5-2018

Digitally Interfaced Analog Correlation Filter System for Object Tracking Applications

Mohsen Judy
University of Tennessee, mjudy@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Mohsen Judy entitled "Digitally Interfaced Analog Correlation Filter System for Object Tracking Applications." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

<div align="right">Jeremy H. Holleman III, Major Professor</div>

We have read this dissertation and recommend its acceptance:

Benjamin J. Blalock, Syed K. Islam, Vasileios Maroulas

<div align="right">Accepted for the Council:
<u>Dixie L. Thompson</u></div>

<div align="right">Vice Provost and Dean of the Graduate School</div>

(Original signatures are on file with official student records.)

Digitally Interfaced Analog Correlation Filter System for Object Tracking Applications

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Mohsen Judy

May 2018

Acknowledgements

I would like to thank my advisor Dr. Jeremy Holleman for his valuable guidance and support during this research.

I am also thankful to my colleagues at the *Integrated Silicon Systems* laboratory from August 2013 to December 2017, especially Nicholas C. Poor, Peixing Liu, and Tan Yang, as well as, Dr. Charles Britton for their contributions in the design of some of the circuit blocks that have been used in this dissertation.

I would also like to thank Dr. Aravind Mikkilineni and Dr. David Blome from Electrical and Electronics Systems Research Division at the Oak Ridge National Laboratory for their valuable contribution in chip-to-computer interfacing as well as providing the test data and analyzing the results.

I am also grateful to Dr. Benjamin J. Blalock, Dr. Syed K. Islam, and Dr. Vasileios Maroulas for serving as my Ph.D. committee members. Their valuable feedback improved the readability and organization of this dissertation.

And last but not least, I would like to thank my family for patiently enduring the inevitable situation of living apart far away during all these years.

The arc of the moral universe is long, but it bends toward justice.

Martin Luther King, Jr.

Abstract

Advanced correlation filters have been employed in a wide variety of image processing and pattern recognition applications such as automatic target recognition and biometric recognition. Among those, object recognition and tracking have received more attention recently due to their wide range of applications such as autonomous cars, automated surveillance, human-computer interaction, and vehicle navigation.

Although digital signal processing has long been used to realize such computational systems, they consume extensive silicon area and power. In fact, computational tasks that require low to moderate signal-to-noise ratios are more efficiently realized in analog than digital. However, analog signal processing has its own caveats. Mainly, noise and offset accumulation which degrades the accuracy, and lack of a scalable and standard input/output interface capable of managing a large number of analog data.

Two digitally-interfaced analog correlation filter systems are proposed. While digital interfacing provided a standard and scalable way of communication with pre- and post-processing blocks without undermining the energy efficiency of the system, the multiply-accumulate operations were performed in analog. Moreover, non-volatile floating-gate memories are utilized as storage for coefficients. The proposed systems incorporate techniques to reduce the effects of analog circuit imperfections.

The first system implements a 24×57 Gilbert-multiplier-based correlation filter. The I/O interface is implemented with low-power D/A and A/D converters and

a correlated double sampling technique is implemented to reduce offset and low-frequency noise at the output of analog array. The prototype chip occupies an area of $3.23mm^2$ and demonstrates a 25.2pJ/MAC energy-efficiency at 11.3 kVec/s and 3.2% RMSE.

The second system realizes a 24×41 PWM-based correlation filter. Benefiting from a time-domain approach to multiplication, this system eliminates the need for explicit D/A and A/D converters. Careful utilization of clock and available hardware resources in the digital I/O interface, along with application of power management techniques has significantly reduced the circuit complexity and energy consumption of the system. Additionally, programmable transconductance amplifiers are incorporated at the output of the analog array for offset and gain error calibration. The prototype system occupies an area of $0.98mm^2$ and is expected to achieve an outstanding energy-efficiency of 3.6pJ/MAC at 319kVec/s with 0.28% RMSE.

Table of Contents

Chapter 1

Introduction

1.1 Correlation Filter

Advanced correlation filters (CFs) have been employed in a wide variety of image processing and pattern recognition applications such as automatic target recognition (ATR) and biometric recognition[24]. Among those, object recognition and tracking [9, 10, 28] have received more attention recently due to their wide range of applications such as autonomous cars, automated surveillance, video indexing, human-computer interaction, traffic monitoring, and vehicle navigation. Figure 1.1 summarizes the CF utilization in pattern recognition applications.

Advances in designing robust but simple CFs that show a better performance at discriminating object and background have paved the way for implementing efficient object tracking systems using fewer computational resources.

The object tracking is performed in two main steps:

1) Detection of the objects of the interest, and

2) Tracking of such objects from frame to frame.

The focus of this work is implementing a CF system to carry out the first step. The object detection procedure using a CF is illustrated in Figure 1.2(a). First, in an off-line training process, the filter coefficients are designed. Images of the object
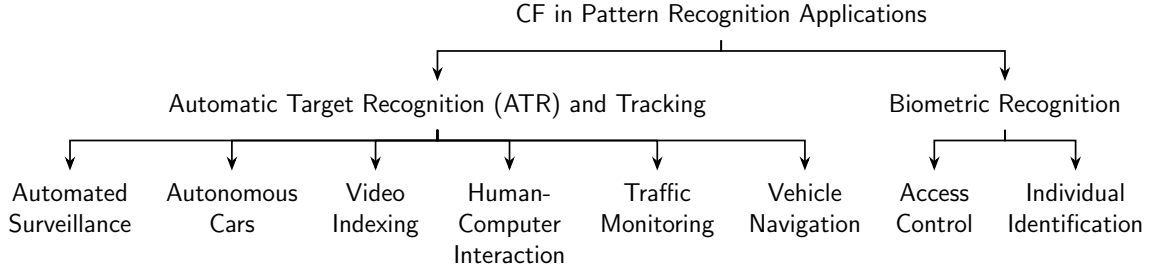
Figure 1.1: CF utilization in pattern recognition applications.

of interest at different backgrounds, orientations and, points of view are fed to an algorithm which trains the coefficients to generate a sharp peak at the output of the CF wherever a match is detected. In this work, training images are from DARPA VIVID dataset and the target object is a vehicle. The algorithm used here is minimum output sum of squared error (MOSSE) [9]. After obtaining filter coefficients for the target object, the system will respond to images containing the target object by generating a sharp peak at the output. It should be noted that images go through a preprocessing step beforehand in order to reduce shadows and intense lighting effects. Figure 1.2(a) shows simulated responses of the CF implemented in MATLAB to presence and absence of a vehicle on a road.

