources of error, e.g. amplifier non-idealities, so that an inaccuracy of 0.20°C can be achieved for the whole design. As in [1], the discharge times are, t_1 = 32µs and t_2 = 1µs, which result in a minimum sampling capacitor of 490fF.

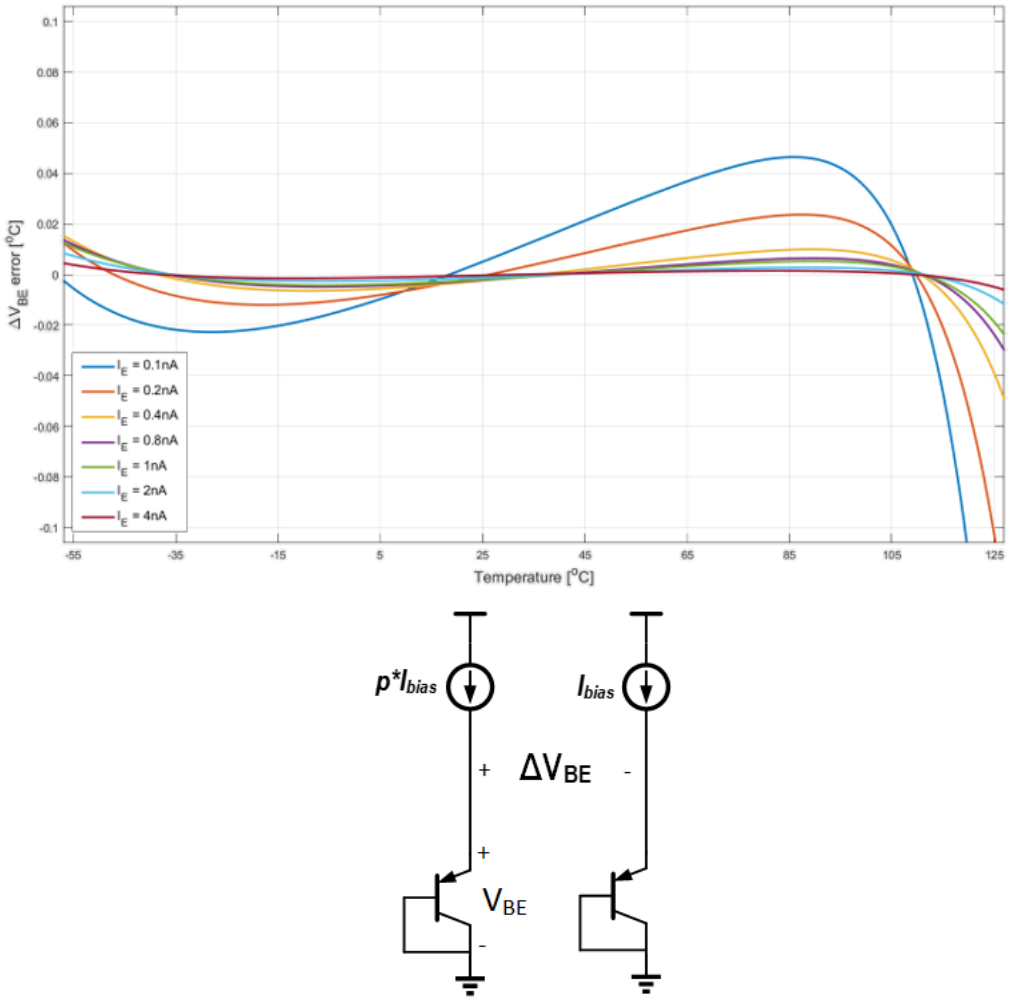


Figure 2.5 a) Temperature error in ΔVBE at different Ibias currents; b) Static biasing of two BJTs.

Besides the accuracy, another thing to consider when decreasing the size of C_S is the increase in the sampled noise (kT/C). For an inaccuracy of 0.18°C, it would be reasonable to design a sensor with ~10 times better resolution, meaning 18mK at most. In the original SC CBD-based front-end design in [1], C_S of 4pF results in a resolution of 1.7mK. This means that for a resolution limit of 18mK, C_S can be reduced by a factor of 10^2 to a minimum value of 40fF. This is much smaller than the limit on C_S set by the accuracy requirements, and so, resolution will not be a limiting factor.

In conclusion, C_S can be chosen as 490fF to reduce the front-end's area and power as well as the following stages of the modulator while maintaining good accuracy.

2.3 Charge balancing readout scheme

The last thing to be considered is the effect of reducing I_E on the charge balancing scheme. This decreases the sampled V_{BE}, which may cause the ADC to clip, as shown in Chapter 1. This can be prevented by scaling μ appropriately, as will be discussed in the following section.

For C_S = 490fF at 125°C, simulations show that:

$$\frac{\Delta V_{BE}}{V_{BE1}} \sim 0.35$$

so a gain of k=3 will make the ADC clip. To prevent this, k=2 was chosen in this design, as shown in *Figure 2.6*. The new charge balancing scheme with two front-ends is shown in *Figure 2.7*.

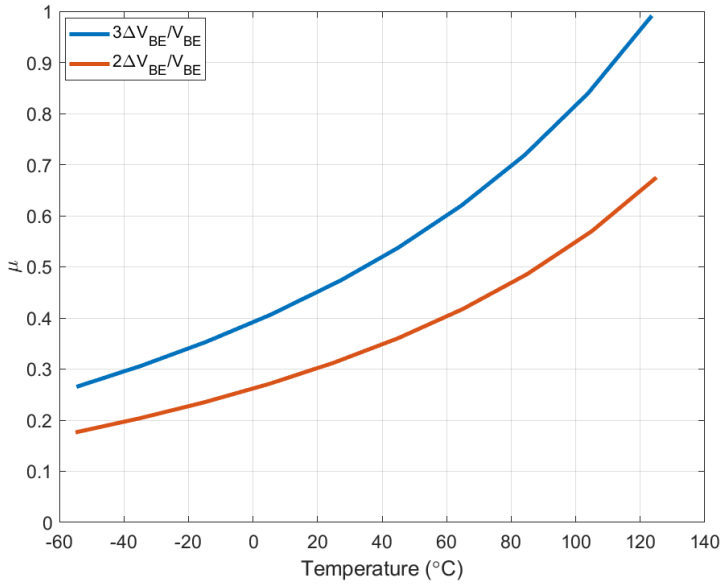


Figure 2.6 Simulated value of $\mu = \frac{3\Delta V_{BE}}{V_{BE}}$ and $\mu = \frac{2\Delta V_{BE}}{V_{BE}}$ across temperature, for p = 32.

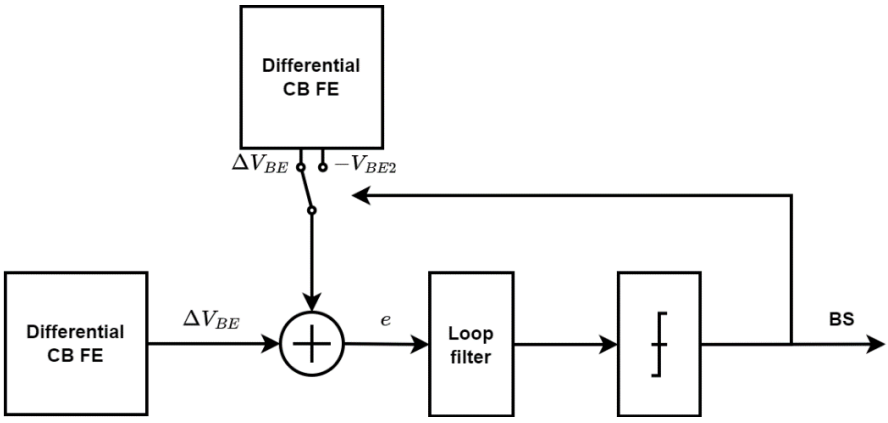


Figure 2.7 Proposed charge balancing scheme.

2.4 Scaling the autozero capacitors C_{AZ}

From the die micrograph shown in *Figure 2.1* and *Figure 2.2*, the second largest elements are the autozero capacitors, C_{AZ}, of the first integrator; therefore, they are the next candidate for area reduction.

Figure 2.8 shows the front-end and the first integrator of the modulator. An inverter-based amplifier is used because it can operate from a sub-1V supply. The capacitors C_{AZ} work as level shifters, enabling the amplifier's virtual ground (VirtP and VirtN) to be set at a 0V input common mode (V_{CM}). They also act as autozero capacitors, sampling the offset from the output during one phase and subtracting it during the integration phase.

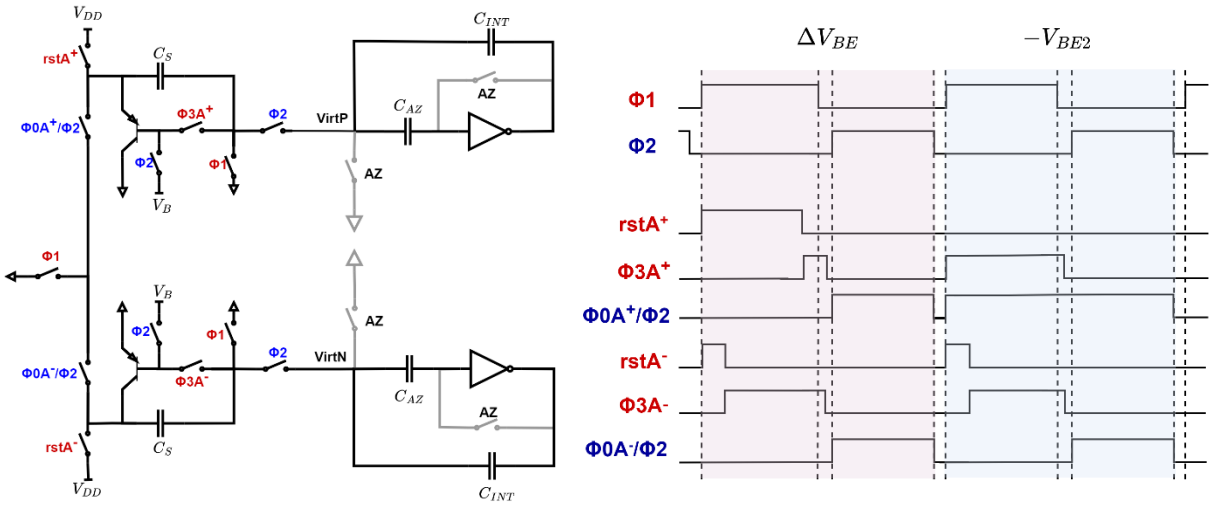


Figure 2.8 a) Front-end and the inverter-based integrator and b) their timing diagram.

The single-sided operation of the autozeroed inverter-based amplifier is shown in more detail in *Figure 2.9*. During the autozero (also biasing) phase (*Figure 2.9a*), the front-end is disconnected, and the PHAZ switches are closed. The OTA's bias current (originally 160nA) is set by the NMOS current mirror, and a voltage Vbn1 is stored on the $C_{AZ,N}$ capacitor. The bias current is generated using a constant-gm circuit. Similarly, the PMOS input transistor's biasing voltage is set on the $C_{AZ,P}$ capacitor. Since the amplifier is connected in unity gain, along with the bias voltage, the offset voltage is also stored on C_{AZP}.

During the integration phase (*Figure 2.9b*), the PHAZ switches are open: the biasing circuit is disconnected, and the front-end and integration capacitor C_{INT} are connected for integration, while the autozero capacitors are connected in series with the input to remove offset and provide the bias voltage to the input pairs. The cascode transistors ensure a large gain >80dB across PVT. Here, it can be seen that the minimum supply voltage for the OTA is $V_{gs} + 2V_{dsat}$, thus permitting sub-1V operation.

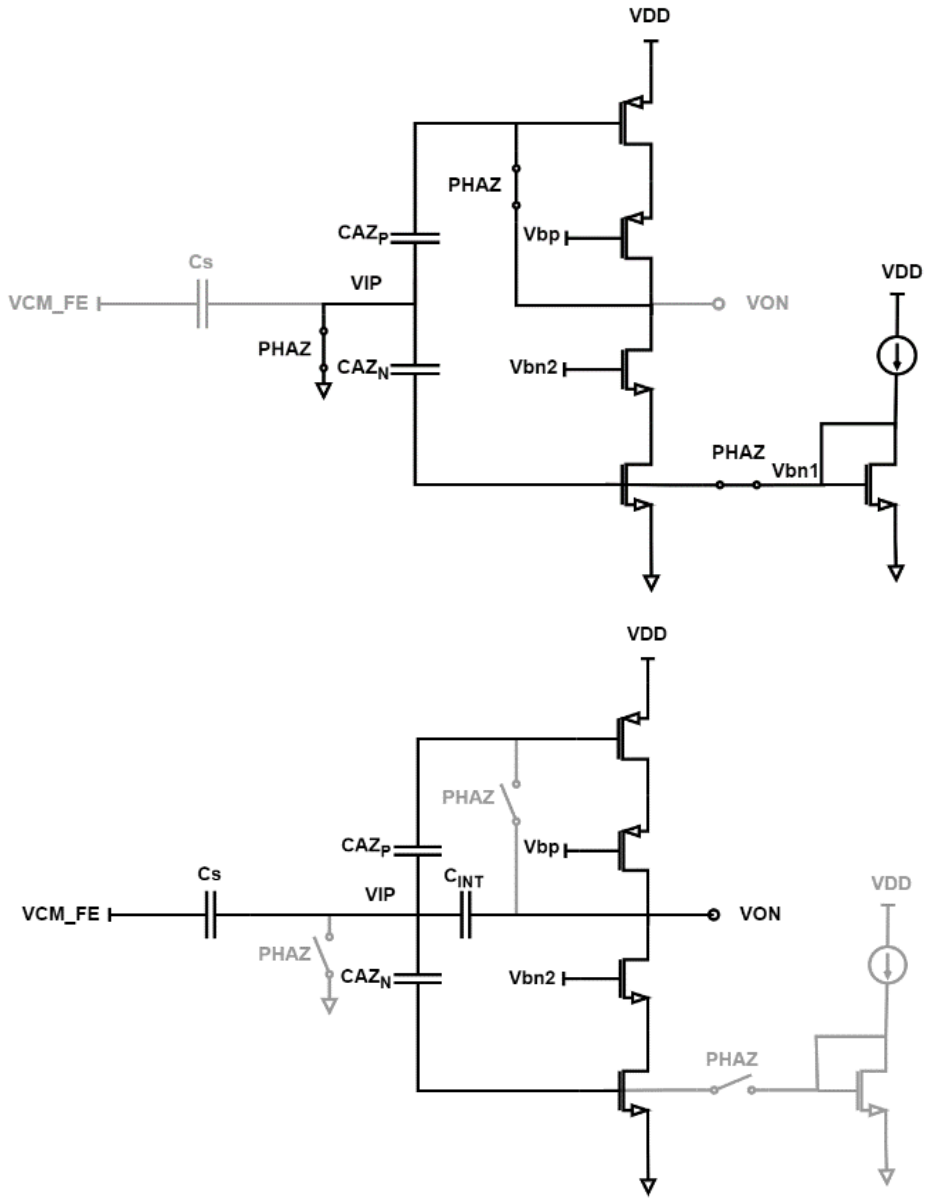


Figure 2.9 Single-sided inverter-based amplifier in a) autozero phase; b) integration phase.

The size of the autozero capacitors impacts the amplifier's noise and bias stability as will be discussed in the following.

Due to the sampling nature of the autozeroing, some of the amplifier's noise will be under-sampled and folded back to low frequencies (*Figure 2.10*). The magnitude of the folded-back noise is proportional to the bandwidth of the amplifier during the autozero phase (f_{BW}) and inversely proportional to the autozeroing frequency (f_{AZ}), and the noise amplification factor at low frequencies is given by [14]:

$$N = \frac{2f_{BW}}{f_{AZ}} \tag{2.9}$$

Furthermore, f_{BW} directly depends on the size of the autozero capacitor as follows:

$$f_{BW} = \frac{1}{2\pi}\frac{Gm}{C_{AZ}} \tag{2.10}$$

This means that the smaller the C_{AZ}, the more noise falls in the band due to autozeroing. To ensure that the front-end sensor dominates the noise, the autozero capacitors in [1] were made 4x larger than the sampling capacitors.

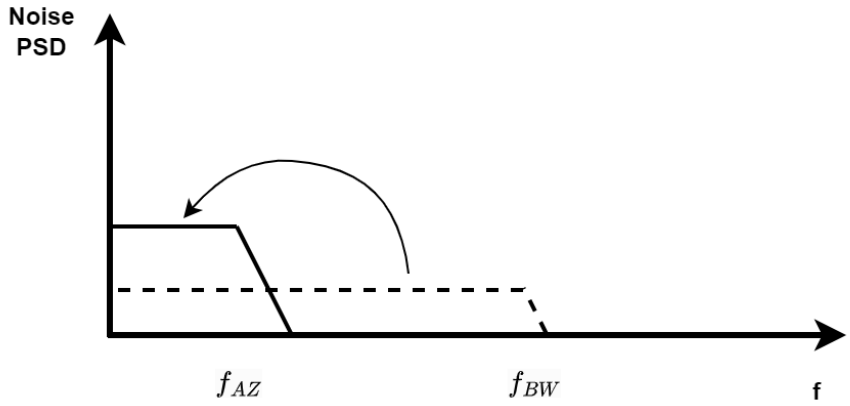


Figure 2.10 Noise folding due to autozeroing.

The size of C_{AZ} also determines how long it can hold the bias voltage in the presence of the leakage currents flowing through the open switches at the gates of the input transistors. The smaller the capacitors, the shorter the duration. Two types of drift are possible: differential drift and common mode drift. Differential drift unequally changes the voltage at the inputs of the differential amplifier's positive and negative sides and causes offset. This is due to the mismatch between the differential half circuits. Common mode drift causes the bias gate voltages of the input pair to change in the same direction, taking the input transistors out of their desired operating region and potentially turning them off. This problem can be resolved either by choosing a large enough C_{AZ}, optimising the autozero switches for low leakage or reducing the duration for which C_{AZ} has to hold its charge.

Lastly, the PHAZ switches will inject charge onto the C_{AZ} capacitors when they open. The smaller the C_{AZ}, the larger the resulting voltage error. Similar to the leakage of the bias voltage from C_{AZ} described above, charge injection can lead to a differential drift, resulting in an offset, or a common mode drift. Since charge injection is dependent on the size of the switches, as well as the size of parasitic capacitors at the gate of the amplifier's input pair, C_{AZ} scaling is a matter of circuit implementation and will be discussed in Chapter 3. However, the issues of noise down-sampling and the offset due to differential drift and charge injection can be resolved by chopping, as will be explained below.

2.3.1 Amplifier choppers

To enable the use of smaller autozero capacitors, the main amplifier can be chopped, as shown in *Figure 2.11*.

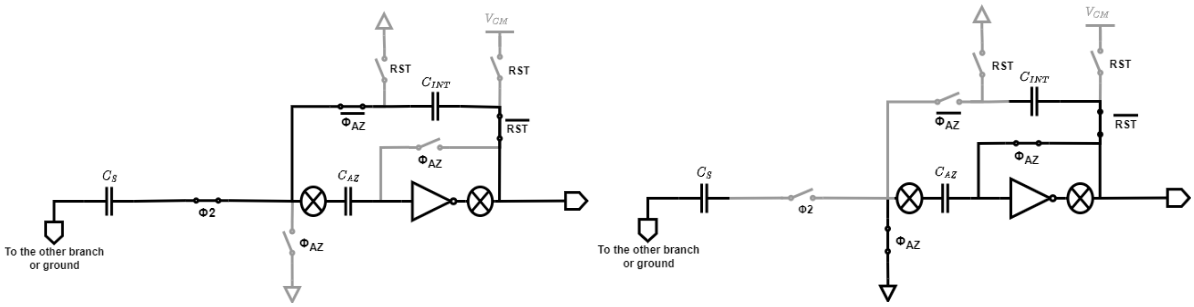


Figure 2.11 Chopper added to the first integrator: a) autozero phase; b) integration phase.

The choppers will upmodulate the offset and noise sampled on C_{AZ} to the chopping frequency $f_{CH,}$ as shown in *Figure 2.12*. The chopping frequency should be at least as high as the autozero frequency. If the autozero frequency is larger than f_{CH}, then as shown in *Figure 2.12b*, some noise will fall in band (around DC). By choosing the chopping and autozeroing frequencies to be at least the same, *Figure 2.12c* shows how the sampled noise on C_{AZ} is upmodulated out of band.

Figure 2.13 shows in the time domain how the choppers eliminate the noise and offset sampled on C_{AZ} for a frequency $f_{CH} = f_{AZ}$. Here, during one integration cycle, choppers are in positive polarity, and the noise sampled on C_{AZ} is positive. During the next integration cycle, the choppers change their polarity, and so does the noise on C_{AZ}. On average, the noise cancels out across one chopping period (which is two integration cycles). In the next cycle, the amplifier is autozeroing again, and new kT/C noise is sampled on C_{AZ}, which is once again cancelled out at the end of one whole chopping period.

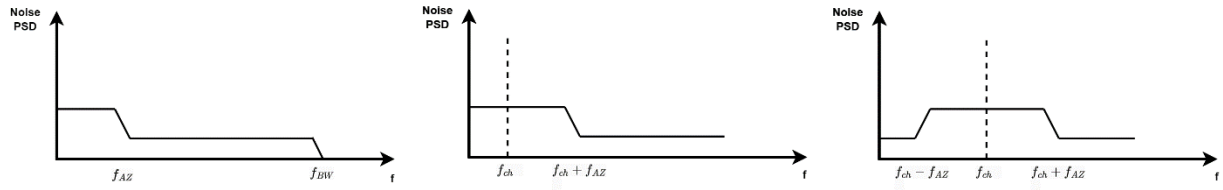


Figure 2.12 Noise power spectral density (PSD) versus frequency with a) autozero; b) autozero and chopping for fCH < fAZ ; c) autozero and chopping for fCH ≥ fAZ.

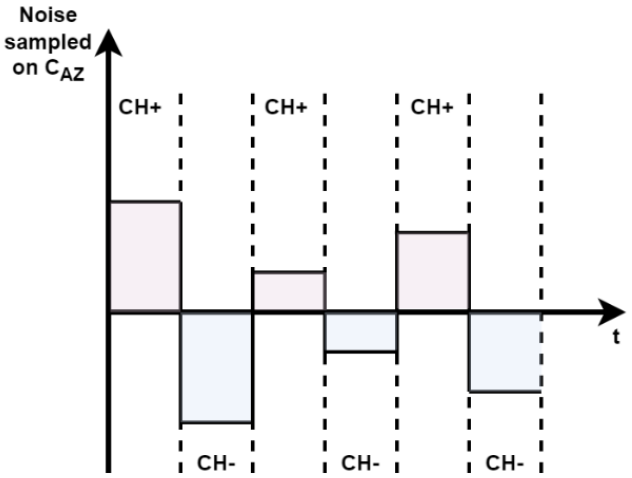


Figure 2.13 Chopping sampled noise on CAZ in the time domain.

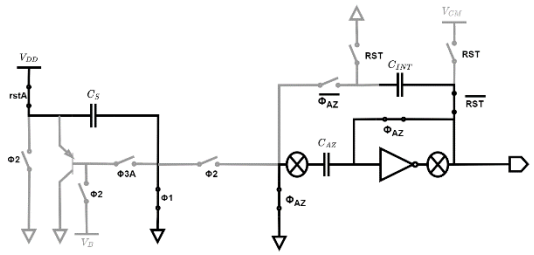
One additional benefit of the choppers is that they can remove the amplifier's flicker noise. By reducing the sampling and integration capacitors, the amplifier's bandwidth increases, or for the same bandwidth requires less current. To reduce the inverter-based amplifier's current, while keeping the same gm/Id, the transconductance (gm) decreases as well. To reduce gm, the input pair's size decreases, increasing the flicker noise since it is proportional to $\frac{1}{WL}$, the inverse of the MOSFET area. The flicker noise corner sets the lower boundary on f_{CH} (and with that f_{AZ}).

The integrator operates once during each state of the choppers, setting the chopping frequency to be at most $f_S/2$. Consequently, the modulator's behaviour can be depicted in *Figure 2.14*, as explained below.

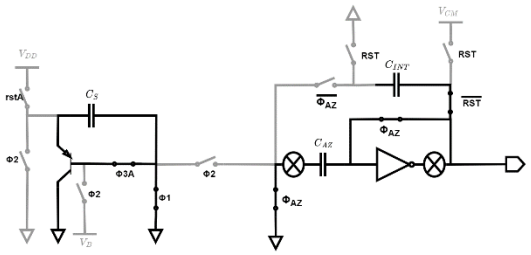
Here we see that during Φ_1 (*Figure 2.14(1,2)*) the front-end is sampling the temperature-dependent voltage by precharging the sampling capacitor to supply voltage V_{DD} and then discharging it across the diode-connected PNP for a specific time. In the meantime, the integrator that is disconnected from the front-end and is autozeroing. Its offset and biasing voltage are stored on the C_{AZ} capacitor. Then, during Φ_2 (*Figure 2.14(3)*) the sampling capacitor is connected to the integrator and the charge is transferred for integration.

In the next Φ_1 (*Figure 2.14(4,5)*), the choppers change their polarity, and once again, the front-end samples the temperature-dependent voltage. This time, the integrator is not autozeroing. Instead, it is set in a "do nothing" mode where the charge on the integration capacitors is stored, and no events are happening with the amplifier. Then, once again, in Φ_2 (*Figure 2.14(6)*), the charge from C_S is integrated in an incremental mode. After this, the choppers' polarity is reversed to its original values, and the offset, flicker and thermal noise sampled on the C_{AZ} are chopped. Afterwards, the amplifier can be autozeroed again.

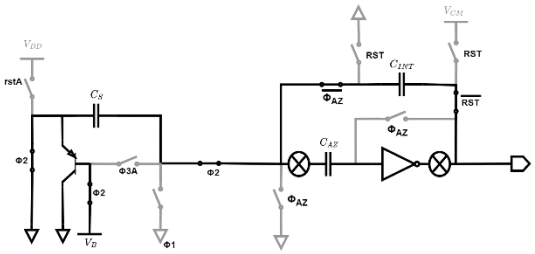
(1) Φ1: rstA, AZ



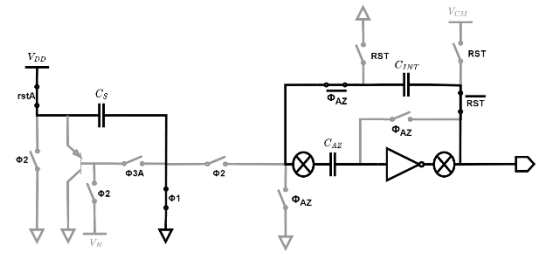
(2) Φ1: Φ3A, AZ



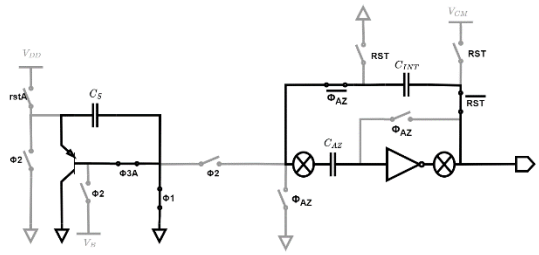
(3) Φ2: INT



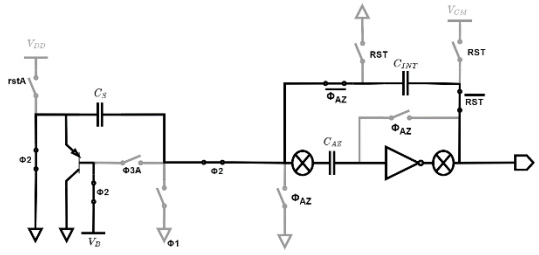
(4) Φ1: rstA, Do Nothing



(5) Φ1: Φ3A, Do Nothing



(6) Φ2: INT



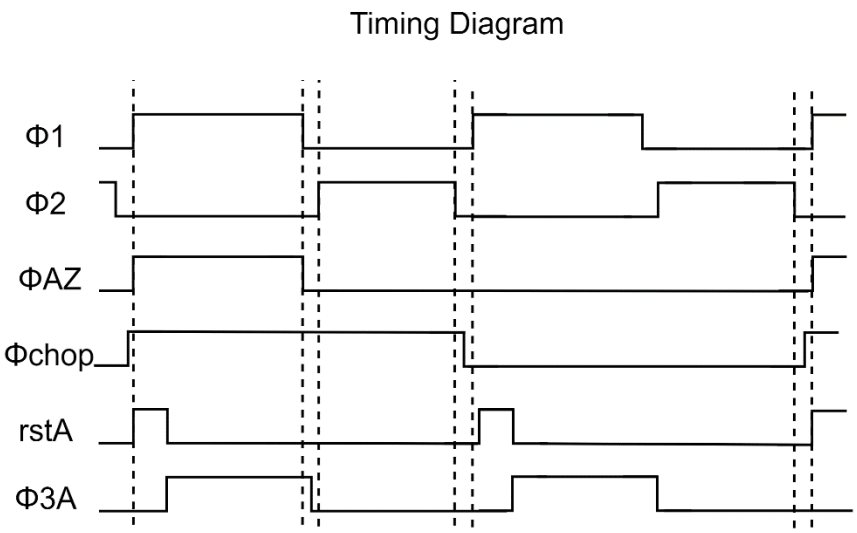


Figure 2.14 a) Integrator behaviour; b) Timing diagram.

2.5 Summary

This chapter discussed the system-level design of the proposed temperature sensor built upon the CB temperature sensor given in [1]. In addition to the advantages of the proposed changes, drawbacks were also discussed with proposed methods to mitigate them. These will be taken into consideration when designing the circuit in the following chapter.

3 Circuit implementation

In this chapter, the circuit implementation of two main blocks is discussed: the CBD front-end and the delta-sigma modulator readout. After this, the clock control logic is described.

3.1 CBD Front-end

Most of the front-end area is occupied by the six sampling capacitors; therefore, as discussed in Chapter 2, they should be made as small as possible (estimated C_S = 490fF) while maintaining good accuracy. However, the choice of BJT also affects the accuracy so this must be decided first.

3.1.1 Scaling down the PNP

The original design employed 5x5µm^2 PNPs [1]. Implementing the CB front-end shown in *Figure 2.8* with C_S = 490fF, a 5x5 µm^2 PNP and an ideal readout results in a 3σ inaccuracy of 0.15°C across the temperature range of -55°C to 125°C for a time ratio 32 (t_1=32µs, t_2=1µs), as seen in *Figure 3.1*. This is in excellent agreement with the measurement results, indicating that the BJTs in this process are well-modelled.

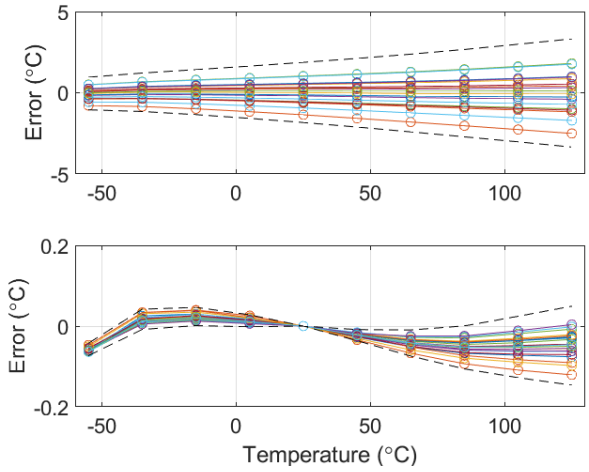


Figure 3.1 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with ideal switches and C_S = 490fF, VDD=0.9V and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom).

Using a smaller PNP (2x2µm^2) increases the emitter current density, and thus increases V_{BE}, as shown in *Figure 3.2.* Compared to the 5x5µm^2 device, the accuracy is quite similar (3σ = 0.16°*Figure 3.3a*), but the ADC range is reduced, resulting in a longer conversion time for equal resolution*Figure 3.4*).

A larger BJT, such as a 10x10 µm^2 PNP, would be expected to spread less, and thus achieve better accuracy at the cost of area. Additionally, V_{BE} will be smaller (*Figure 3.2*) resulting in larger ADC range and shorter conversion time (*Figure 3.4*). However, a larger BJT will have larger parasitic capacitors, C_p, between emitter and ground, which will decrease the initial voltage across C_S. If this drops below V_{BE}, which is highest at low temperatures, it will result in significant errors, as shown in *Figure 3.3b*. To avoid this, the supply voltage needs to be increased above 1V to compensate for the initial voltage drop.

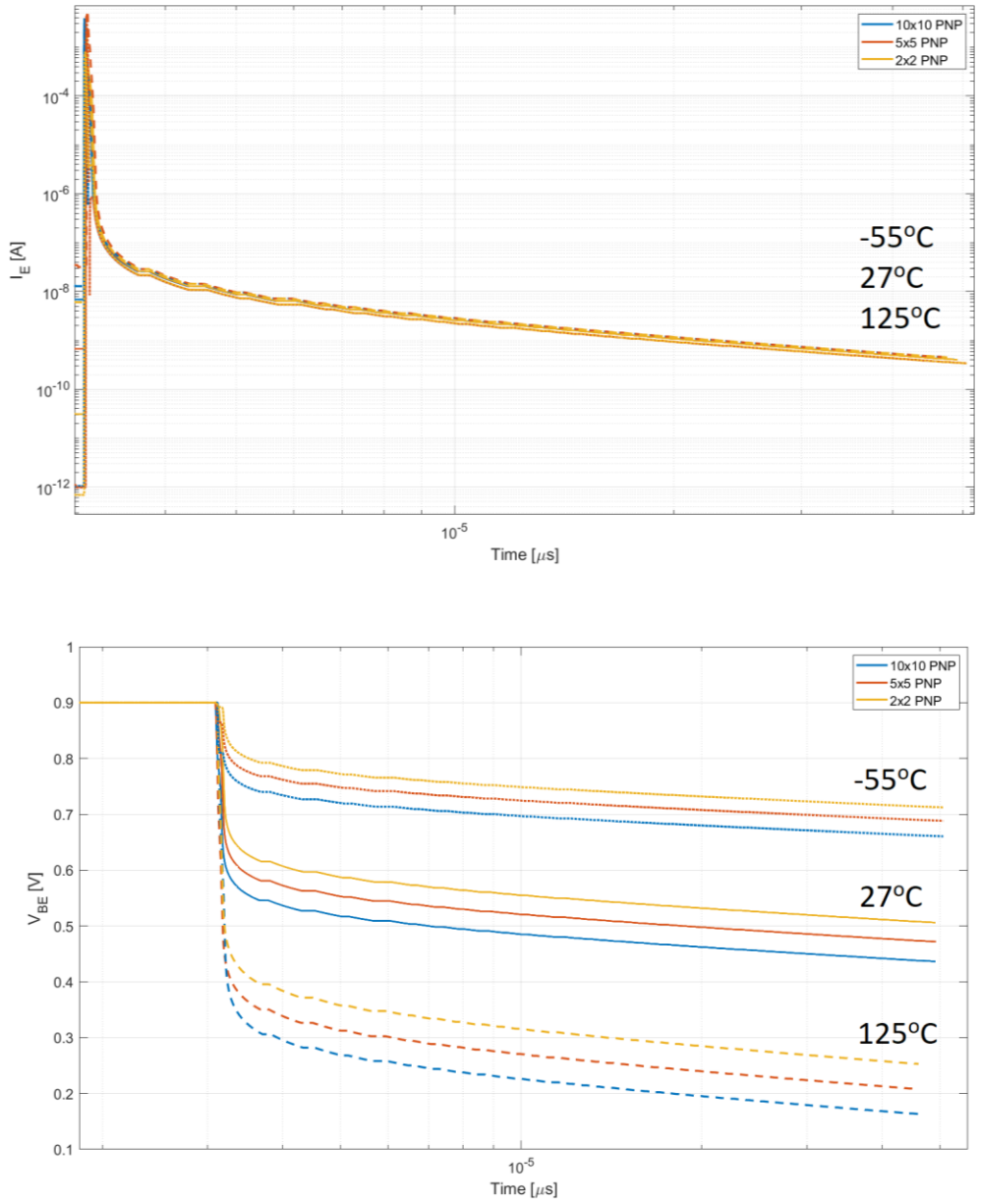


Figure 3.2 a) I_E and b) V_{BE} versus time for PNPs of different areas with C_S = 490fF and VDD=0.9V across temperature.

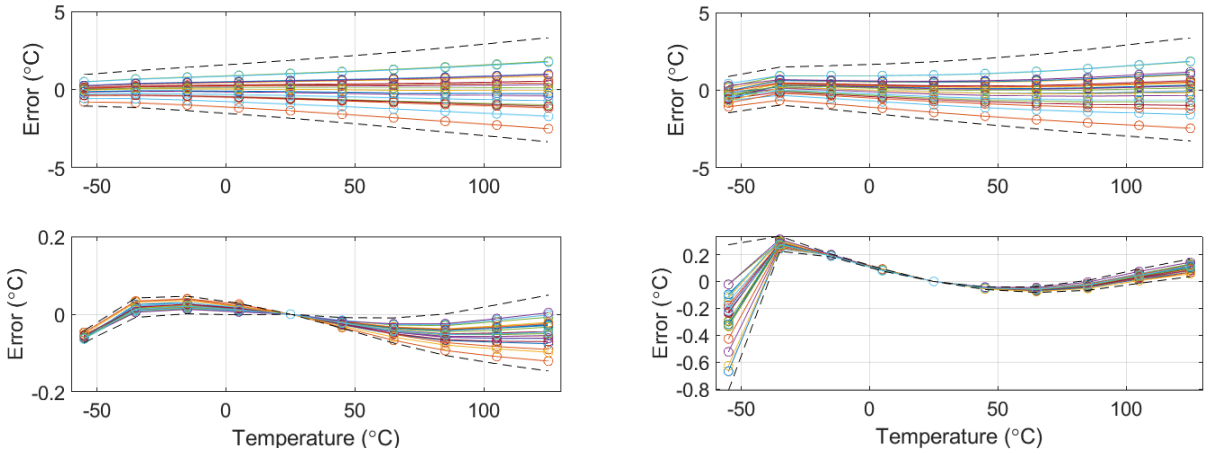


Figure 3.3 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with real switches and C_S = 490fF, VDD=0.9V and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom) for a) 2x2 PNP; b) 10x10 PNP.

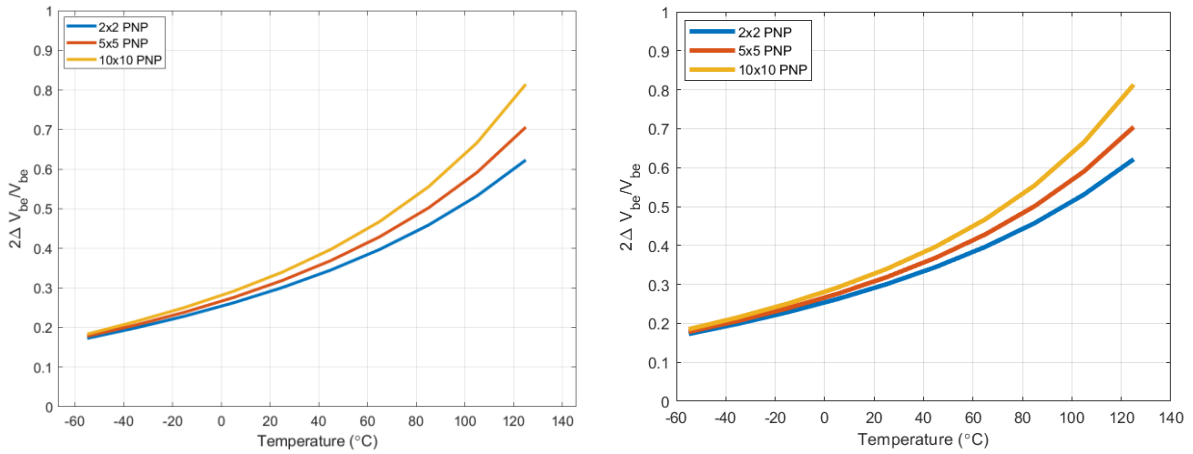


Figure 3.4 Decimated output over temperature for a 2x2, 5x5 and 10x10 PNP device with C_S = 490fF and p = 32 for a) VDD = 0.9V and b) VDD = 0.95V.