1.2   Analog Versus Digital Implementation

Digital signal processing (DSP) has long been used in a wide variety of application including audio and speech processing, image processing, telecommunications, control systems, etc. However, for some application e.g portable devices, the energy efficiency of the DSP systems is a point of concern. The fastest digital supercomputer to-date (Sunway TaihuLight - 2016) [7] has reached the computational power of a human brain, however, it dissipates a power equivalent to the output of a typical power
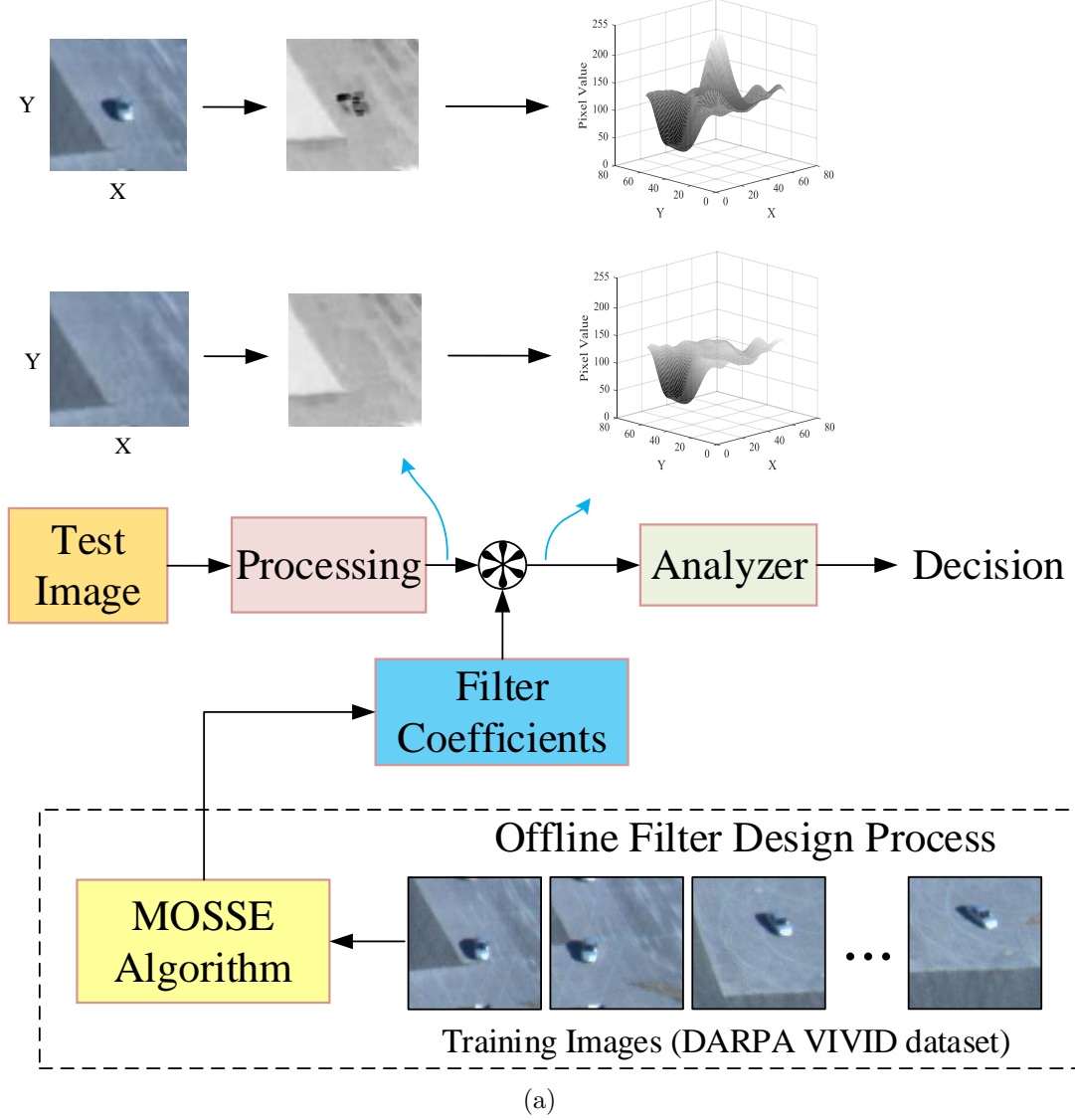
Figure 1.2: Object detection procedure using a CF. Filter coefficients are designed using the MSSOE algorithm and training images from the DARPA VIVID dataset. After obtaining the filter coefficient, the CF is ready to be tested. If a test image includes the same object, the correlation output exhibits a sharp peak, otherwise, the correlation output should not have any significant peak. Simulated responses of the CF implemented in MATLAB to presence and absence of a vehicle on a road is shown at the top of this figure.

station![1] In fact, most of the energy efficiency of the brain comes from the fact that the biological systems are analog. Unlike the digital world, where each wire only

---

[1]Sunway TaihuLight has a computational power of 93×10^{15} Ops./s at 15 MW power, while human brain has an estimated computational power of more than 10^{15} Ops./s at less than 10 W.