Alternatively, an NPN can be used to generate the temperature-dependent voltages V_{BE} and ΔV_{BE}, as shown in *Figure 3.5*. This was tried in an earlier tape-out, but as can be seen from the measurements shown in *Figure 3.6*, the resulting sensor exhibits significant non-linearity at high temperatures due to the leakage through the parasitic diode between the deep-n-well collector and the p-substrate. Considering that the area of this diode was minimized by using a custom device with a 1x1 mm^2 emitter area, it was decided not to use NPNs in this design.

To conclude, a 5x5μm^2 PNP will still be used as the sensing device in the SC front-end.

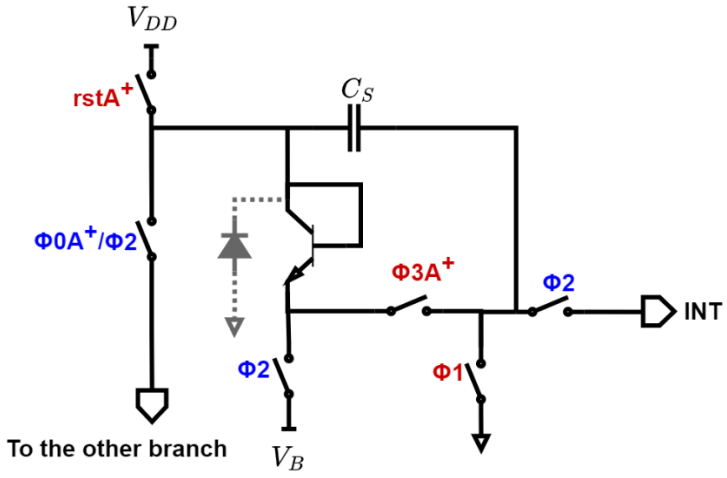


Figure 3.5 CBD Front-end with NPN device.

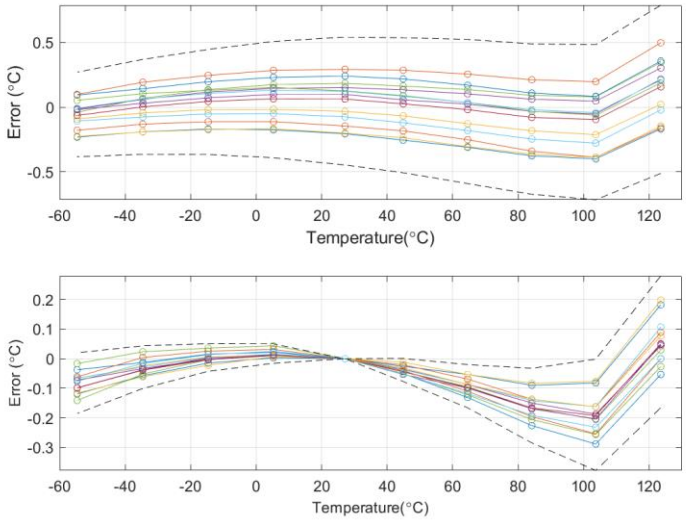


Figure 3.6 Measured inaccuracy of a 1x1 NPN-based CBD front end without (top) and with a 1-point PTAT trim at 25°C (bottom).

3.1.2 Scaling down the front-end sampling capacitor, C_S

Besides the BJT non-idealities discussed in Chapter 2, the minimum size of C_S is also limited by switch leakage. A version of the SC CBD-based temperature sensor given in [1] with 0.8pF C_S was taped out, resulting in an inaccuracy of 0.31°C (3σ) with VDD = 0.95V (*Figure 3.7*). Apart from the poor performance of two chips at low temperatures, the sensor's accuracy was limited by leakage at high temperatures, mainly through SW4 and SW5 (*Figure 3.8*), which will be optimized in this design.

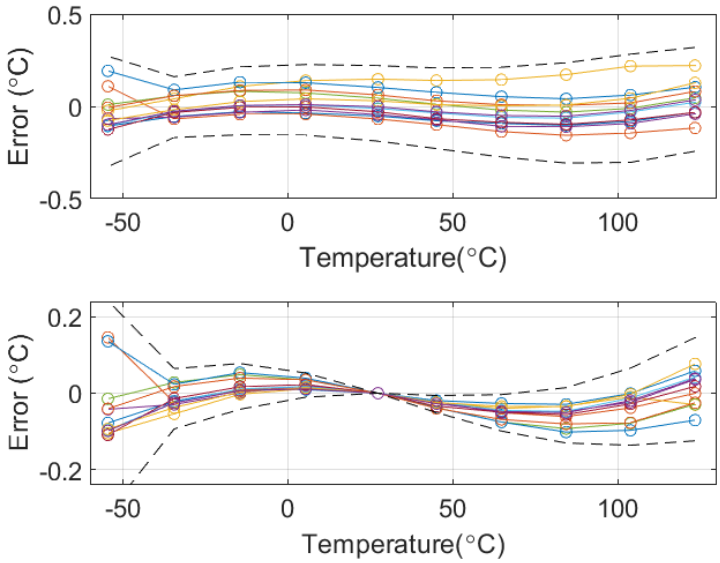


Figure 3.7 Measured inaccuracy of a 5x5 PNP-based CBD front end with Cs = 0.8pF without (top) and with a 1-point PTAT trim at 25oC (bottom).

There are two types of leakage: subthreshold leakage from the drain to the source of the switches and leakage through the source/substrate and drain/substrate PN junctions, which become exponentially worse at higher temperatures. To reduce their leakage, the switches were implemented with long NMOS transistors of size W/L = 220nm/700nm. Since V_{BE} is larger at lower temperatures, an additional PMOS switch of the same size is added in parallel to take over the switch's function at colder temperatures. Additionally, SW5 is implemented with a high-threshold voltage (HVT) NMOS device driven by a boosted $2V_{DD}$ clock signal. This way, over PVT, the largest leakage current is 1.12pA through SW4 and 4.5pA through SW5. They are much smaller than the minimum I_E of 880pA at 125°C.

The reset switch SW1 also shows high leakage and is implemented with a minimum-size PMOS T-switch, resulting in a leakage current of maximum 1.8pA across PVT.

The SW3 switch was implemented with an NMOS switch of size W/L = 4µm/350nm, with a boosted clock signal of $2V_{DD}$ for low on-resistance of 270Ω to minimise the voltage drop due to the base current. Additionally, SW3 is surrounded by half-sized dummy switches for low charge-injection.

Lastly, switches SW2 and SW6 are implemented with simple NMOS devices of sizes W/L = 500nm/180nm and W/L = 220nm/180nm, respectively. SW2 was made wider to reduce its on-resistance and the associated voltage drop across it when C_S is discharging through the BJT.

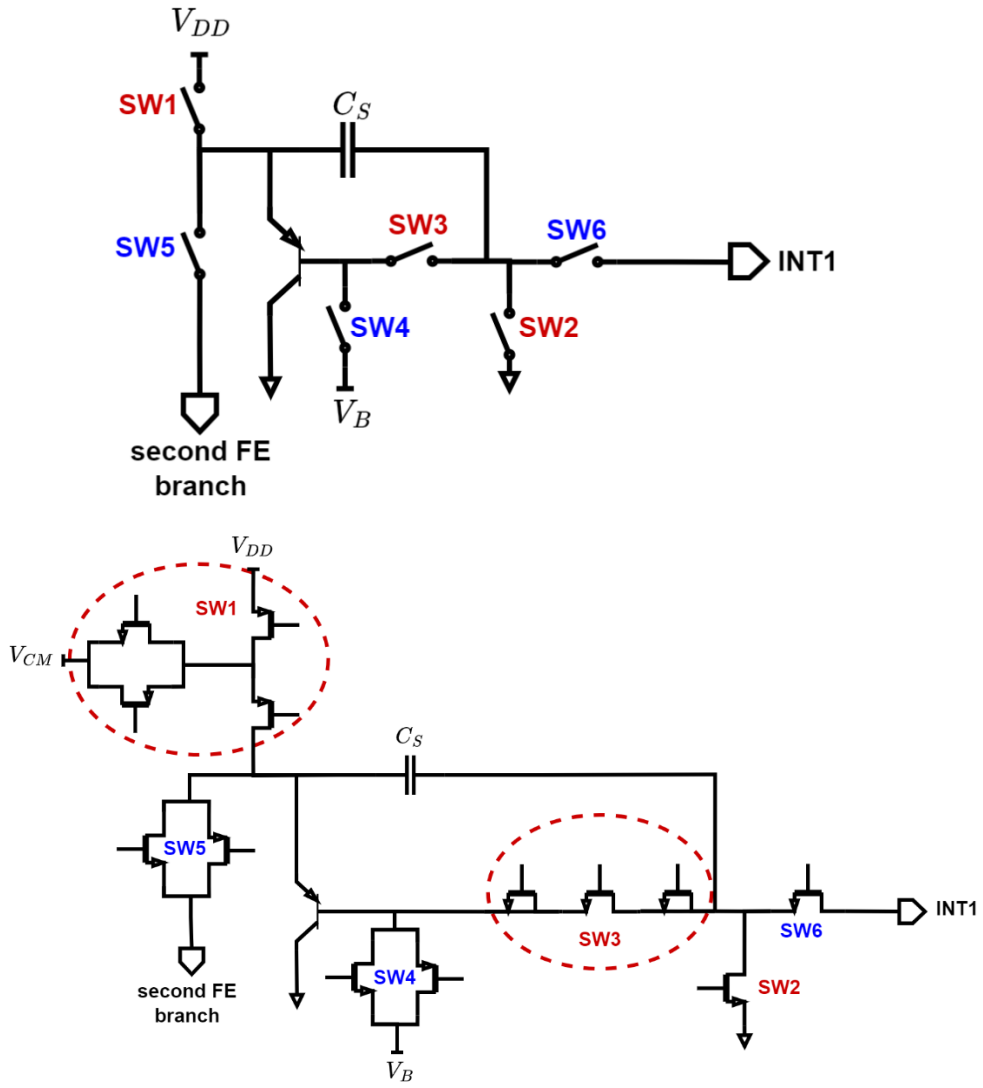


Figure 3.8 Single-sided front-end with switches a) representation; b) real implementation.

Despite optimising the switch leakage, simulations showed that the bias current had to be increased to 650pA to achieve the target inaccuracy. This choice results in C_S = 800fF, and a front-end inaccuracy of 0.18°C after a single-point PTAT trim at a supply voltage of 0.9V, (*Figure 3.9a)*. Similarly, the inaccuracy at a supply voltage of 0.95V is 0.19°C (*Figure 3.9b*).

For the case of shorter discharge times of t_1 = 16µs and t_2 = 0.5µs, the inaccuracy plots are shown in *Figure 3.10* and result in a 3σ inaccuracy of 0.17°C at a supply voltage of 0.9V.

As a result, the front-end area is approximately five times less than that in [1].

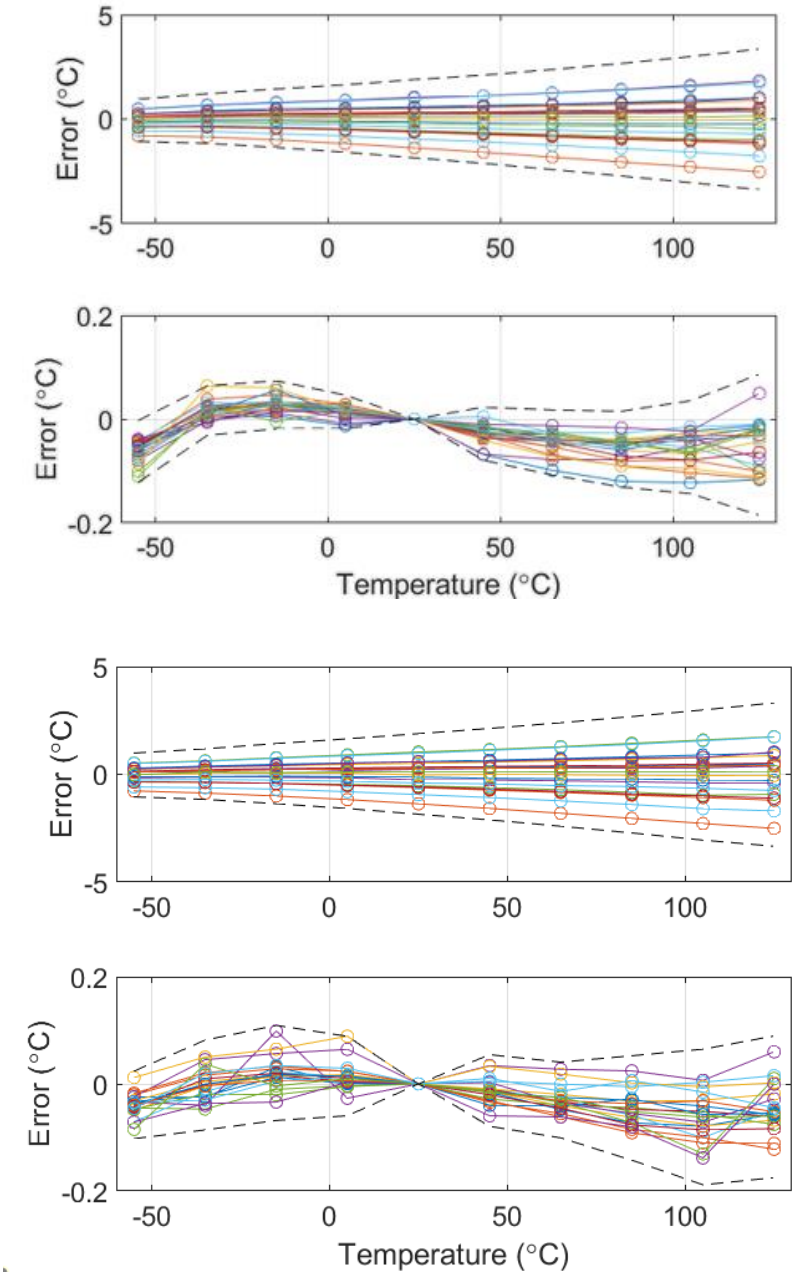


Figure 3.9 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with real switches and C_S = 800fF and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom) for a) VDD = 0.9V and b) VDD = 0.95V.

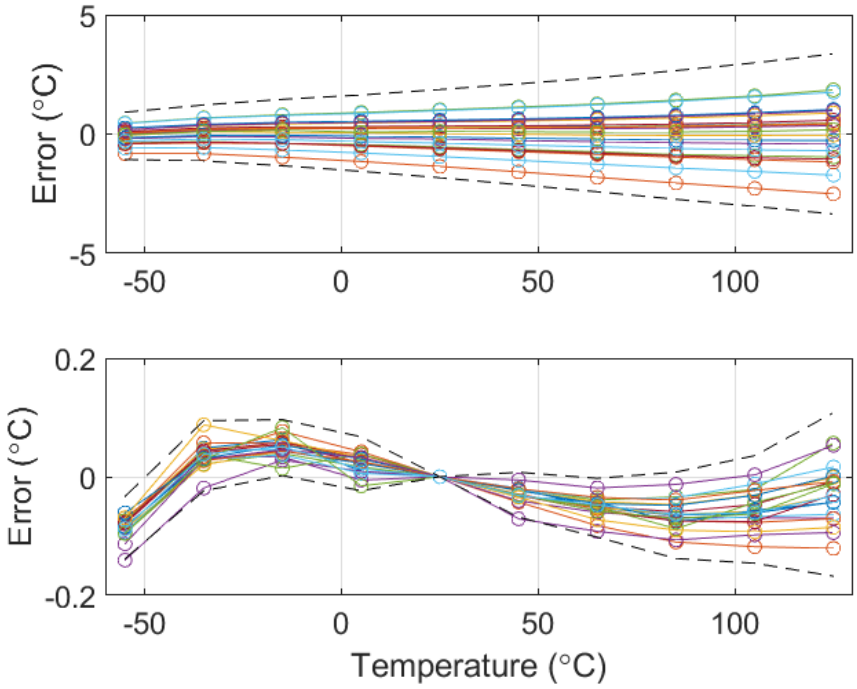


Figure 3.10 Monte Carlo (20 points, mismatch and process spread) SC CBD inaccuracy with real switches and CS = 800fF and p = 32, for halved discharge time t1 and t2 without (top) and with a 1-point PTAT trim at 25oC (bottom) for VDD = 0.9V.

3.1.3. Front-end biasing voltage generator

To turn off the PNP during Φ_2, its base is connected to a biasing voltage, V_B, (*Figure 2.8*). This voltage is generated by a simplified replica of the front-end, as shown in *Figure 3.11* [1]. The biasing circuit is shared by both CBD front-ends. Since V_B does not need to be accurate, the switches' requirements are relaxed.

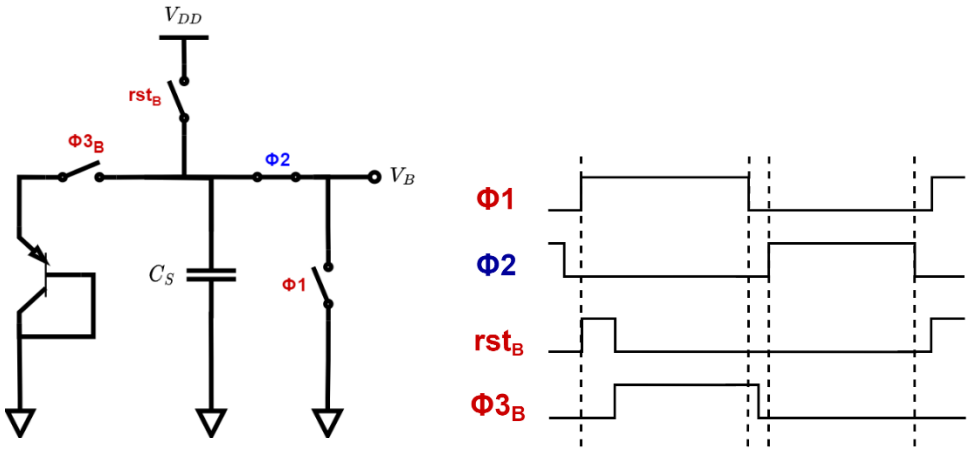


Figure 3.11 Front-end biasing circuit and its timing diagram.

3.2 Delta-sigma modulator readout

3.2.1 DSM topology

As in [1], a 2nd order delta-sigma ADC, as depicted in *Figure 3.12*, is used to balance the PTAT and CTAT voltages from the front-end and digitise their ratio. A feed-forward topology was employed to ensure a small swing at the output of the 1st integrator and, with that, better circuit linearity. The modulator's coefficients were scaled such that a swing of ±200mV is kept at the output of each integrator. This swing requirements are further discussed in the following sections. The coefficients are shown in *Table 3.1*.

With f_s = 15.625kHz, a target conversion period of 128ms, and an OSR of 1000, given the thermal noise is limited by C_S=800fF, the modulator provides an SNR = 96dB. This is equivalent to a resolution of 2.8mK, determined by the front-end capacitors, over a temperature range of -55°C to 125°C, as given by:

$$SNR = 20\log\left(\frac{Temp\ range}{resolution}\right) \tag{3.1}$$

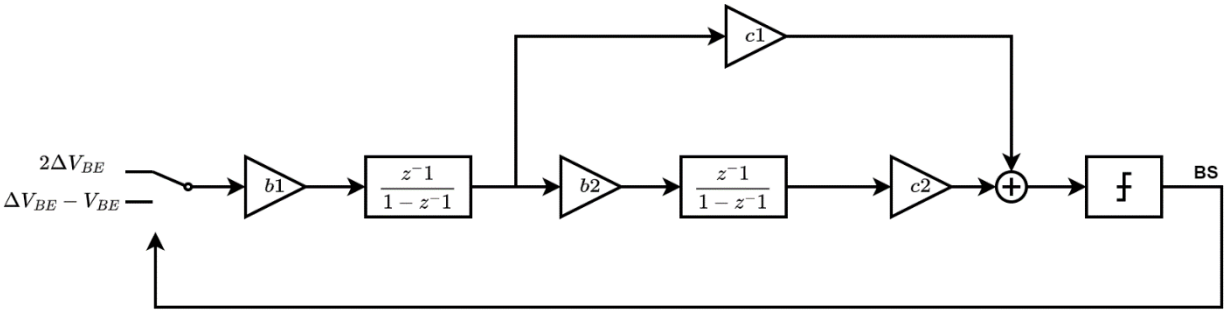


Figure 3.12 DSM Block architecture.

Table 3.1 Coefficients for the implemented DSM.

DSM Coefficient	Value
b_1	0.6
b_2	0.4
c_1	2
c_2	2.5

3.2.3 Amplifier design

As explained in Chapter 2, the first integrator is built around a pseudo-differential inverter-based amplifier. The single-sided operation of the amplifier during autozero and integration phase is shown in *Figure 3.14*.

The required amplifier's open-loop gain is determined by the required DSM open-loop gain. To achieve an SNR = 96dB and, with that, a resolution of ~2.8mK, the DSM open-loop gain (A_{DSM}) should be [15]:

$$A_{DSM} \sim 10 \times SNR = 116dB \tag{3.2}$$

Since it is a 2nd order modulator, the open-loop gain of one amplifier should be >58dB across temperature and corner. To suppress any amplifier non-linearities, the amplifier was designed for a higher gain of 80dB, with a minimum value of 65dB across PVT, as seen in *Figure 3.15a*.

The amplifier's gain-bandwidth (GBW) is set at 137.5kHz with a load capacitance of 1.9pF, constituting of the sampling capacitors of the first and second integrator, first integration capacitor, feed-forward capacitors and the common-mode feedback capacitors, as shown in *Figure 3.13*. This bandwidth ensures that the output settles within 5τ resulting in an inaccuracy error of less than 6mK (5τ), which is equivalent to 25µV settling error in V_{BE}.

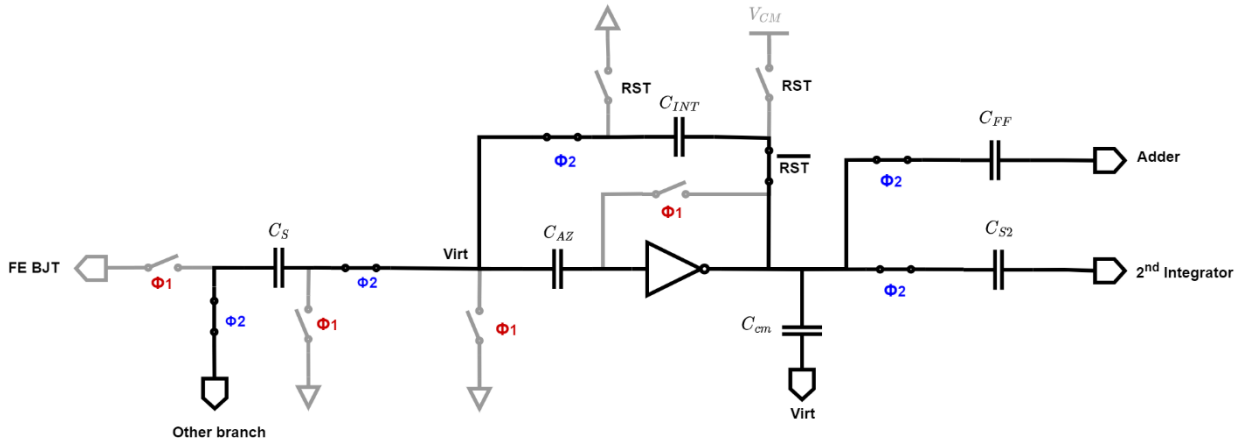


Figure 3.13 Switch-capacitor implementation of first integrator.

The transistors are set in saturation with a minimum saturation margin, V_{DSAT}, of 30mV across PVT. This allows for a sub-1V operation of the amplifier. The input pair is biased in weak inversion for high current efficiency with a gm/Id value of 20. Since C_S is 5x smaller than in [1], the amplifier is operating at a current of 32nA (per OTA branch), which is also 5x smaller than the current used in the previous design [1].

To reduce the inverter-based amplifier's current, for the same gm/Id = 20, the size of the input pair has to be scaled down as well, resulting in a smaller transconductance. This increases the $1/f$ and thermal noise of the amplifier. The thermal noise voltage will be √5x larger than the previous design, at 0.16µV/√Hz. Despite this, the kT/C noise of the front-end still dominates the noise floor. As for the flicker noise, with its corner set at 1.1kHz (*Figure 3.15b*), it can be removed by chopping the amplifier, as discussed in the following subsections.

The biasing voltages for the gates of the NMOS input transistor as well as the cascodes are provided by a replica biasing circuit shown in *Figure 3.16*. Here, the current of 6nA is copied to the main amplifier via a

current mirror. The same biasing is reused in the 2nd integrator. The current through the biasing circuit is provided by a constant-gm circuit [1].

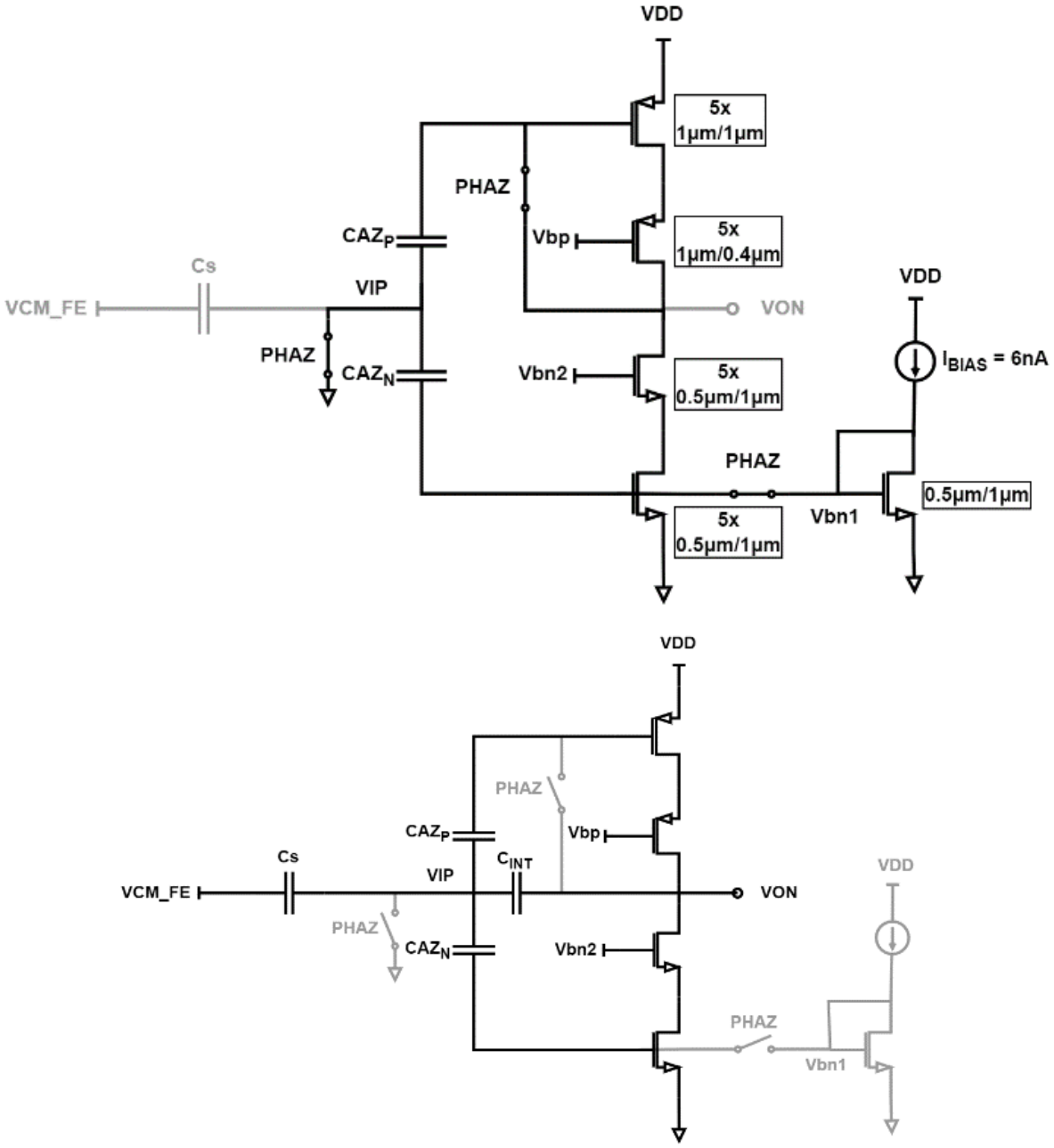


Figure 3.14 Single-sided inverter-based amplifier in a) autozero phase; b) integration phase.

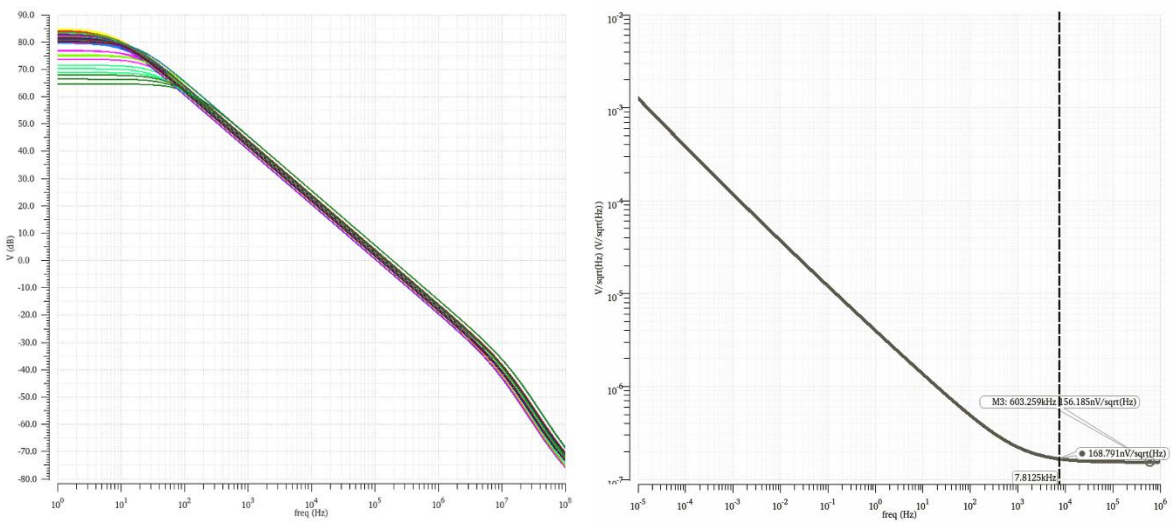


Figure 3.15 a) Monte Carlo amplifier's DC gain; b) Amplifier's input-referred noise at room temperature.

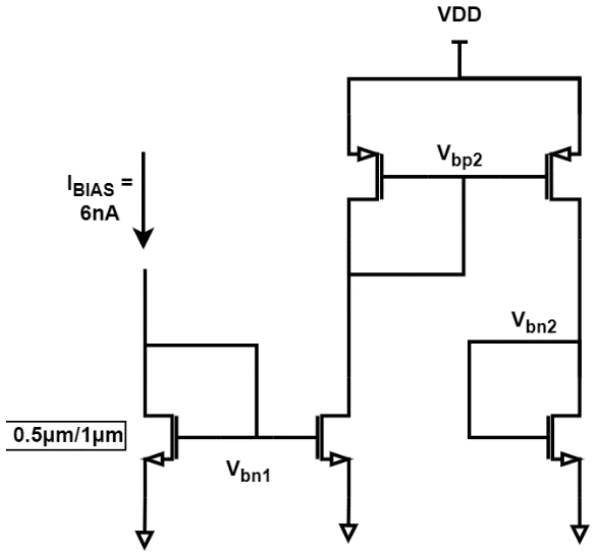


Figure 3.16 OTA Biasing circuit.

3.2.3.1 Reducing the autozero capacitors

To reduce C_{AZ}, we look at the error in accuracy caused by decreasing its value. Firstly, to achieve a 3σ inaccuracy of 0.20°C, with the front-end designed above of an inaccuracy of ~0.18°C, the DSM should contribute an error no larger than:

$$e_{amp} < \sqrt{0.2^2 - 0.18^2} \approx 0.09^o C \tag{3.3}$$

To be on the safe side, the amplifier error budget is set to 0.06°C. With this, the error the amplifier can make on ΔV_{BE} is then:

$$V_{error,\Delta V_{BE}} = \frac{0.06}{S_{\Delta V_{BE}}^{D_{out}}} \tag{3.4}$$

where $S_{\Delta V_{BE}}^{D_{out}}$ is the sensitivity of the decimated output to ΔV_{BE} given by Equation *(2.7)* and which has a value of approximately 526. With this, the differential error in ΔV_{BE} is 114µV. Assuming a 10% mismatch between the transistors in the two OTA branches, this results in an allowable error of 1.2mV in the sampled voltage on C_{AZ}.

There are three error sources to be considered here: the effect of sampled noise on C_{AZ} on the accuracy, the charge injection and drift due to leakage.

Assuming these error sources are maximum, each error source has a budget of $\frac{1.2mV}{\sqrt{3}} \sim 690\mu V$. Therefore, the sampled kT/C noise for 5σ is given by:

$$5 \times \sqrt{\frac{kT}{C_{AZ}}} < 690\mu V \tag{3.5}$$

From this, we get that C_{AZ} > 260fF at 125°C.

The charge injection and leakage are determined by both C_{AZ} and the type and size of the autozero switch. Therefore, we can first choose a large enough C_{AZ} that satisfies the noise and area requirements and determine the AZ switch topology and size based on the leakage and charge injection requirements. In addition to the noise requirement, C_{AZ} must be ~10x larger than the gate parasitic capacitance of the amplifier's input pair (~38fF) to avoid any charge sharing. Therefore, C_{AZ} of 400fF is chosen.

For this C_{AZ} value, the allowed leakage current through the AZ switches is given by:

$$\frac{I_{leak} \times \frac{3}{4} t_{AZ}}{C_{AZ}} < 690\mu V \tag{3.6}$$

where t_{AZ} is the autozero period of $1/f_{AZ} \approx 133\mu s$, and AZ has a duty cycle of 25%. From this, we get that I_{leak} < 2.7pA across temperature. To minimise the leakage, the autozero switches were implemented with minimum-size HVT NMOS devices. The largest leakage current is then 1.4pA across PVT, and so the voltage across C_{AZ} drifts by 320µV at most over one autozero period across PVT.

Lastly, charge injection due to these AZ switches results in an amplifier output common mode voltage (V_{CM}) difference of 625µV, which is tolerable, considering that the output V_{CM} is at $V_{DD}/2$ ~ 450mV. With a 10% mismatch between the two branches of the amplifier, the charge injection will cause only a small offset of 6.25µV after chopping, fitting within the accuracy budget for the amplifier.

However, the smaller C_{AZ} results in larger sampled noise, which will affect the resolution. To mitigate that, choppers are added around the amplifier, as discussed in the following section.

3.2.3.2 Choppers

The choppers are added to the amplifier as shown in *Figure 3.17*. Input choppers are implemented using minimum-size NMOS switches for small charge injection. For the output choppers, there are two options: either implementing them between the cascodes and the input pair or at the output. Here, they were implemented in cascode to ensure fast settling. The drawback is that at this position, they cannot remove

the 1/f noise of the cascode transistors. However, they contribute smaller noise in comparison to the input pair. In addition, they use minimum-size switches for small charge injection. Because of the voltage level across the choppers, the choppers on the PMOS side were implemented using PMOS devices, while on the NMOS side, they used NMOS devices.

The frequency at which the choppers operate should be larger than the corner frequency of the amplifier's flicker noise, which is at 1.1kHz. The chopping frequency is set at $f_S/2$ = 7.5kHz, which allows for two conversion cycles per chopping period.

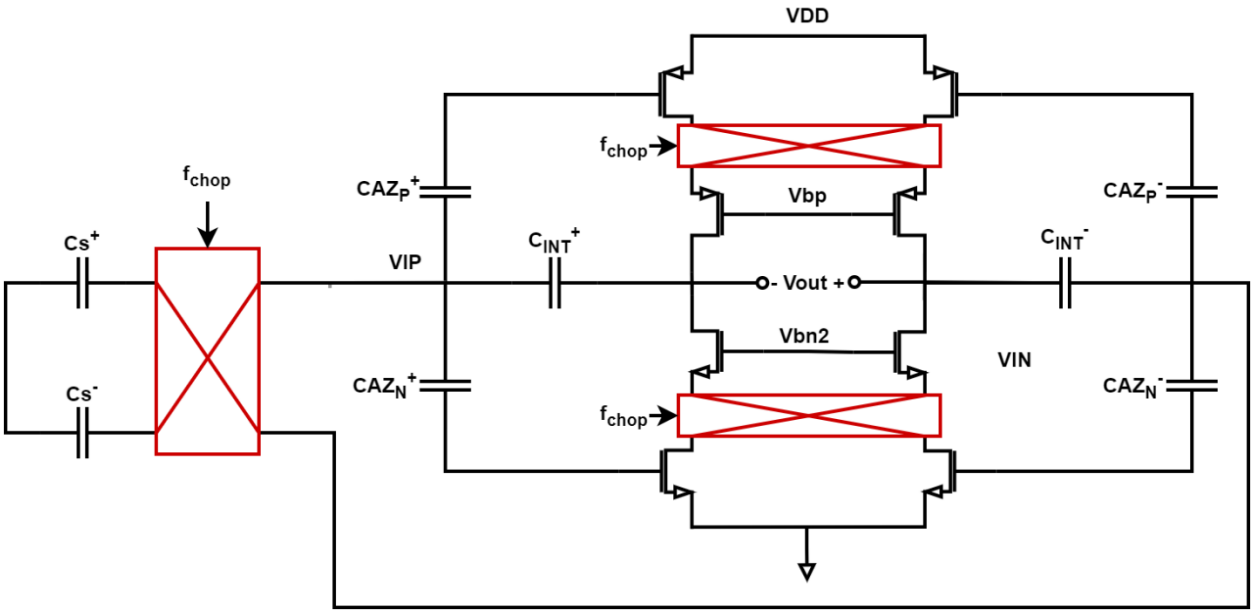


Figure 3.17 First integrator with choppers.

3.2.3.3 CMFB

The common-mode feedback circuit (CMFB) is shown in *Figure 3.18* [17].

During Φ_1, the common-mode feedback capacitors, C_{cm}, are connected to the ground, V_{SS}, and the supply voltage, V_{DD}. Looking at a single side, the charge on the two capacitors is:

$$Q_{\Phi 1} = C_{cm} \times 0 + C_{cm} \times V_{DD} = C_{cm} \times V_{DD} \tag{3.7}$$

During Φ_2, C_{cm} is connected between the virtual ground at the input of the amplifier (VirtP and VirtN) and the output of the amplifier (VOP and VON). Looking at a single side, the charge on the two capacitors is:

$$Q_{\Phi 2} = C_{cm} \times (V_{OP} - Virt_P) + C_{cm} \times (V_{ON} - VirtP) \tag{3.8}$$

$$Q_{\Phi 2} = C_{cm} \times (V_{OP} + V_{ON} - 2Virt_P) \tag{3.9}$$

Through conservation of charge, we have that the voltage at the input of the amplifier, VirtP, is regulated as:

$$Virt_P = \frac{V_{OP} + V_{ON}}{2} - \frac{V_{DD}}{2} \tag{3.10}$$

This way, the output V_{CM} is regulated back to $V_{DD}/2$ by correcting the charge at the virtual node of the amplifier.