Table 1.1: Analog versus digital realization of multiplication and addition [39]

Realization	Addition	Multiplication
Digital	240 Transistors	3000 Transistors
Analog	A wire (KCL)	4-8 Transistors

represents one bit of information, a wire in the analog world can represent multiple bits. On the other hand, analog implementation can exploit basic laws of physics in order to perform mathematical operations. For instance, addition can be easily implemented using the Kirchhoff's Current Law (KCL) for current-mode signals, and a pair of MOSFETs biased in the sub-threshold region can perform the multiplication, thanks to the semiconductor device physics. Table 1.1 compares estimated hardware resources for the realization of basic mathematical operations (multiplication and addition) for two 8-bit numbers in analog and digital domains[39]. As a matter of fact, it has been shown that computational tasks that require low to moderate Signal-to-Noise Ratios (SNRs) are more efficiently realized in analog than digital in terms of area and power consumption [39].

1.3   Challenges in Analog Implementations

In spite of all the advantages mentioned above, the analog implementation of computational systems has certain issues. One of the most important problems with analog computing systems is the noise and offset accumulation which could result in a significant degradation of accuracy. A common way to compensate for the offset in such systems is to manually calibrate the biasing current of the analog memories [34, 11] or measure the offset for each output and store them in a separate array of analog memories and then subtract them from the output signals [5]. The former requires one analog memory per array element and the latter can only compensate for one of two inputs by fixing the other one. In both of these methods, the remaining offset depends on the programming accuracy.
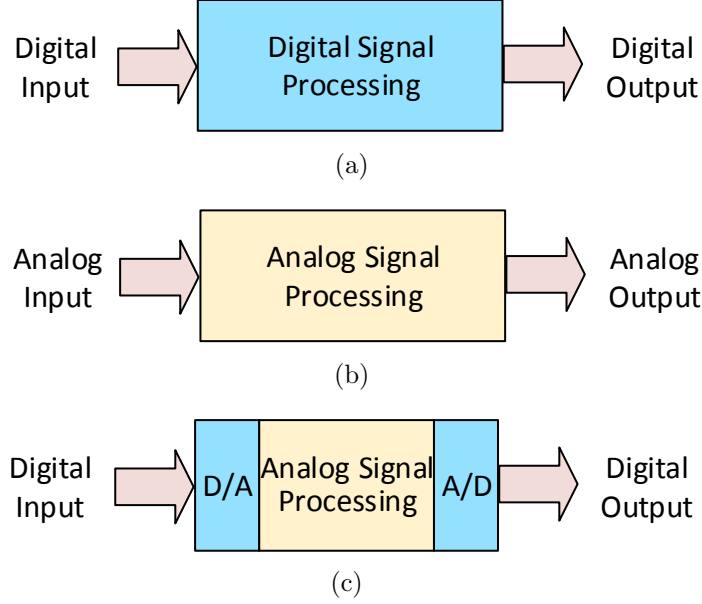
Figure 1.3: Signal processing methods: (a) Digital, (b) analog, and (c) analog with digital I/O interfacing.

Apart from the noise and offset-related issues, fully parallel implementation of analog processing operations requires input/output interface circuitry capable of supplying/acquiring a large number of analog data simultaneously to/from the computational block. Although implementing such interfaces is essentially a challenging task for any system, this becomes more of an issue when dealing with noise- and mismatch-sensitive analog signals. Therefore, it is a key to develop configurable digital interfaces that can easily scale with the number of inputs and efficiently communicate with other pre- and post-processing blocks.

The three signal processing methods discussed above are visualized in Figure 1.3. Figure 1.4 outlines the pros and cons of analog and digital implementation of I/O interfacing versus signal processing.

Several architectures implementing such hybrid/mixed-signal approaches have been reported in the literature. In [19] a bit-serial input/bit-parallel output architecture has been proposed that demands one flash ADC per output and needs further off-chip processing. Moreover, it utilizes DRAM memories which require

	I/O Interfacing	**Signal Processing**
Digital	✓ High-speed ✓ Noise-tolerant ✓ Easily scalable ✓ Standard communication	x Power hungry x Bulky design x Complex interconnection
Analog	x Low-speed x Noise-sensitive x Complicated scaling x Non-standard communication	✓ Energy-efficient ✓ Area-efficient ✓ Simple interconnection

Figure 1.4: Pros and cons of analog versus digital implementation of I/O interfacing and signal processing.

constant refreshing. The architecture proposed in [22] employs a large number of SRAM memories and other supporting digital logic circuits which add excessive power overhead to the system.

1.4   Research Goals

This dissertation investigates the implementation of a correlation filter system with analog signal processing and digital I/O interface. In summary, desired specifications of the system are:

- Energy and area efficiency.

  An 8-bit digital multiply-accumulate (MAC) circuit in a 130 nm CMOS process operating at a data-rate of higher than 1MHz consumes ∼5 pJ energy and occupies an active area of $80\mu m \times 80\mu m$. Adding data registers, decoders and other digital blocks for I/O interfacing, clock buffers, etc. which are needed to build a system such as the correlation filter, and considering the leakage power will significantly add to the energy consumption per MAC. In fact, a post-layout

evaluation of a 24×57 fully-digital CF system operating at 1.5 V supply and 50-500 MHz data-rate shows an energy-efficiency of ∼280 pJ/MAC. The total layout area was 6mm×4.75mm.

- High throughput to keep up with the 25-30 frames/sec produced by modern cameras (real-time processing requirement).

  Processing an 80×80 pixel frame with an 80-input CFS means running 137 vectors through the system. This means in order to have 30 frames/sec rate, the CFS should have a throughput of more than 4.11 kVector/s.

- Programmable analog non-volatile storage.

  Analog non-volatile floating-gate memories are capable of storing values even without power in contrast to SRAM/DRAM. Additionally, floating-gate memories occupy less area compared to digital storage, making it an attractive solution for storage in analog signal processing systems.

- 8-bit digital interface to communicate with pre- and post-processing blocks that:

  - Easily scales with the number of inputs and outputs.