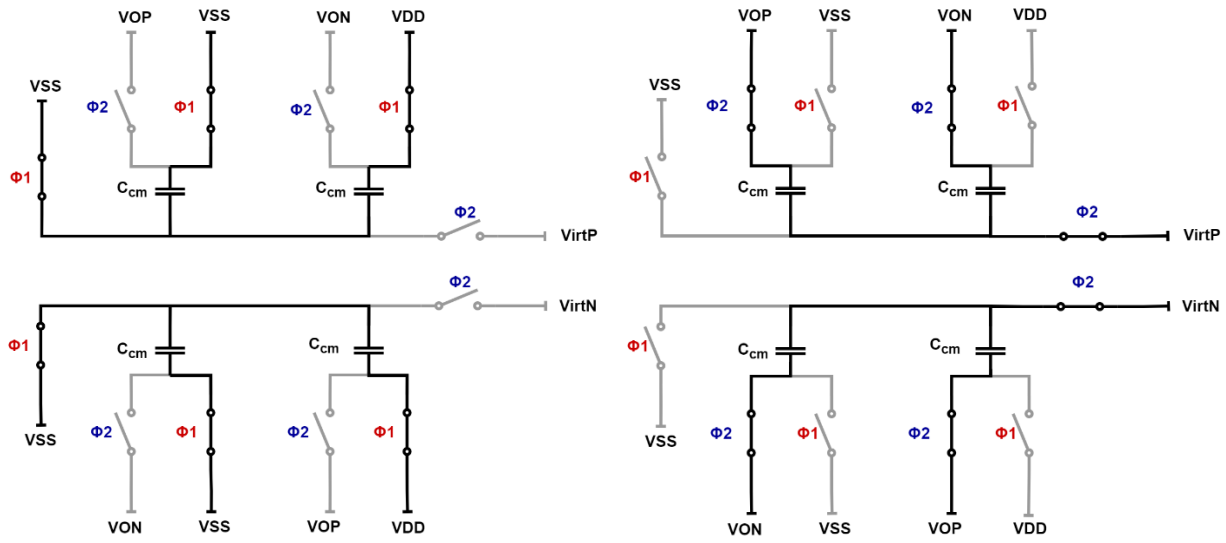


Figure 3.18 CMFB circuit during a) Φ1 and b) Φ2.

The advantage of the switch-capacitor CMFB at the virtual ground of the amplifier is the correction of the V_{CM} of the front-end, shown in *Figure 3.19*. During the sampling phase, Φ_1, the CTAT voltage V_{BE} is sampled across both the C_S and the BJT parasitic capacitor C_{be}. Then, during Φ_2, when the charge is transferred to the integrator, it is transferred with respect to the front-end $V_{CM,in}$. Ideally, this voltage is:

$$V_{CM,in} = \frac{V_{BE1} + V_{BE2}}{2} \tag{3.11}$$

However, due to the presence of the BJT parasitic capacitors, it is:

$$V_{CM,in} = \frac{V_{BE1} + V_{BE2}}{2} + \frac{C_{be}}{C_S + C_{be}} V_B \tag{3.12}$$

In this implementation, $V_B \sim V_{BE} \approx 600mV$, while $C_S = 0.8pF$. For a 5x5 PNP, the parasitic capacitor $C_{be} \approx 32fF$, resulting in a shift in the front-end $V_{CM,in}$ of 23mV.

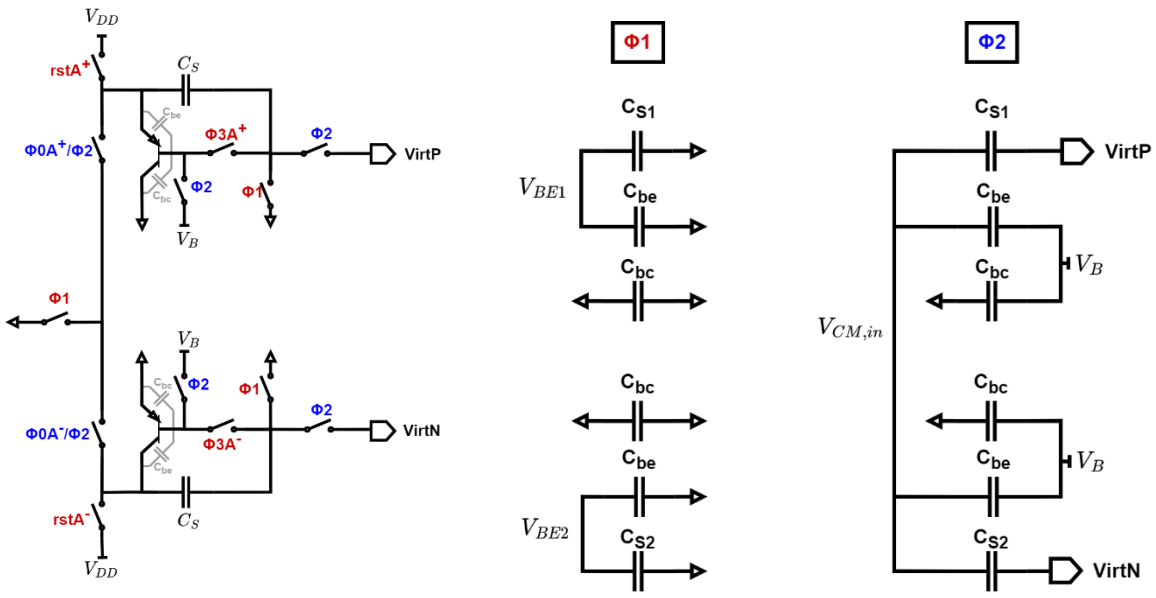


Figure 3.19 a) Front-end VCM shift circuit diagram; b) Voltage across CS and parasitic capacitors during sampling and integration phase.

To fix this shift in the input $V_{CM,in}$, we need to choose carefully the common-mode capacitor values, C_{cm}, from *Figure 3.18*. On the one hand, these capacitors should be large to compensate for the input $V_{CM,in}$ shift as follows:

$$2C_S \times \Delta V_{CM,in} = 2C_{cm} \times \Delta V_{CM,out} \tag{3.13}$$

On the other hand, C_{cm} should be as small as possible to reduce their noise contribution.

With a 900mV supply voltage, 100mV V_{DS} across each transistor in the amplifier and a ±200mV output swing, the amplifier can tolerate a common-mode shift of ±40mV at the output. This results in a C_{cm} size of 0.4pF, which is two times smaller than in the previous design, reducing both the area as well as the noise contribution of the CMFB.

To conclude on the first integrator, through added chopping as well as switch optimisation, without compromising the accuracy, the amplifier area was significantly reduced. The area-dominating autozero capacitors were scaled down by a factor of 42.5 in comparison to [1].

3.2.3 Implementation of the 2nd integrator

The second stage of the modulator is a scaled version of the first integrator with smaller capacitors (C_{S2} = 0.1pF, C_{INT2} = 0.25pF) and smaller power consumption of the OTA (6.2nA per branch) (*Figure 3.20*). The noise of the second stage is suppressed by the first stage, meaning there is no requirement for choppers.

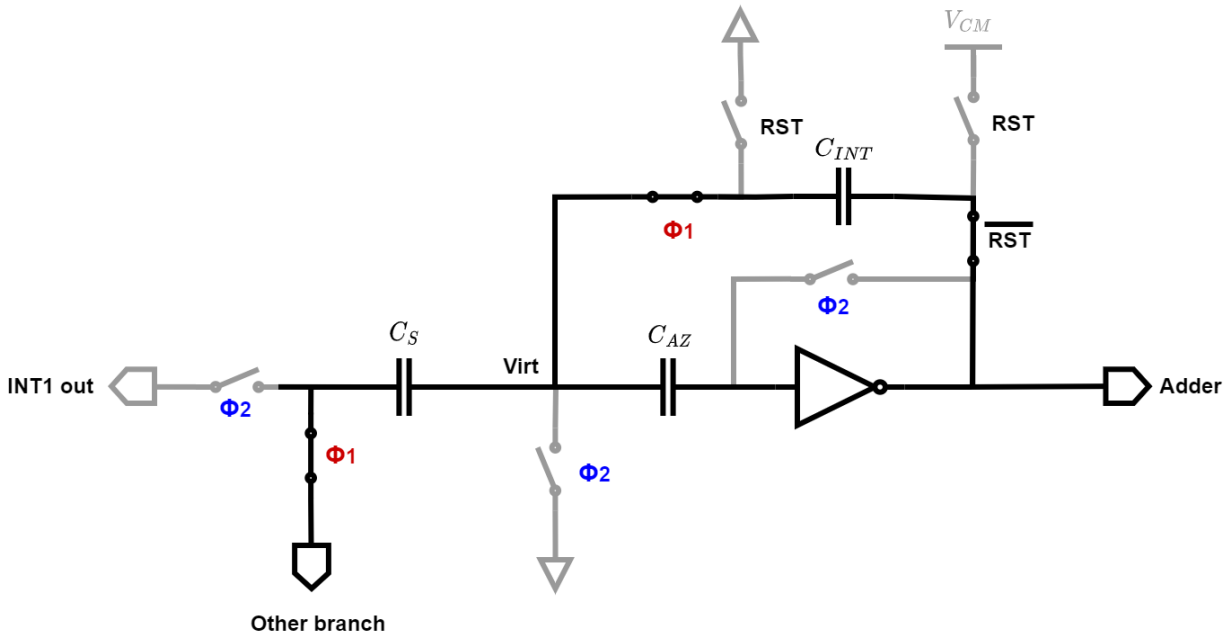


Figure 3.20 Second integrator.

The outputs of the first and second stages are added using a switch-capacitor adder with small capacitors of 0.1pF and 0.125pF, respectively.

3.3 Timing control

The clock signals for the front-end and the DSM are generated from an external clock of 1MHz via digital logic. Flexibility for the time ratio p = 8,16,32,64 is programmed on the chip. Furthermore, to reduce the mismatch between the BJTs and small C_S in the front-end, system-level chopping and dynamic element matching (DEM) were used.

System-level chopping is shown in *Figure 3.21*. One conversion of 64ms is conducted with V_{BE} generated from the BJT in the positive half of the differential front-end, and in the second conversion of 64ms, V_{BE} is generated by the BJT in the negative half of the differential front-end.

The timing diagram for a bitstream-controlled DEM is shown in *Figure 3.22*. The comparator makes a decision just before the end of Φ_2, and the available bitstream determines whether in the next cycle (in Φ_1) $2\Delta V_{BE}$ or $\Delta V_{BE} - V_{BE}$ are sampled. When the bitstream output is 1, in one cycle, the $-V_{BE}$ voltage is generated from front-end FE1; in the other cycle, it is generated from front-end FE2.

Both chopping and DEM ensure that all four BJTs in the front-end generate VBE, minimising mismatch errors and improving accuracy.

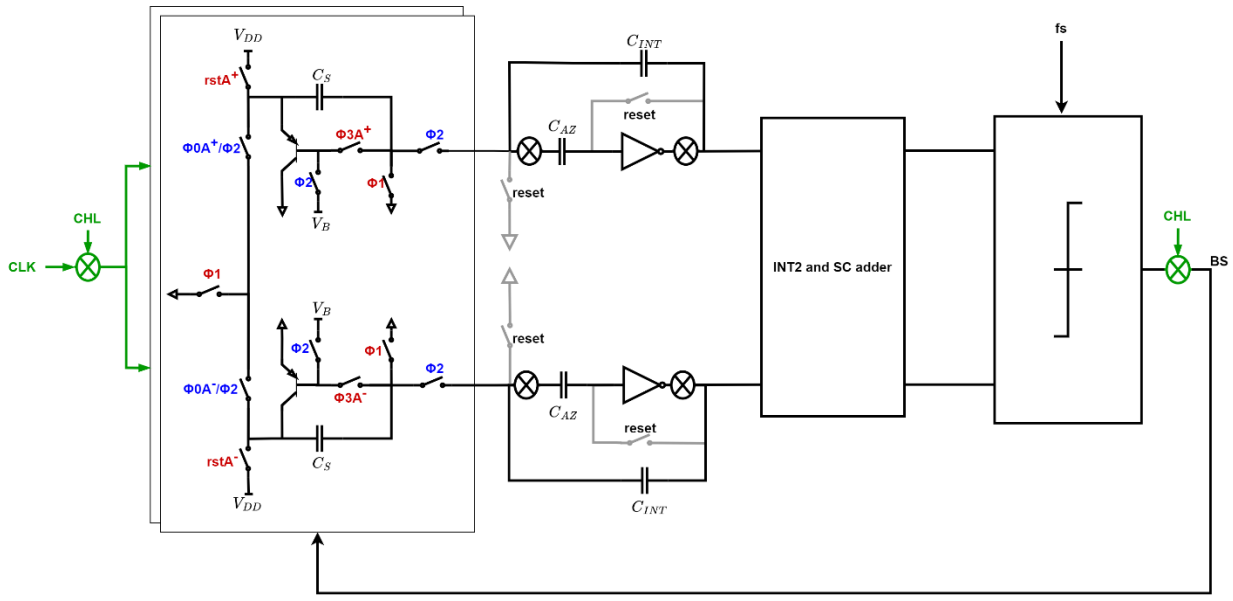


Figure 3.21 System-level chopping.

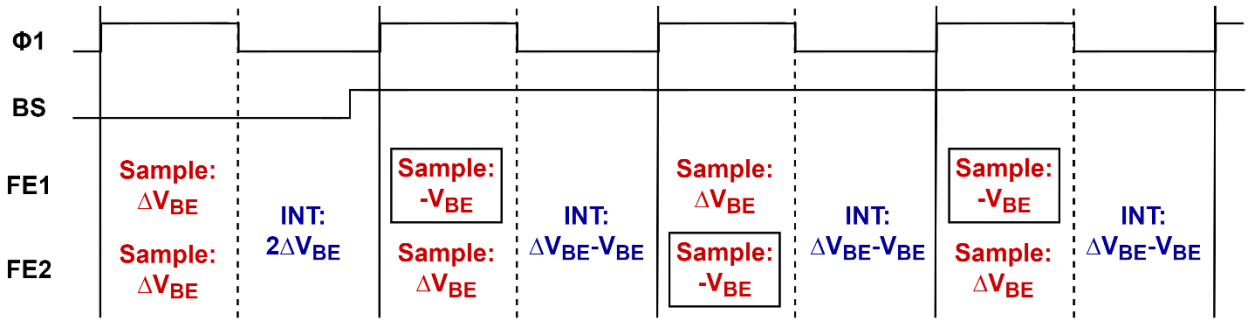


Figure 3.22 Dynamic element matching (DEM) timing diagram.

3.3.1 Clock boosters

Lastly, some of the switches in the sensor require clock signals boosted beyond rail to $2V_{DD}$ due to their leakage and settling requirements. The advantages of the conventional clock boosters used in [1] are their fast start-up and single-clock phase. However, they occupy a large area (as seen in *Figure 2.1* and *Figure 2.2*) using two capacitors. Therefore, a new clock booster circuit [16] is used in this design, as shown in *Figure 3.23*.

The main advantage of these clock boosters in comparison to the original design is that they use only one capacitor, reducing the total capacitance and area, in addition to using a smaller number of transistors. With a C_B = 0.4pF, the voltage is boosted from 0.9V (V_{DD}) to 1.72V ($2V_{DD}$) across PVT with a turn-on time of 15ps.

This way, the total capacitance of the clock boosters has been reduced by 3.2x compared to the original design (*Figure 3.24*), resulting in an area and power reduction, with no cost to the clock boosters' performance.

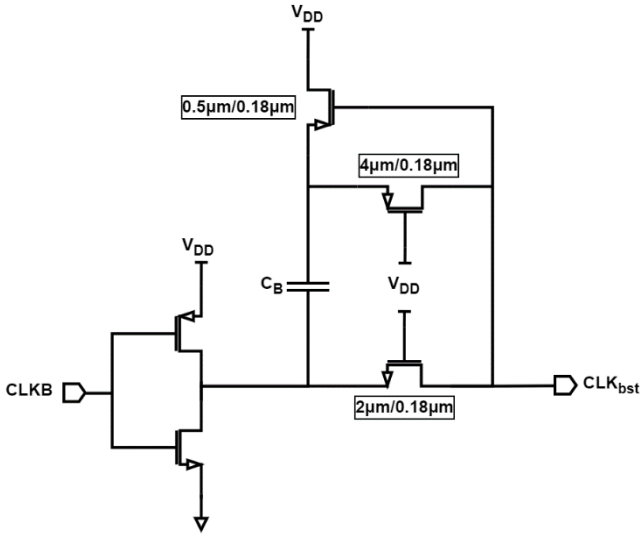


Figure 3.23 Circuit schematic of the implemented single-capacitor clock booster [16].

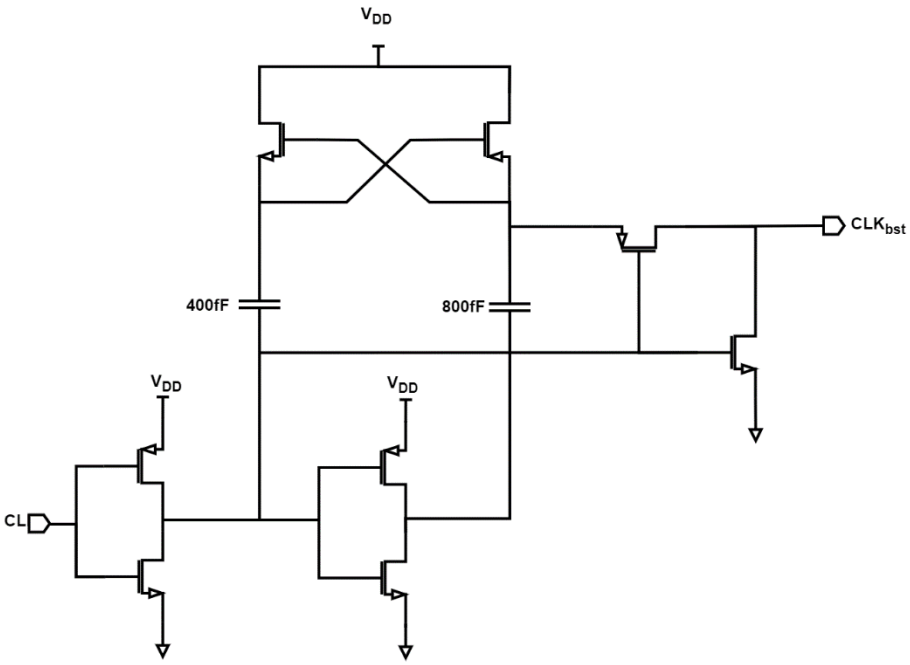


Figure 3.24 Circuit schematic of the clock boosters in [1].

3.3 Area estimation

After going through different techniques to reduce the area of the chip, a summary of the area reduction of each block is shown in *Figure 3.25*. As seen, the total block area (without top-level wiring) is estimated to be reduced by a factor of 3.9, from 0.17mm^2 to 0.044mm^2. With the top-level wiring, the total block area is estimated to increase 1.5 times, resulting in an estimated total area of 0.066mm^2, which meets the design goal. Additionally, the total capacitance was reduced by a factor of 8.6, from 69pF to 8pF, as seen in *Table 3.2*.

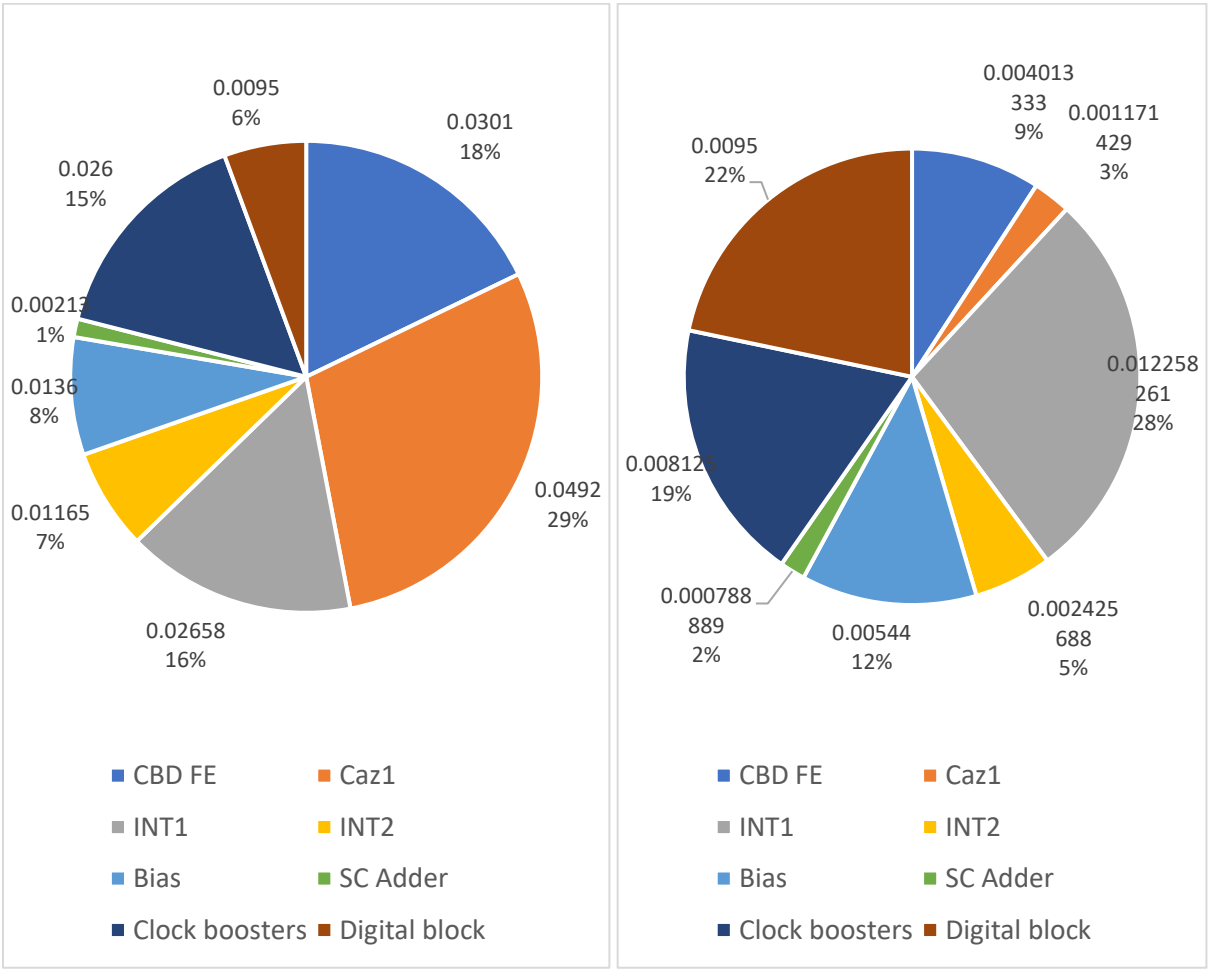


Figure 3.25 Area distribution in mm^2 of each block in a) design in [1]; b) this design.

Table 3.2 Dominant capacitor values in the original [1] and this design.

Single-sided capacitors	Original design [pF]	Scaling down factor	This design [pF]
C_{S1}	3×4	7.5	2×0.8
C_{INT1}	12	4.3	2.6
C_{S2}	0.45	4.5	0.1
C_{INT2}	1.8	7.2	0.25
C_{FF1}	0.2	2	0.1
C_{FF2}	0.4	3.2	0.125
C_{cm1}	2×0.8	2	2×0.4
C_{cm2}	2×0.05	1	2×0.05
C_{AZ1}	2×16.8	42	2×0.4
C_{AZ2}	2×1.6	4	2×0.4
$C_{S,bias}$	4	5	0.8
Total capacitance	69.35	8.6	8.075

3.4 Summary

In this chapter, the circuit implementation of the main blocks of the sensor was discussed, mainly the CBD front-end and the DSM. They were designed to reduce the total area of the chip, as shown above. The next chapter will discuss the simulation results of the chip in Cadence Virtuoso.

4 Simulation results

This chapter shows the simulated results of the CBD-based temperature sensor described in the previous chapter.

4.1 Modulator behaviour

The output swings of the first and second integrator and the adder across temperature and corner are shown in *Figure 4.1* and *Figure 4.2*, separately. They vary across temperature since the sampled input voltages ΔV_{BE} and V_{BE} also vary with temperature, with V_{BE} being largest at cold temperatures. Over the military temperature range, both integrators were designed to have output swings of ±200mV, which is within the given headroom of the sub-1V OTAs. The output swing results across temperature and corner are shown in Table 4.1.

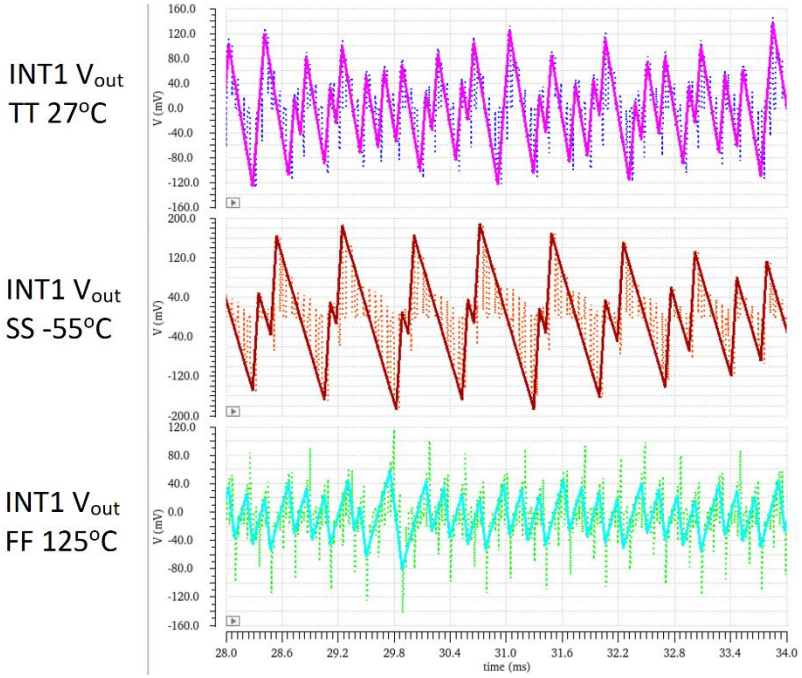


Figure 4.1 Output swing of the first integrator across corner and temperature.

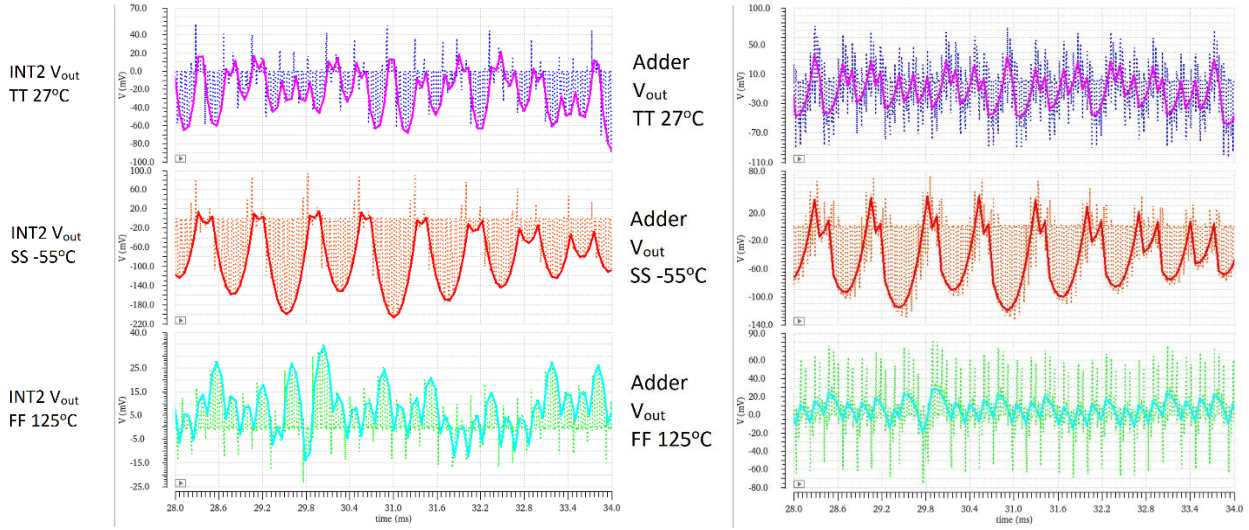


Figure 4.2 Output swings of the a) 2nd integrator; and b) the SC adder across corners and temperature.

Table 4.1 Output swings of the a) 1st integrator; b) 2nd integrator; and c) the SC adder across corners and temperature.

		Max swing [mV]	Min swing [mV]
TT 27°C	**INT1**	142.8	-145.7
	INT2	27.0	-96.7
	Adder	41.1	-62.9
SS -55°C	**INT1**	186.7	-187.0
	INT2	16.0	-208.5
	Adder	43.5	-119.7
FF 125°C	**INT1**	58.5	-79.3
	INT2	36.5	-14.2
	Adder	30.1	-19.2

The output common mode voltage (V_{CM}) of the first and second integrator across temperature and corner is shown in *Figure 4.3*. The output V_{CM} of the first integrator shifts by ~30mV at most in the SS corner at cold temperatures. This corresponds to the input V_{CM} shift of the front-end (described in Chapter 3), as seen in *Figure 4.4*. The front-end input V_{CM} shifts by 17.7mV at SS -55oC at most, which causes the 30mV shift in the V_{CM} at the output of the first integrator. The second integrator's output V_{CM} shifts by 35mV at most. The large fluctuations in the V_{CM} present at FF 125°C in both integrators are due to the effect of leakage on the biasing voltage stored on the autozero capacitors, C_{AZ}. Considering the V_{CM} drift over temperature and corner, as well as the output swing, the amplifier can work well at a supply as low as 0.9V.

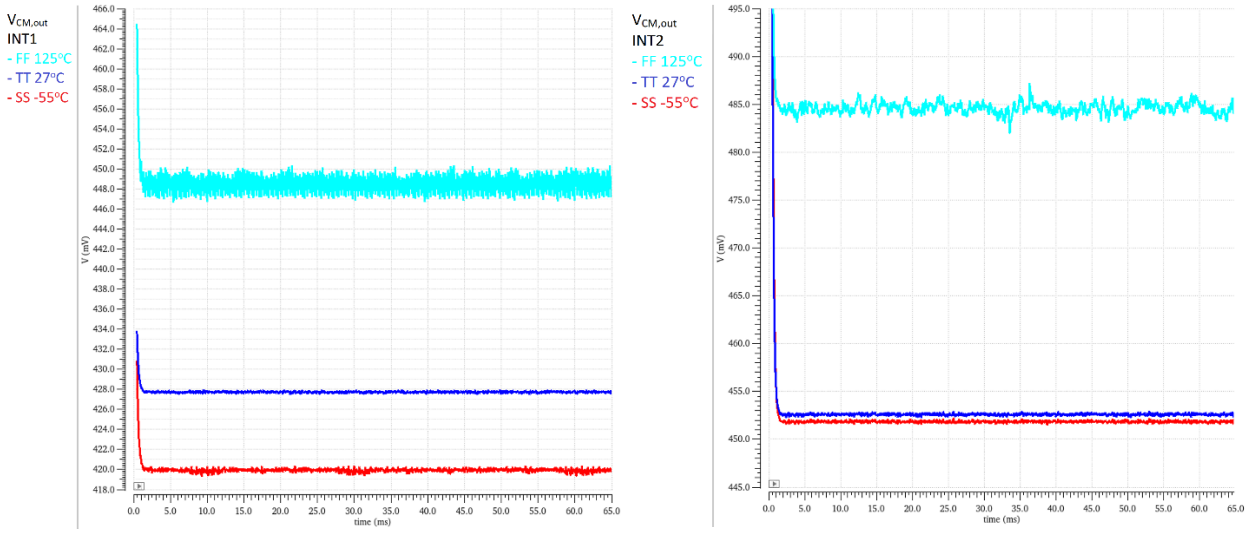


Figure 4.3 Output VCM of the a) 1st integrator and b) 2nd integrator across temperature and corner.

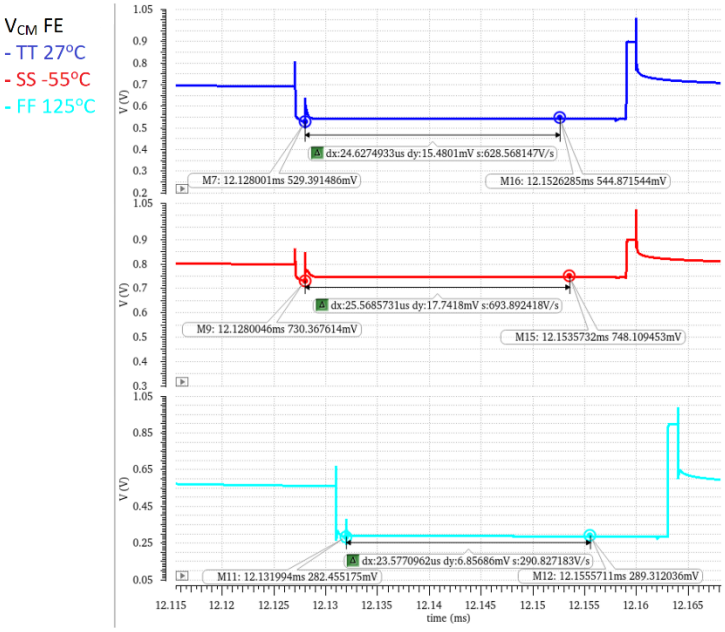


Figure 4.4 Input VCM of the front-end across temperature and corner.

4.2 Accuracy

In Chapter 3, the front-end was designed to achieve an inaccuracy of 3σ = 0.18°C after one-point trim with an ideal readout circuit. The readout circuit should not worsen this result to an inaccuracy larger than 0.2°C. Due to the high gain of each stage of the DSM, the first integrator will be the most dominant error source in the readout circuit. The transient Monte Carlo accuracy results of the SC front-end and the

implemented first integrator with and without a one-point trim at room temperature are shown in *Figure 4.5*. The inaccuracy goes up to 0.20°C after a one-point trim at a supply voltage of 0.9V. The inaccuracy does not worsen further after adding choppers to the amplifier in the first integrator, as can be seen in *Figure 4.6*. Similarly, at a supply of 0.95V, the inaccuracy is 0.19°C, as seen in *Figure 4.7*.

The fitting coefficients for converting the decimated output μ into temperature are:

$$T = A\frac{\alpha\mu}{\alpha\mu + 1} - B = 566.5 \times \frac{3.8\mu}{3.8\mu + 1} - 281.8 \tag{4.1}$$

Lastly, the decimated output obtained with an ideal second stage and quantiser and real ones is compared, resulting in a temperature error of less than 0.1°C across corner and temperature. Finally, the entire sensor has a 3σ inaccuracy of 0.2°C.

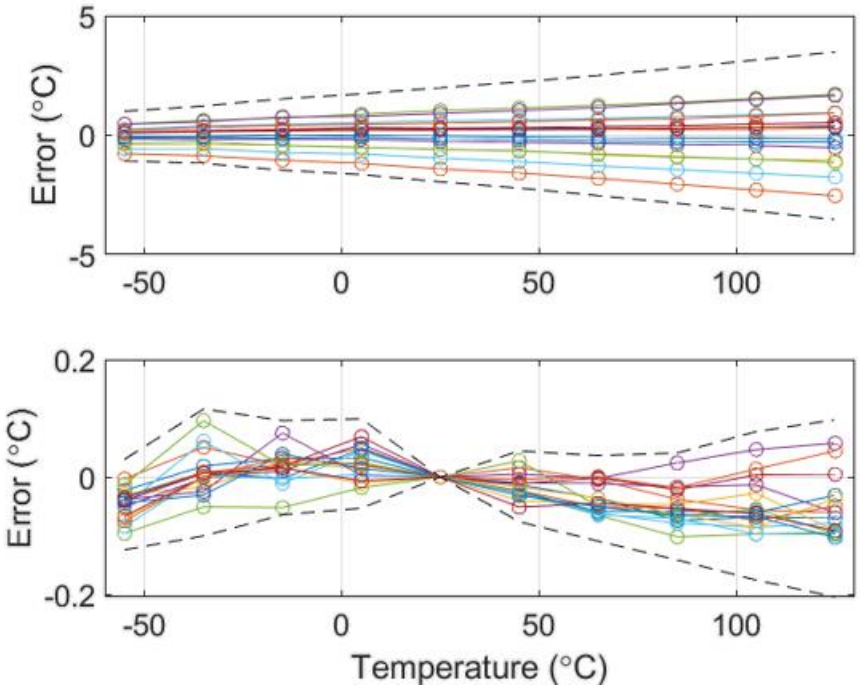


Figure 4.5 Monte Carlo (20 points, mismatch and process spread) inaccuracy of the front-end and first integrator (without OTA choppers) with C_S = 800fF, V_{DD}=0.9V and p = 32 without trimming (top) and with a 1-point PTAT trim at 25°C (bottom).

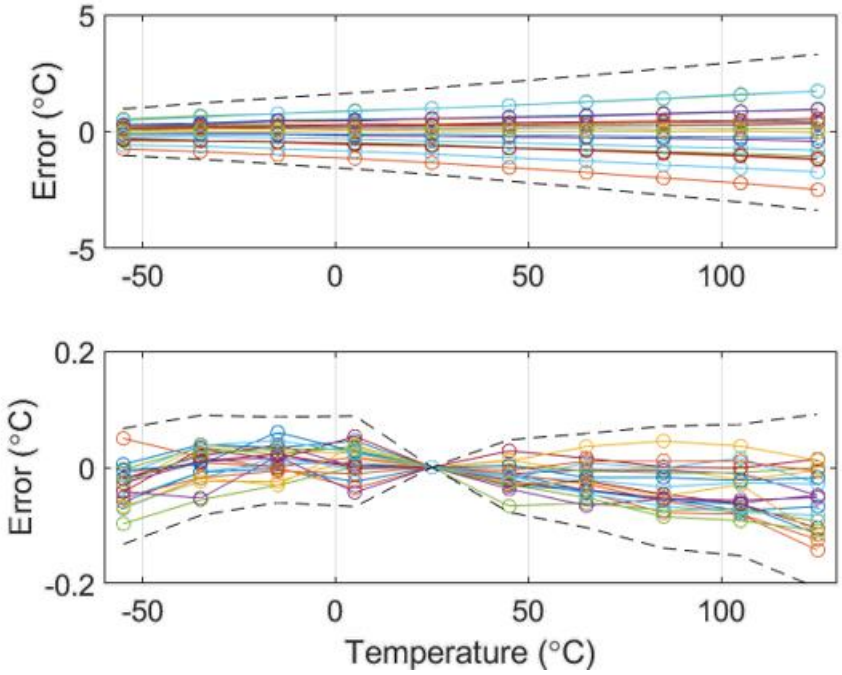


Figure 4.6 Monte Carlo (20 points, mismatch and process spread) inaccuracy of the front-end and first integrator with OTA choppers with C_S = 800fF, V_{DD}=0.95V and p = 32 without (top) and with a 1-point PTAT trim at 25°C (bottom).

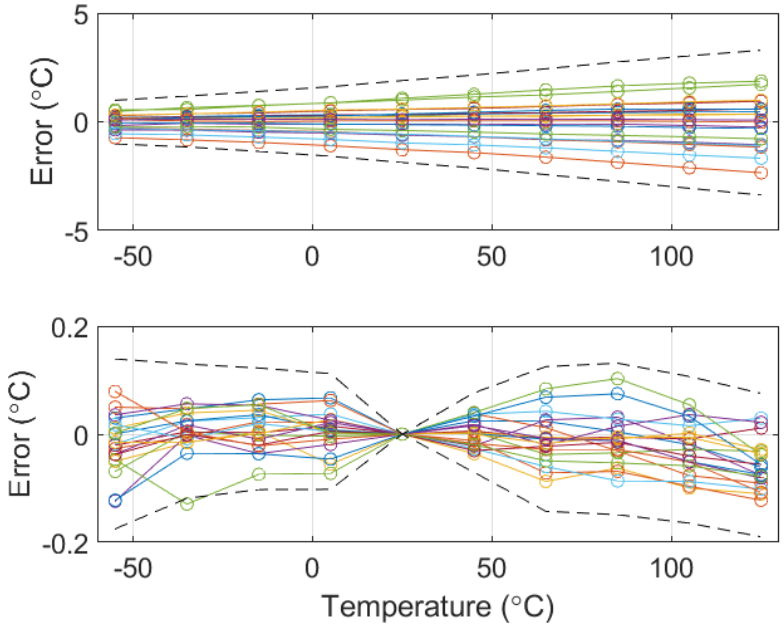


Figure 4.7 Monte Carlo (20 points, mismatch and process spread) inaccuracy of the front-end and first integrator with OTA choppers with CS = 800fF, VDD=0.95V and p = 32 without (top) and with a 1-point PTAT trim at 25oC (bottom).