  - Does not add significant power overhead to the system.(The entire system should maintain its energy-efficiency.)

- 8-bit input linearity

- The collective RMS error < %4

1.5   Original Contributions

In this work, two digitally-interfaced analog correlation filter systems for object tracking applications are proposed. The original contributions are summarized below:

- Proposed novel architecture and circuits to realize a Gilbert-multiplier-based correlation filter system with digital I/O interface, non-volatile storage, and techniques to reduce the offset and low-frequency noise of the analog array. The proposed system incorporates power management techniques to further increase the energy-efficiency of the system.

- Proposed novel architecture and circuits to realize a PWM-based correlation filter system with digital I/O interface, non-volatile storage, and techniques to reduce the offset and gain error of the analog array. The proposed system incorporates power management techniques to further increase the energy-efficiency of the system.

- Proposed a fast convergent calibration technique for dynamic comparators.

- Proposed a linearized pseudo-differential OTA with ± 800 mV linearity range at 1 V power supply.

1.6 Dissertation Organization

The remaining chapters of this dissertation will cover the design of circuits and architectures to implement the CFS systems described above.

Chapter 2 describes the design and presents the experimental results of a Gilbert-multiplier-based CFS. In chapter 3 development of a PWM-based CFS is presented as an alternative architecture which addresses limitations of the Gilbert-multiplier-based CFS and achieves better performance in terms of energy-efficiency and operating speed. Chapter 4 concludes the dissertation and proposes potential future works.

Chapter 2

A Gilbert-Multiplier-Based Correlation Filter System

An energy-efficient digital I/O interface solution for an analog correlation operator for linear filtering is presented which maintains the power and area efficiency of the entire system and easily scales with the number of inputs. Furthermore, the proposed system utilizes non-volatile floating-gate memories as storage devices, eliminating the need for DRAM/SRAM memories. Also, a correlated double sampling (CDS) technique has been implemented to cancel offset and to reduce the low-frequency noise at the outputs of the array.

2.1 System Description

The correlation of two vectors $w(m)$ and $x(k)$ is defined as:

$$y(n) = \sum_{i=-M/2}^{M/2} w(i) \cdot x(n - i) \tag{2.1}$$

In this equation $w(m)$ is the filter coefficient vector, where $-M/2 \leq m \leq M/2$, $x(k)$ is the input vector, where $-M/2 \leq k \leq N - 1 + M/2$ and $y(n)$ is the output vector,

9
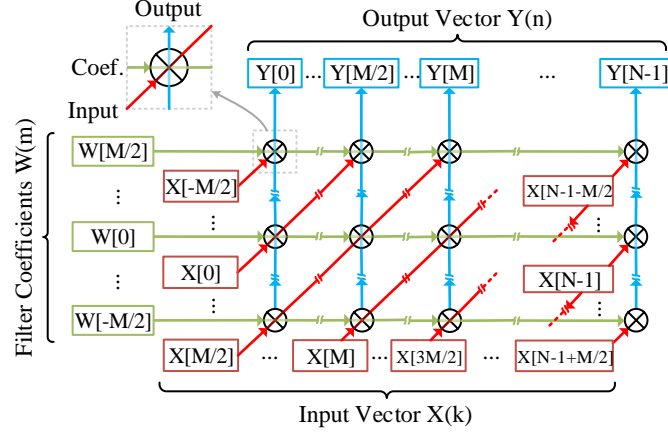
Figure 2.1: An M×N fully parallel analog correlation filter realized with four-quadrant multipliers and floating-gate memories. The multipliers at each row share the same weight inputs (green lines) and diagonal multipliers share the same signal inputs (red lines). The multiplier outputs at each column are wire-summed to yield a single pixel (blue lines).

where $0 \leq n \leq N - 1$. It should be noted that in an array with M coefficients and N outputs, the input vector has $M - 1$ more elements than the outputs. Figure 2.1 illustrates a fully parallel implementation of a M×N correlation filter.

Figure 2.2: The proposed architecture for a 24×57 fully parallel correlation filter system with digital I/O interface.

10

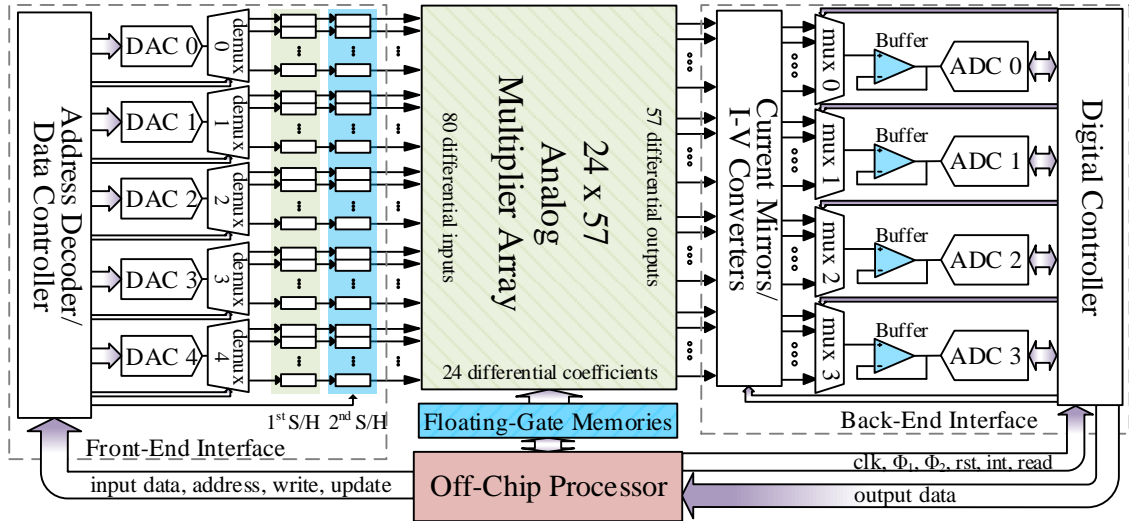The proposed architecture for a digitally interfaced fully parallel analog correlation filter system (CFS) is shown in Figure 2.2. The filter coefficient vector is stored in an array of analog floating-gate memories (FGMs) which are connected to voltage inputs of multipliers in each row. In other words, the FGMs are shared across N multipliers. A front-end interface converts the digital input vector to analog current vector and delivers them to the multiplier array simultaneously. Thanks to the utilization of current-mode signals, the summation in (2.1) is performed by simply connecting the output of multipliers in each column together according to Kirchoff's current law (KCL). After computation, a back-end interface converts the output vector back to digital. In this paper, we present a prototype CFS with $M = 24$, $N = 57$, and 80 inputs.

2.2 The Front-End Interface

The front-end interface comprising five digital-to-analog converters (DACs) followed by eighty sample-and-hold (SH) circuits. Every 16-channels share one DAC using a time-domain multiplexing (TDM) scheme. It should be noted that an 8-bit address space is chosen for input data to account for future growth. The lower 4 bits choose the DAC and the upper 4 bits select the SH channel. Figure 2.3 shows the timing diagram of the front-end interface.
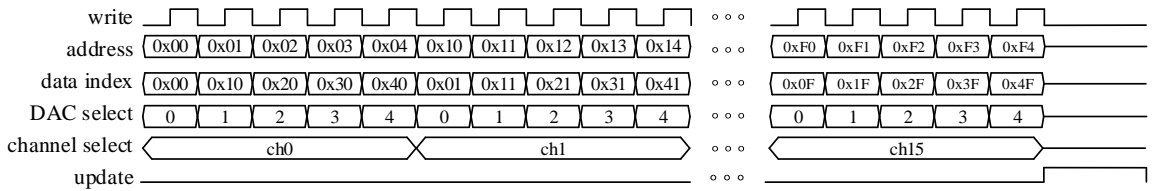

Figure 2.3: The front-end timing diagram: In the write phase 80 8-bit data samples are sent to the chip. The five DACs operating in TDM scheme convert digital data to analog and store them in the SHs.

(a)

(b)
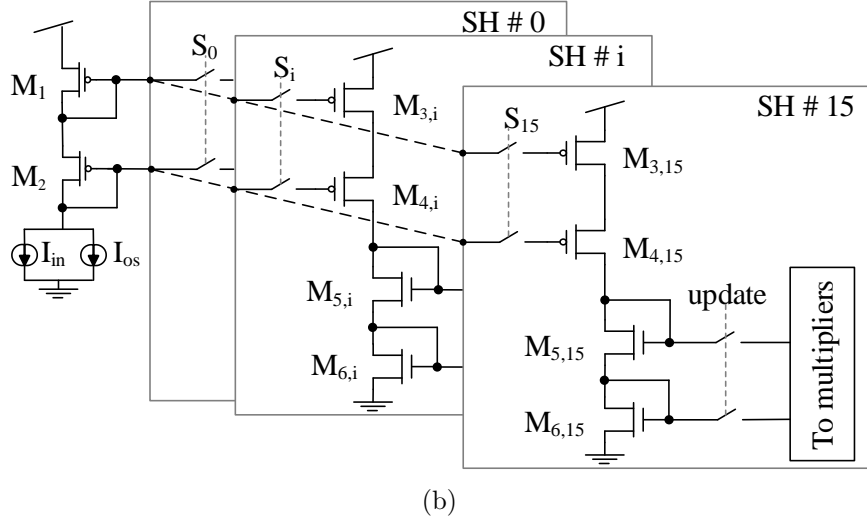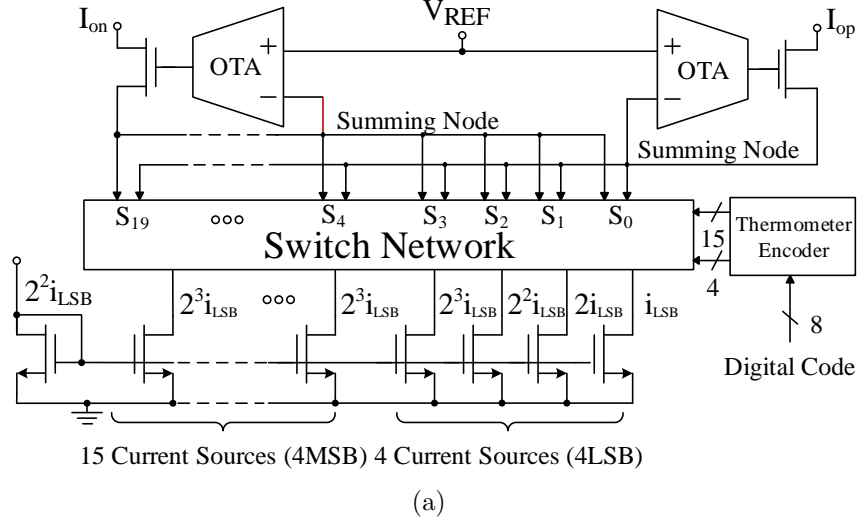
Figure 2.4: The front-end interface: (a) the segmented current-steering DAC with differential outputs, and (b) the current-mode demultiplexer/SHs. The switches S_0-S_{15} are controlled by the channel select signal shown in the timing diagram of Figure 2.3.

2.2.1 Digital-to-Analog Converter

An 8-bit current-steering DAC was designed to convert the digital input signals to analog signals for the analog computation block. As shown in Figure 2.4(a), the DAC uses segmented topology: 4 MSBs are thermometer coded, and the 4 LSBs are binary weighted. The segmentation helps to reduce differential non-linearity (DNL) and integral non-linearity (INL). An operational transconductance amplifier (OTA)

12

is used in a cascode current sink configuration to increase the output resistance. A detailed description of this design can be found in [33].