4.2.1 Power supply sensitivity of the front-end

Monte Carlo simulations were done at three temperatures (-55°C, 27°C, 125°C) with supply voltages from 0.9 V to 1.4 V. The sensor achieves a power-supply sensitivity (PSS) of 0.08°C/V at room temperature and 0.24°C/V across temperatures (*Figure 4.8*), which align with the measured PSS in [1]. The large spread at 125°C is due to the leakage of the front-end switches.

For a shorter discharge time of t_1 = 16µs and t_2 = 0.5µs, the PSS from 0.9 V to 1.4 V is 0.44°C/V, shown in *Figure 4.9*. However, accepting the fact that the sensor no longer works well at 0.9V, the PSS from 0.95 to 1.4V is at most 0.11°C/V across temperatures (0.03°C/V at room temperature), compared to 0.10°C/V (0.07°C/V at room temperature) with the longer discharge time.

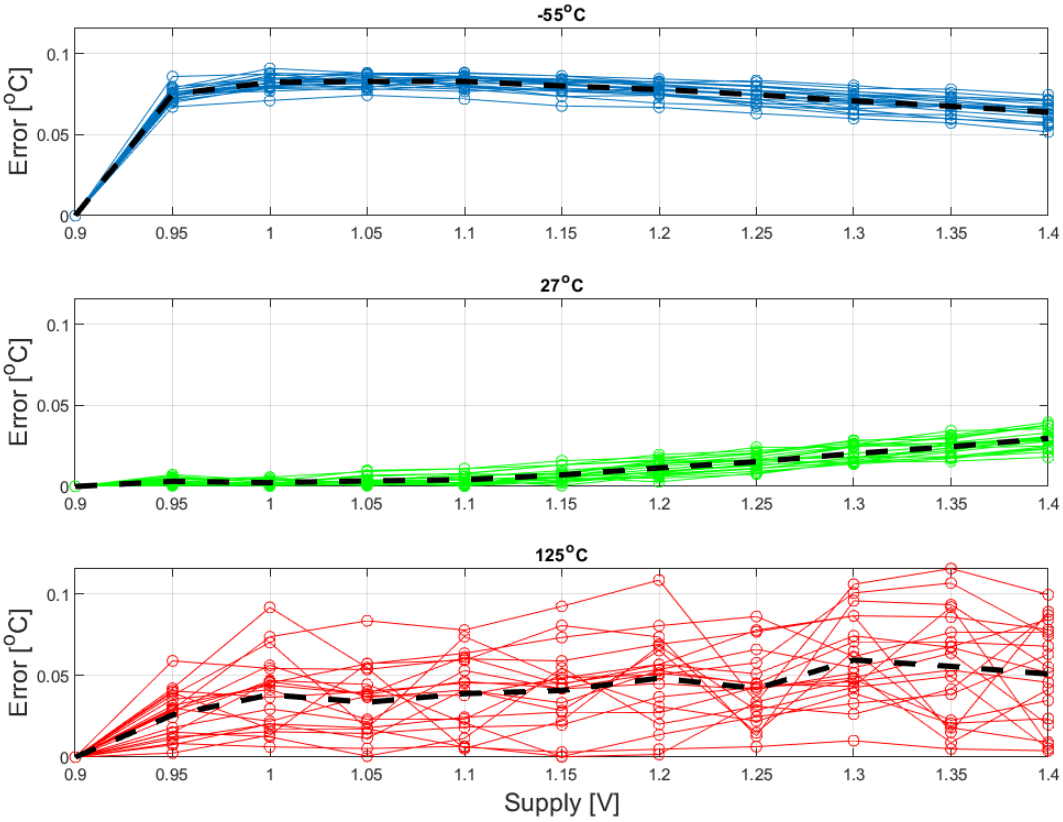


Figure 4.8 Monte Carlo (20 points, process spread) simulation of power-supply sensitivity (PSS) at a) SS -55C; b) TT 27oC; c) FF 125oC.

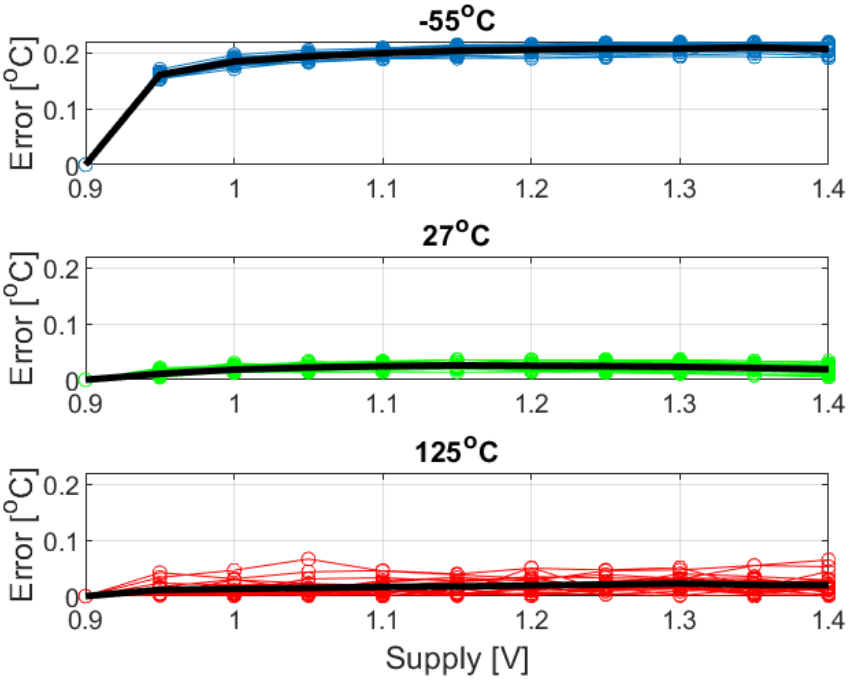


Figure 4.9 Monte Carlo (20 points, process spread) simulation of power-supply sensitivity (PSS) with halved discharge time at a) SS -55C; b) TT 27oC; c) FF 125oC.

4.3 Layout and area estimation

One of the main goals of this design was to reduce the chip area. Originally, the area was 0.25mm^2 in a standard 180nm CMOS process [1], and the objective was to halve it at least. Since most of the area is occupied by large capacitors such as the sampling capacitors, C_S, and the autozero capacitors in the amplifier, C_{AZ}, the main task was reducing the total capacitance of the circuit. This was reduced by a factor of 8.6, from 69pF to 8pF. The area without top wiring is estimated to be reduced by a factor of 3.8 to 0.044mm^2. The estimated layout of the design in comparison to the chip in [1] is shown in *Figure 4.10*. With wiring, this ultimately results in a total area estimation of 0.066mm^2 and an area improvement by 3.8x, as per the specifications.

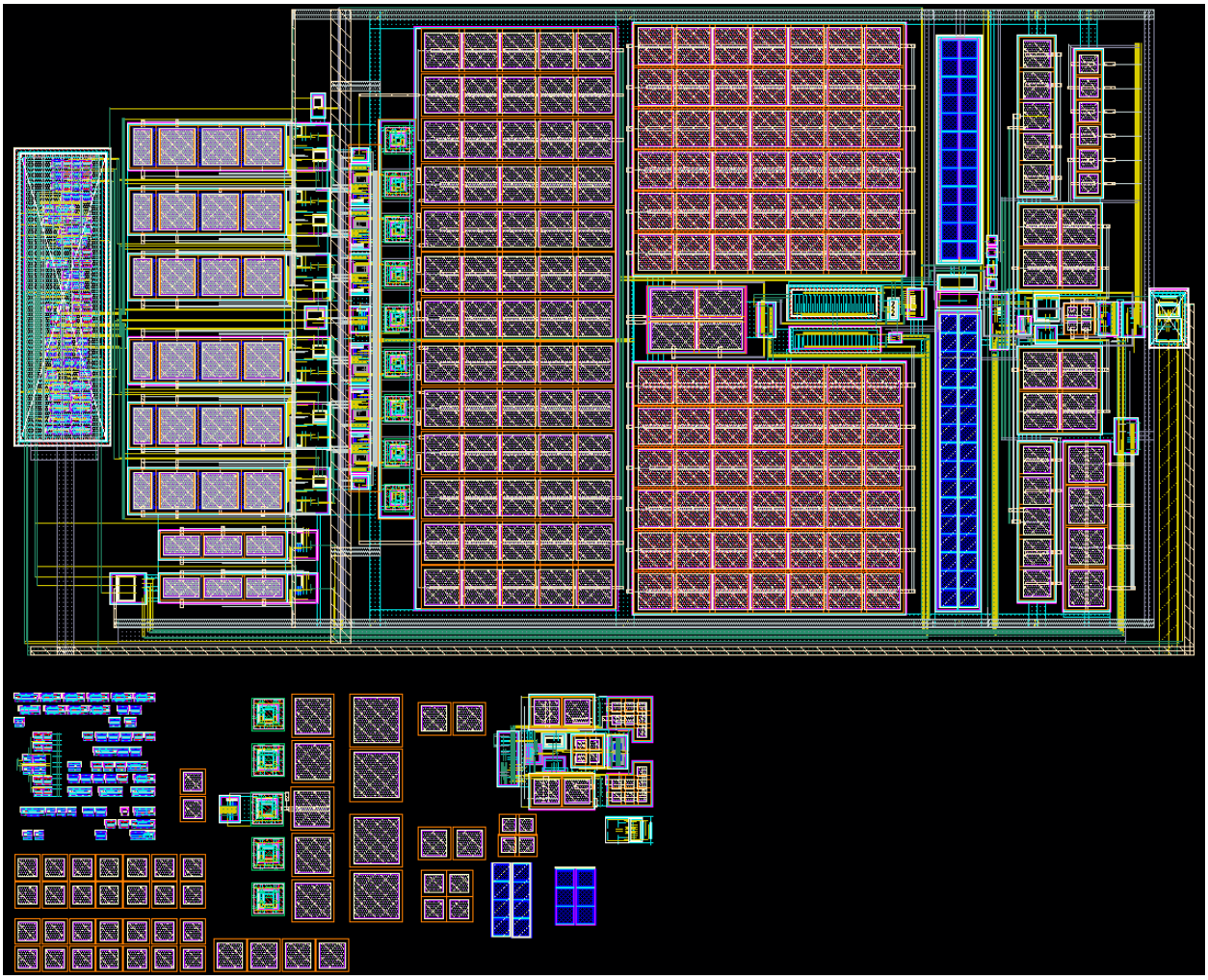


Figure 4.10 Area comparison of design in [1] (top) and estimated layout of this design (bottom).

4.4 Power dissipation

By reducing the design's total capacitance, the circuit's dynamic power decreases as well. Since this dominates, a reduction of the analogue power by a factor of two is expected. *Figure 4.11* and *Figure 4.12* show the simulated power breakdown for a supply voltage of 0.9V, which is the minimum supply voltage. As seen from *Figure 4.12*, the digital circuitry now consumes most of the power (52%). Compared to the original design [1], the digital power was reduced from 225nW to 176nW (at 1V supply, Table 4.2), mainly because the number of CBD FEs, and hence the required clock signals, was reduced from three to two. As for the analogue sub-blocks, the total power was reduced by a factor of 4.3, from 685nW to 159nW (at 1V supply, Table 4.2). Most of the analogue power is consumed by the first integrator (*Figure 4.11*), the clock boosters and the CBD front-end.

Since the design operates at 0.9V supply voltage, the total simulated power here is 275.7nW, with the analogue part consuming only 132.7 nW. With this, the total power has been reduced from 810nW [1] by a factor of 3.

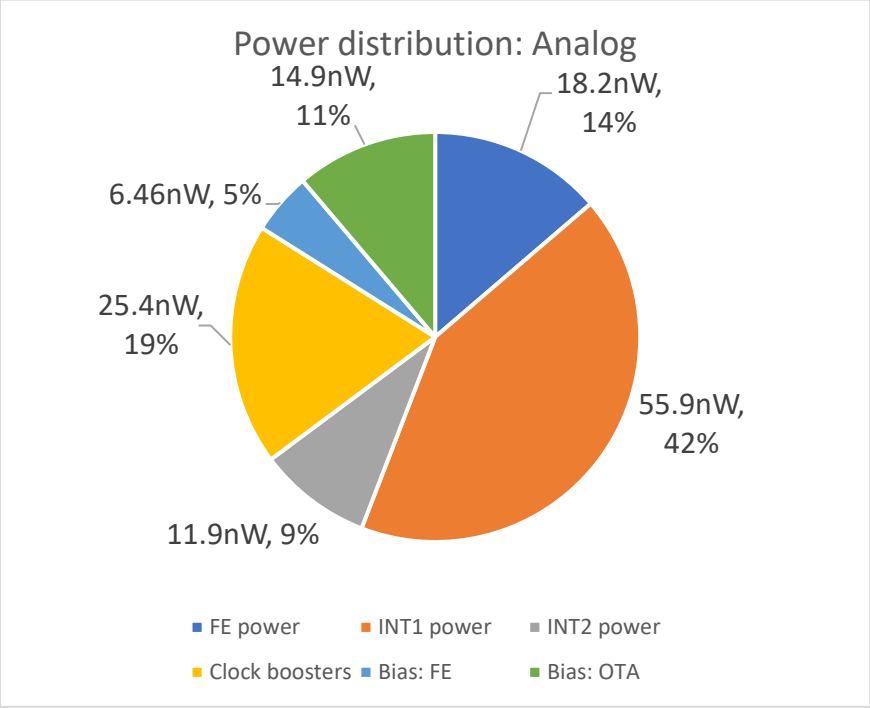


Figure 4.11 Simulated analogue power consumption breakdown per block at 0.9V supply.

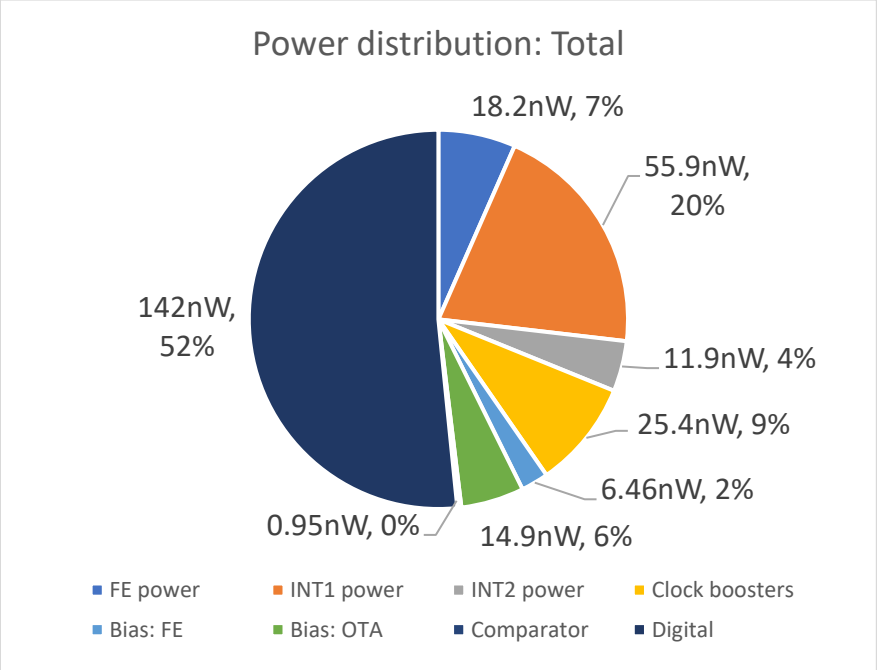


Figure 4.12 Simulated total power consumption breakdown per block at 0.9V supply.

Table 4.2 Comparison of power consumption between the original [1] and this design at VDD = 1V.

V_{DD} = 1V	Original design [1]	This design
Analog power [nW]	685	159
Digital power [nW]	225	176
Total power [nW]	890	335

4.5 Resolution

The FFT of the modulator's bitstream output at room temperature is shown in *Figure 4.13*. Clear second-order noise-shaping in addition to the input signal at DC can be observed. Additionally, *Figure 4.13* compares the effect of the chopping of the first amplifier on the folded-back noise due to the small C_{AZ} size. An improvement of ~5dB in the noise floor can be seen.

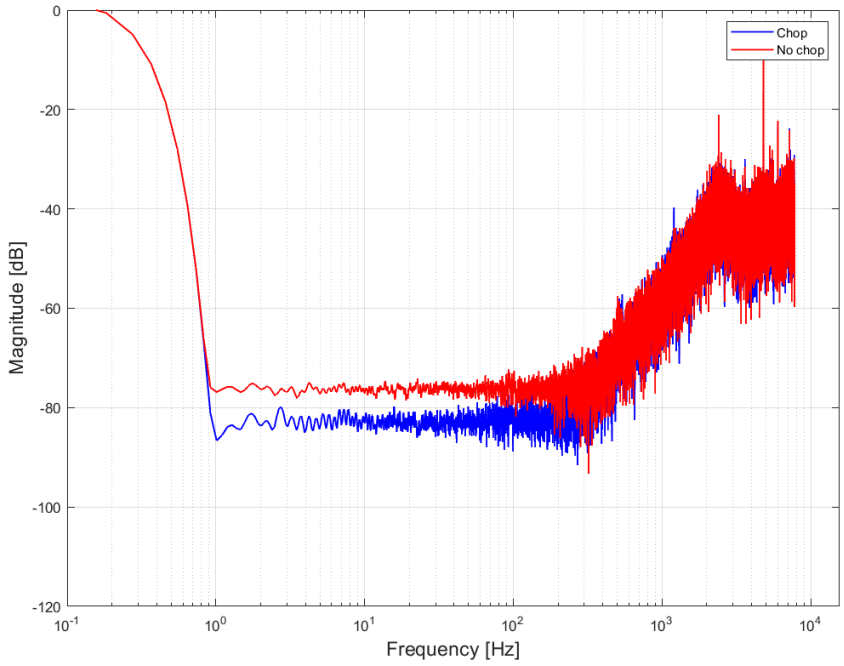


Figure 4.13 Modulator's frequency response with and without chopping.

The simulated resolution versus conversion time plot at room temperature is shown in *Figure 4.14*. The plot shows that for short conversion periods < 60ms, the resolution is limited by the quantisation noise, following a slope of $\frac{1}{T_{Conv}^2}$. After ~60ms, the resolution becomes thermal noise limited, with a slope of $\frac{1}{\sqrt{T_{conv}}}$. At a conversion time of 128ms, the design achieves a resolution of 2.9mK. This corresponds to an SNR of:

$$SNR = 20\log\left(\frac{Temp\ range}{Resolution}\right) = 20\log\left(\frac{125-(-55)}{2.9m}\right) = 95.9dB \quad (4.2)$$

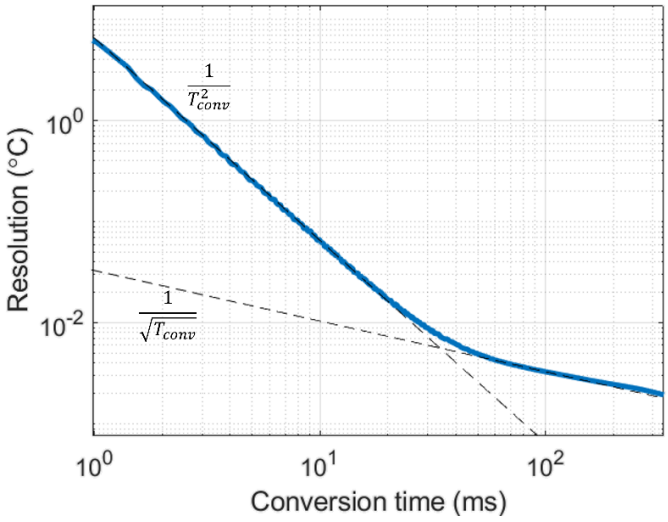


Figure 4.14 Resolution versus conversion time.

4.5.1 Figure of merit

The figure of merit (FoM) for a temperature sensor is calculated as:

$$FoM = Resolution^2 \times T_{conv} \times Power \quad (4.3)$$

For this temperature sensor with a total power dissipation of 276nW, the FoM versus conversion time is shown in *Figure 4.15*. For a conversion time of 128ms and a resolution of 2.9mK obtained, the FoM is 310fK^{2}J, in the same order as the FoM of the original design (340fK^{2}J) [1].

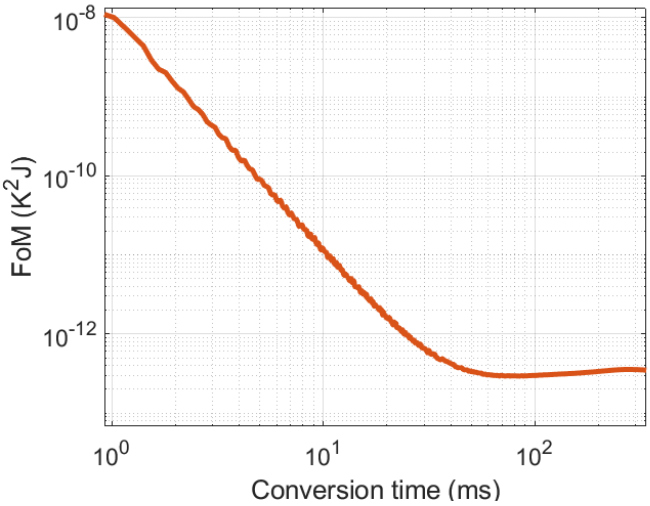


Figure 4.15 Figure of Merit versus conversion time.

4.6 Summary

Simulated results of the CBD-based BJT temperature sensor are shown here. The design achieves a simulated accuracy of 0.2°C over the temperature range of -55°C to 125°C, occupying an area of 0.066mm^2 and consuming a total power of 270nW, while operating at a supply voltage of 0.9V (Table 4.3). Although not based on measurement results, these results indicate that the design should consume lower power than previously reported CBD-based temperature sensors, as well as being the smallest realization in a 180nm process.

Table 4.3 Performance summary and comparison to the state-of-the-art.

	2019 [2]	**2020 [7]**	**2021 [13]**	**2023 [1]**	**This work**
Technology	16nm FinFET	28nm	55nm	180nm	180nm
Type	Bulk diode SAR	DTMOST OSC	P+/Nwell DT ΣΔ	PNP DT ΣΔ	PNP DT ΣΔ
Area [mm^2]	0.0025	0.017	0.021	0.25	0.066
Supply [V]	0.85 – 1	0.85 – 1.15	1 – 1.2	0.95 – 1.4	0.9 – 1.4
T. Range [°C]	-15 – 105	-10 – 90	-55 – 125	-55 – 125	-55 – 125
3σ Error [°C] (Trim point)	+1.5/-2.0 (0)	±2.0 (0) ±0.9 (1)	±1.4 (0) ±0.6 (1)	±0.45 (0) ±0.15 (1)	±0.2 (1)
R.I.A. [%] (Trim point)	2.9 (0)	4 (0) 1.8 (1)	1.6 (0) 0.67 (1)	0.5 (0) 0.17 (1)	0.22 (1)
Power [µW]	18	33.75	2.2	0.81	0.27
Tconv [ms]	0.013	0.1	6.4	128	128
Res. [mK]	300	10.2	15	1.8	2.9
PSS [°C/V]	1.5	0.27	3.7	0.2	0.2
Res. FoM [pJK^2]	21	0.36	3.1	0.34	0.31

5 Conclusion and future work

5.1 Conclusion

A capacitively-biased BJT-based temperature sensor operating at a 0.9V supply and a power consumption of 270nW has been designed. It is an improved and scaled version of the temperature sensor described in [1], in which the total capacitance has been reduced by 8.6x and the estimated area 3.8x. The design achieves a competitive inaccuracy of ±0.2°C (3σ) with one-point trim over a temperature range from -55°C to 125°C. It also achieves a resolution of 2.9mK with a conversion time of 128ms. This was enabled through dynamic error suppression techniques such as chopping, in addition to careful sizing and optimisation. These were necessary because the currents in the sensor were reduced 5x in comparison to [1] resulting in more significant leakage effects. The resulting sensor can be used for IoT applications due to its compact size and low power consumption while maintaining good accuracy and resolution.

5.2 Future work

5.2.1 Reducing the conversion period

In comparison to the other CBD-based temperature sensors included in *Table 4.3*, this design has a relatively large conversion time, T_{conv}, of 128ms. Future work should be done to reduce this.

One potential approach is reducing the discharge time of the front-end BJTs. Care must be taken to not reduce the discharge time to an extent where it reaches the supply-dependent CBD operating region.

Another way the conversion period can be reduced is by utilising all the integrator's phases for integration, as described in the sections below.

5.2.2 Reducing digital power

The focus of this design was to reduce the power consumption of the analog front-end and the ΣΔ modulator. As a result, most of the power (52%) is dissipated by the digital circuits – mainly the clock signal generators' static power. This can be reduced by implementing the circuit in a more advanced technology node. This way, the total power consumption would be dominated by the analog power and could be reduced below 200nW.

5.2.3 Reducing analog power

The first integrator dissipates most of the analog power. However, the integrator does not use the entire conversion cycle to integrate the sampled charge. Instead, the integrator switches between autozero, integration, do nothing, and integration phases over two conversion cycles.

During the “do nothing” phase, the front-end is disconnected and samples the temperature-dependent voltage. The integrator is disconnected from both the front-end and the further stages and stands idle in a closed-loop configuration, holding the previously integrated charge on the integration capacitor. The amplifier is still dissipating energy, and wasting potential integration time. Below two proposals to improve the integrator’s energy efficiency are presented.

5.2.3.1 First integrator in off-mode

The first proposal is to turn off the amplifier during the “do nothing” phase. One readily available way of doing this method would be to turn off the output chopper switches (*Figure 5.1a*). The chopper’s operating behaviour is shown in *Figure 5.1b*. During autozeroing, the choppers are in the positive configuration, as well as during the first integration cycle. Then, they open during the “do nothing” phase, stopping current flow through the OTA. Then, in the second integration cycle, the chopping switches close in the negative configuration, completing one chopping cycle.

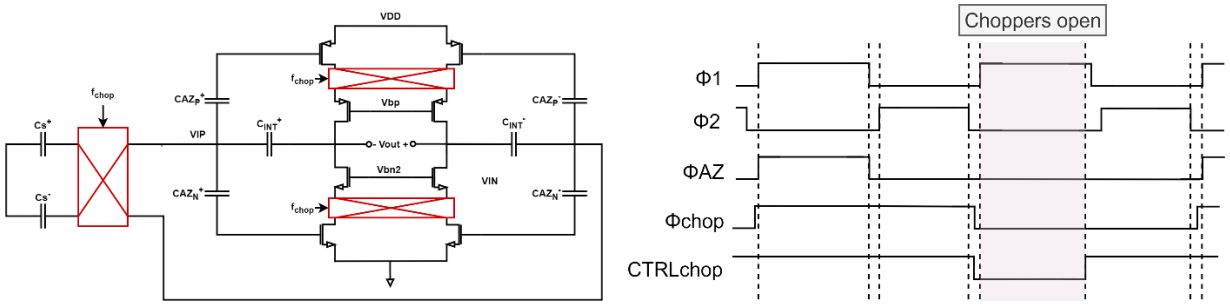


Figure 5.1 a) Schematic of the first amplifier and b) choppers timing signal.

Some essential things need to be taken into consideration with this implementation.

Firstly, one must ensure that the charge on the integration capacitors (C_{INT}) does not leak during this period and is not affected by the common-mode charge injection of the chopper’s switches. This can be achieved by disconnecting C_{INT} from the amplifier.

Secondly, since the output node is floating, the output common mode voltage, $V_{CM,out}$ may drift to VDD or GND. To prevent this, the output nodes could be connected to a fixed voltage during this phase. However, this would require additional switches, which would introduce leakage current and settling considerations.

Thirdly, the amplifier needs to be properly biased before the integration phase starts. This means that it cannot be turned off during the entire “do nothing” phase. However, depending on the settling speed of the chopping switches, this will take only a minor part near the end of the “do nothing” phase (~1µs).

Alternatively, the amplifier can be turned off by biasing the input pair’s gate with a voltage that ensures that they will be turned off, such as ground for the NMOS input transistor and V_{DD} for the PMOS input transistor. But in order not to negate the benefits of chopping, this will have to be done without disturbing the voltage across the autozero caps.

5.2.3.2 Ping-pong front-end

Alternatively, energy efficiency can be improved by using both the autozero phase and the "do nothing" phase for integration. In this way, the amplifier is used continuously. However, since the front-end still requires a discrete-time operation, which samples the V_{BE} voltage in one phase and integrates it in the other, a ping-pong operating mode based on two front-ends is proposed.

Figure 5.2 shows the single-sided connection of the two front-ends to the first integrator. During Φ_1, FE1 is sampling $-V_{BE}$ or ΔV_{BE} while FE2 is connected to the integrator and the temperature-dependent charge is being integrated. Similarly, during Φ_2, FE2 is sampling while FE1 is connected to the integrator. This way, half of the charge is transferred in each phase. At the end of Φ_2 the integrator has integrated $2\Delta V_{BE}$ or $\Delta V_{BE} - V_{BE}$ depending on the bitstream. The rest of the modulator loop further integrates this charge.

The advantages of the ping-pong mechanism are that during each phase, only one front-end C_S is connected to the integrator, loading the amplifier lightly and increasing its bandwidth:

$$C_{load} = C_{s2} + C_{cm} + C_{ff} + C_{int} || C_{s1}$$

where C_{S2}, C_{cm} and C_{ff} are the sampling capacitors of the second stage, the common-mode feedback capacitors, and feed-forward capacitors respectively, not shown in the previous figure. Less capacitance will allow for a smaller amplifier transconductance to achieve similar bandwidth, so that less current is required.

However, the ping-pong brings some disadvantages as well.

Firstly, by transferring a portion of the charge in one phase and another portion in the other, the largest input is now V_{BE} for the ping-pong architecture in comparison to $V_{BE} - \Delta V_{BE}$ for the standard architecture. This will require a slightly (~10%) larger integration capacitor to keep the output of the amplifier from clipping.

Secondly, sampling in both phases requires a greater complexity in clock generation.

Thirdly, whether $\Delta V_{BE} - V_{BE}$ or $2\Delta V_{BE}$ is sampled depends on the bitstream output, which is available at the end of Φ_2. To reduce the mismatch between the BJTs and the small C_S in the front-ends, the V_{BE} generation should be alternated between the two front-ends. However, since one of the front-ends should already have finished sampling VBE before the bitstream result is available, this requires a more complicated sampling technique, where not the entire sampling period is utilised but only a fraction once the bitstream result is known.

Lastly, multiplexing between the two front-ends, while the integrator is always on, may result in cross-talk between the front-ends, which in turn may cause errors. This requires extra complexity in the design of the multiplexer that switches between the two front-ends, as this must provide good isolation as well as low leakage.

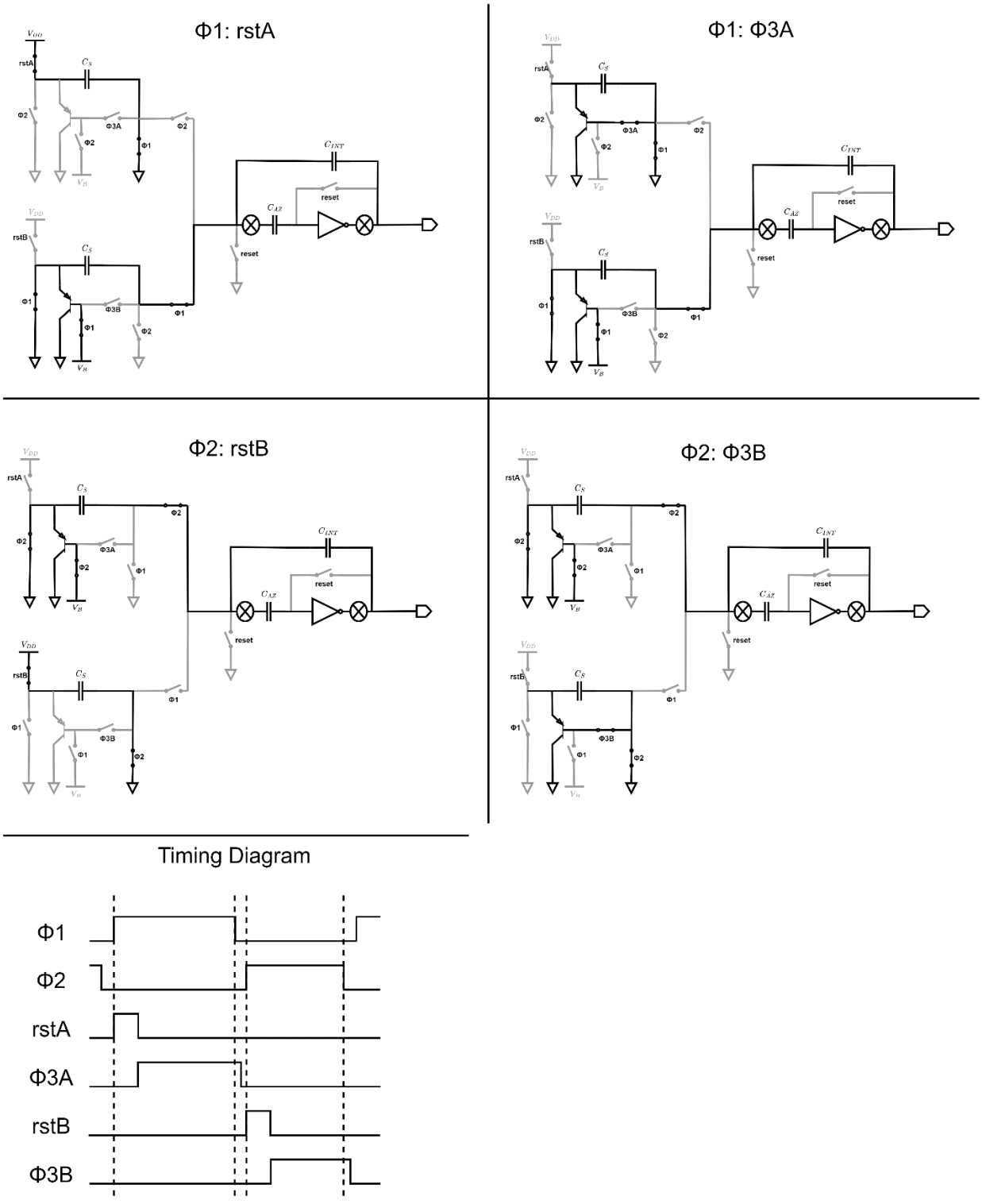


Figure 5.2 Opreation of the proposed ping-pong front-end with its timing diagram.

5.2.3.3 Implementing the second integrator in continuous-time

Building upon the ping-pong strategy, the second integrator can also be used in both Φ_1 and Φ_2. This will increase the energy efficiency of the second integrator as well. This way, the bitstream output is available twice within a sampling period: at the end of Φ_1 and Φ_2, allowing the overall conversion time to be halved, while still maintaining the sampling frequency of the front-end. This would further improve the energy efficiency of the sensor.

Another significant disadvantage of using both phases for integration is that the amplifier can only be biased at the start of a conversion. This means that the charge on C_{AZ} must be held throughout the conversion period, and so the leakage of the autozero switches must be further reduced.

5.2.2.4 Conclusion

The proposed modifications would improve the design's energy efficiency, power consumption and conversion time while maintaining good accuracy and area at the cost of slightly more complex clock signal generation and switch optimisation for low leakage.

Bibliography

[1] Z. Tang, S. Pan, and K. A. A. Makinwa, “23.5 a sub-1v 810nw capacitively-biased bjt-based temperature sensor with an inaccuracy of ±0.15°c (3) from 55°c to 125°c,” in *2023 IEEE International Solid- State Circuits Conference (ISSCC)*, 2023, pp. 22–24.

[2] M. Eberlein and H. Pretl, “A no-trim, scaling-friendly thermal sensor in 16nm finfet using bulk diodes as sensing elements,” in *ESSCIRC 2019 - IEEE 45th European Solid State Circuits Conference (ESSCIRC)*, 2019, pp. 63–66.

[3] E. H. Hellen, “Verifying the diode–capacitor circuit voltage decay,” *American Journal of Physics*, vol. 71, no. 8, pp. 797–800, 2003.

[4] M. A. Pertijs and J. Huijsing, *Precision temperature sensors in CMOS technology*. Springer Science & Business Media, 2006.

[5] Someya, V. van Hoek, J. Angevare, S. Pan, and K. Makinwa, “A 210 nw npn-based temperature sensor with an inaccuracy of ±0.15 °c (3) from 15 °c to 85 °c utilizing dual-mode frontend*,” IEEE Solid-State Circuits Letters*, vol. 5, pp. 272–275, 2022.

[6] K. Souri, Y. Chae, F. Thus, and K. Makinwa, “12.7 a 0.85v 600nw all-cmos temperature sensor with an inaccuracy of ±0.4°c (3) from 40 to 125°c,” in *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 2014, pp. 222–223.

[7] S. Park, Y. Kim, W. Choi, Y. Lee, S. Kim, Y. Shin, and Y. Chae, “A dtmost-based temperature sensor with 3 inaccuracy of ±0.9°c for self-refresh control in 28nm mobile dram,” in *2020 IEEE Custom Integrated Circuits Conference (CICC)*, 2020, pp. 1–4.

[8] S. Pan, Y. Luo, S. Heidary Shalmany, and K. A. A. Makinwa, “A resistor-based temperature sensor with a 0.13 pj · k2 resolution fom,” *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 164–173, 2018.

[9] B. Yousefzadeh, S. H. Shalmany, and K. Makinwa, “A bjt-based temperature-to-digital converter with ±60mk (3) inaccuracy from 70°c to 125°c in 160nm cmos,” in *2016 IEEE Symposium on VLSI Circuits (VLSI-Circuits)*, 2016, pp. 1–2.

[10] S. Pan, J. A. Angevare, and K. A. A. Makinwa, “5.4 a hybrid thermal-diffusivity/resistor-based temperature sensor with a self-calibrated inaccuracy of ±0.25° c(3) from -55°c to 125°c,” in *2021 IEEE International Solid-State Circuits Conference (ISSCC)*, vol. 64, 2021, pp. 78–80.

[11] C. P. van Vroonhoven, D. d’Aquino, and K. A. Makinwa, “A thermal-diffusivity-based temperature sensor with an untrimmed inaccuracy of ±0.2°c (3s) from 55°c to 125°c,” in *2010 IEEE International Solid-State Circuits Conference - (ISSCC)*, 2010, pp. 314–315.

[12] K. Souri, Y. Chae, Y. Ponomarev, and K. A. A. Makinwa, “A precision dtmost-based temperature sensor,” in *2011 Proceedings of the ESSCIRC (ESSCIRC)*, 2011, pp. 279–282.

[13] Z. Tang, Y. Fang, X.-P. Yu, N. N. Tan, Z. Shi, and P. Harpe, “An energy-efficient capacitively biased diode-based temperature sensor in 55- nm cmos,” IEEE Solid-State Circuits Letters, vol. 4, pp. 210–213, 2021.

[14] C. Enz and G. Temes, “Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization,” *Proceedings of the IEEE*, vol. 84, no. 11, pp. 1584–1614, 1996.

[15] S. Pavan, R. Schreier, and G. C. Temes, *Understanding delta-sigma data converters.* John Wiley & Sons, 2017.

[16] M. Gu, Y. Tao, X. He, Y. Zhong, L. Jie, and N. Sun, “A 3.7mw 11b 1gs/s time-interleaved sar adc with robust one-stage correlation-based back- ground timing-skew calibration,” in *ESSCIRC 2023- IEEE 49th European Solid State Circuits Conference (ESSCIRC)*, 2023, pp. 145–148.

[17] Y. Chae and G. Han, “Low voltage, low power, inverter-based switched-capacitor delta-sigma modulator,” *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 458–472, 2009.