2.2.2 Current-Mode Demultiplexer/SH

Figure 2.4(b) shows the schematic of the current-mode demultiplexer/SH circuit. Transistors M_1 and M_2 are diode-connected and have a sink current directly from the DAC output (I_{in}). An externally-controlled DC offset current (I_{os}) was added in order to decrease the time-constant of input stage especially for small input currents. The gate connections of M_1 and M_2 are shared with 16 SHs through switches controlled by $S_i, i = 0, 1, ..15$. When a set of switches is turned on, a current mirror is formed between transistors $M_1 - M_2$ and the transistors on the other side of the given switch, $M_{3,i} - M_{4,i}$. The sampled voltages are held on the gate capacitance of $M_{3,i} - M_{4,i}$ when their switches are turned off. The second SH is formed between $M_{5,i} - M_{6,i}$ and the multiplier tail transistors. By activating *update* signal, input currents held in the first SHs are passed to the multiplier array to start the computation. In order to minimize the charge injection error dummy switches have been used whose gates are driven by inverted clock signals.

2.3 Analog Multiplier Array

The four-quadrant Gilbert multiplier circuit is shown in Figure 2.5(a). The multiplier is formed with NMOS transistors operating in the sub-threshold region. The multiplier linear region was extended using voltage-controlled degeneration technique [23]. Using the equation provided in [45], the differential transconductance can be written as:

$$G_{m0} \approx \frac{I_{in}}{nU_T} \cdot \frac{2}{L + 4} \tag{2.2}$$

where I_{in} is the input current, n is the sub-threshold slope factor, U_T is the thermal voltage, and L is the ratio of transconductance parameters of M_i and $M_{i,a}$ i.e. $\beta_i/\beta_{i,a}$,
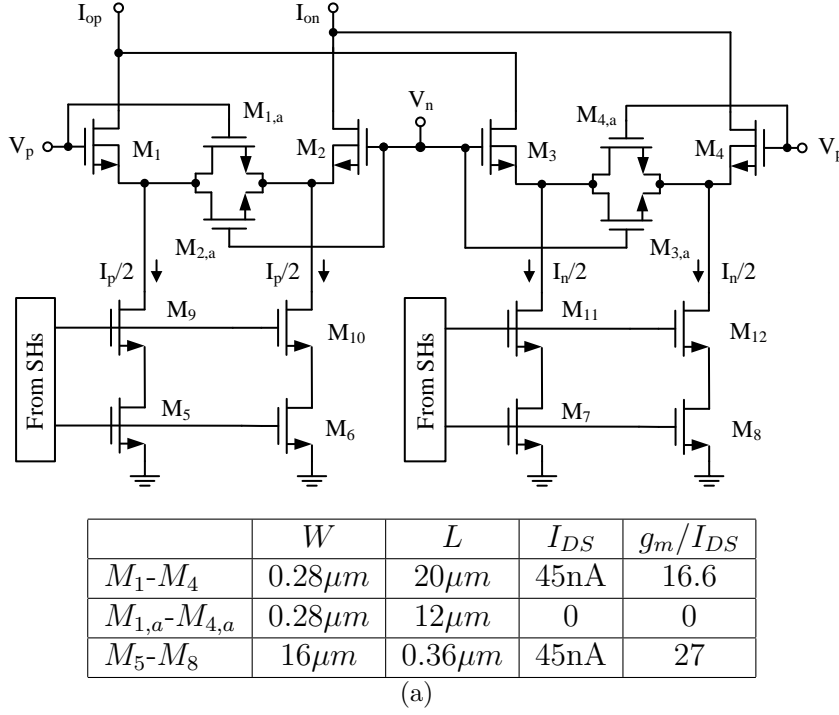
	W	L	I_{DS}	g_m/I_{DS}
M_1-M_4	$0.28\mu m$	$20\mu m$	45nA	16.6
$M_{1,a}$-$M_{4,a}$	$0.28\mu m$	$12\mu m$	0	0
M_5-M_8	$16\mu m$	$0.36\mu m$	45nA	27

(a)

Figure 2.5: The linearized four-quadrant multiplier circuit. Transistors W, L, I_{DS} and g_m/I_{DS} (for $V_p = V_n$) are given in the table.

$i = 1, 2, 3, 4$. The linearity was found to be maximum for $L \approx 2.5$. Transistors M_i were chosen to be *triple-well* devices to eliminate the body effect and hence, to improve multiplier linearity with respect to the current input. Moreover, isolation from the substrate improves the noise performance. It should be noted that achieving an 8-bit dynamic range and tolerable mismatch-related errors come at the expense of using long-channel devices, with the result that each multiplier cell occupies an area of $30\mu m \times 40\mu m$. Nevertheless, this design is markedly smaller than a digital counterpart.

2.3.1 Power and Speed Performance

The input stage shown in Figure 2.6(a) dominates the frequency response of the CFS. From a small-signal analysis of the self-biased cascode current mirror, the

14
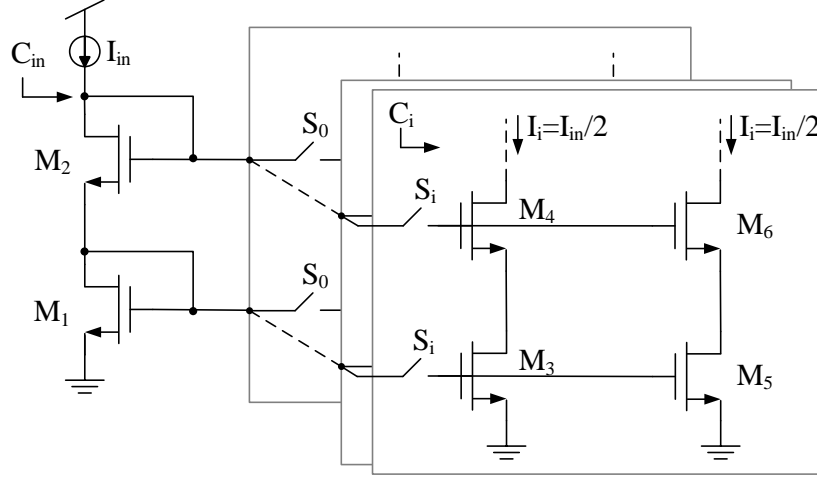
(a)
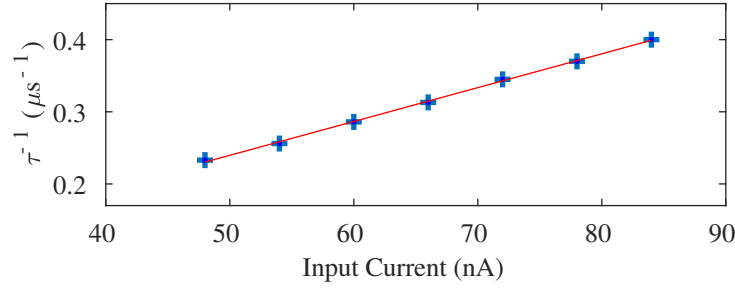
(b)

Figure 2.6: (a) The input stage of the multiplier array. The switches S_0-S_i are controlled by the update signal. (b) The inverse of multiplier array time constant varies linearly with input current as predicted by (2.5).

dominant pole is:

$$P = \frac{-g_{m1}}{C_{in}} \qquad (2.3)$$

where g_{m1} is the transconductance of the transistor M_1 and

$$C_{in} = (j+1)C_i \qquad (2.4)$$

and j is the number of multipliers connected to the input node and the worst case is when the input node is connected to multipliers in all of the rows ($j = M$). By substituting the sub-threshold equation for g_{m1}, the inverse time constant can be

15

written as:

$$\tau^{-1} = \frac{I_{in}}{nU_T\,C_{in}}.$$ (2.5)

This equation shows that the inverse time constant linearly scales with the input current level and the number of rows in the array. Figure 2.6(b) plots the measured inverse time constant for a given input current level. Total power consumption of the $N \times M$ multiplier array is:

$$P_{MA} = 2\,NMI_{in}V_{dd}.$$ (2.6)

Assuming a settling time of 5τ for the array, the power-delay product is then:

$$P_{MA}(5\tau) \approx 10\,NM^2nU_T\,C_iV_{dd}.$$ (2.7)

This equation shows that the power-delay product for multiplier array is a linear function of N and a quadratic function of M. It is also independent of input current, suggesting that operating at higher speed by increasing the input current level does not reduce the energy efficiency. However, increasing the input current level increases the non-linearity because transistors start moving out of the sub-threshold region.

2.3.2 Noise Performance

In the multiplier circuit shown in Figure 2.5(a) the shot noise is the dominant noise source due to the large size of the devices. The noise spectral density of transistors in sub-threshold region is given by [38]:

$$\overline{i^2} = 2\,qI_s(1 + \exp(-V_{DS}/U_T))$$ (2.8)

where I_s is the sub-threshold saturation current. Considering that transistors $M_i(i = 1, 2, ..., 8)$ operate in the saturation region ($V_{DS} >> U_T$), their noise spectral density

16

can be approximated by:

$$\overline{i_{M_i}^2} = 2qI_{s,i} = qI_{in}, \quad i = 1, 2, ..., 8 \tag{2.9}$$

where I_{in} is the SH input current. Transistors $M_{i,a}$, on the other hand, operate in the triode region and according to (2.8) their noise contribution reach their maximum of $4qI_{s_{i,a}}$ when $V_{DS_{i,a}} = V_p - V_n = 0$. However, because $I_{s_{i,a}}/I_{s_i} = \beta_{i,a}/\beta_i = 0.4$ ($i = 1, 2, 3, 4$), therefore $I_{s_{i,a}} = 0.2I_{in}$ and the noise spectral density of $M_{i,a}$ can be written as:

$$\overline{i_{M_{i,a}}^2} = 4qI_{s_{i,a}} = 0.8qI_{in}, \quad i = 1, 2, 3, 4. \tag{2.10}$$

If we consider the current noise of the tail transistor M_5, as shown in Figure 2.7(a), the current gain from the node A to the the positive and negative outputs can be written as:

$$A_p = \frac{L}{L+1} = 0.71, \qquad A_n = \frac{1}{L+1} = 0.29. \tag{2.11}$$

Therefore, the noise contribution of M_5 to the output is:

$$(A_p - A_n)i_{n5} = 0.42i_{n5}. \tag{2.12}$$

As depicted in Figure 2.7(b), the current noise of transistor M_1 can split into two correlated noise sources with the same values, one at the drain of M_1 and the other at the node A. Thus, the noise contribution of M_1 to the output is:

$$(1 - A_p + A_n)i_{n1} = 0.58i_{n1}. \tag{2.13}$$

And finally, the current noise of transistor $M_{1,a}$ can split into two noise sources at node A and node B as shown in Fig 2.7(c). Hence, the noise contribution of $M_{1,a}$ to the output is:

$$2(A_n - A_p)i_{n1a} = 0.84i_{n1a}. \tag{2.14}$$

17
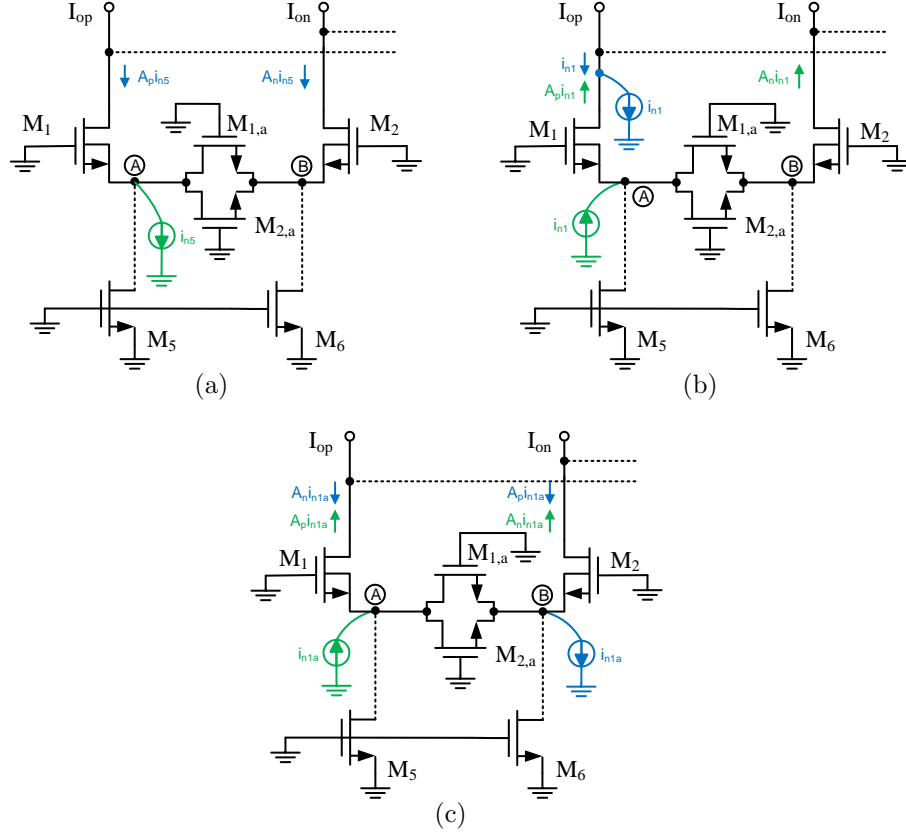
Figure 2.7: The simplified schematic of the multiplier circuit for noise analysis (the cascode transistors are not shown). The Noise contribution of (a) M_5, (b) M_1, and (c) $M_{1,a}$.

Consequently, the noise contribution of all the transistors to the output can be calculated as:

$$
\begin{aligned}
\overline{i_{out}^2} &= (0.58)^2 \sum_{i=1}^{4} \overline{i_{M_i}^2} \\
&+ (0.84)^2 \sum_{i=1}^{4} \overline{i_{M_{i,a}}^2} + (0.42)^2 \sum_{i=5}^{8} \overline{i_{M_i}^2} \\
&= 1.34\,\overline{i_{M_i}^2} + 2.82\,\overline{i_{M_{i,a}}^2} + 0.7\,\overline{i_{M_i}^2}.
\end{aligned}
\tag{2.15}
$$

Hence, the total output noise is given by:

$$
\overline{i_{out}^2} = (1.34\,\overline{i_{M_i}^2} + 2.82\,\overline{i_{M_{i,a}}^2} + 0.7\,\overline{i_{M_i}^2})\,\Delta f = 4.3\,qI_{in}\,\Delta f.
\tag{2.16}
$$

18

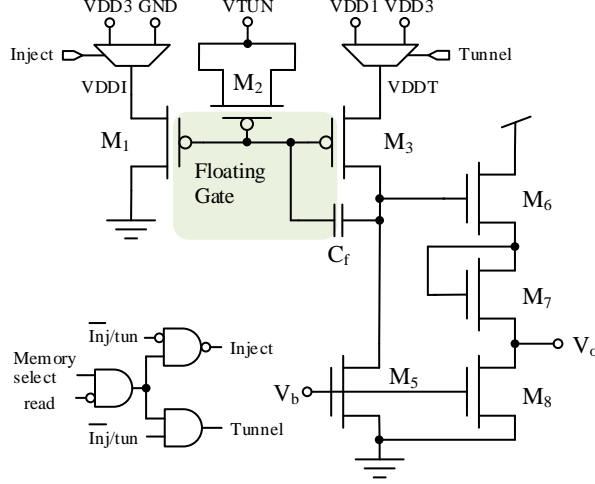Figure 2.8: The floating-gate memory cell and digital control logic.

Substituting the equivalent noise bandwidth (ENB) of $1/4\tau$ in the above equation we obtain:

$$\overline{i_{out}^2} \approx \frac{4.3\, qI_{in}^2}{4nU_T\,(M+1)C_i}. \tag{2.17}$$

Thus, the RMS signal-to-noise ratio (SNR) can be written as:

$$SNR = \frac{I_{od}}{\overline{i_{out}}} = \frac{G_{m0} \cdot V_{id}}{\overline{i_{out}}} \approx \frac{4V_{id}}{\sqrt{4.3}(L+4)}\sqrt{\frac{(M+1)C_i}{n\,kT}} \tag{2.18}$$

where k is the Boltzmann constant and T is the absolute temperature. Based on this equation, increasing the input capacitance leads to higher SNR, which is indeed nothing but the well-known trade-off between the bandwidth and SNR. From (2.5) and Figure 2.6(b) the C_i is estimated to be 325 fF. With a $V_{id} = 0.1$ V, $kT = 4.11 \times 10^{-21}$ J, and $n = 1.42$ from simulation, expected SNR is 60.8 dB.

2.4 Nonvolatile Floating-Gate Memories

An array of floating-gate (FG) memories is employed to store the analog filter coefficients in a differential mode. The schematic of the FG analog memory cell is shown in Figure 2.8 [26]. The gate of M_1, M_2 and M_3 and the top plate of capacitor

Table 2.1: Control Signals for Different Operation Modes

Control Signal	Injection	Tunneling	Read	VDDT	VDDI
VTUN	L	H	L	1	0
$\overline{inj}/tun$	L	H	L	3	3
read	L	L	H	3	0

form the FG. The stored charge on the FG is modified by the injection process through M_1 and the tunneling process through the transistor M_2. Tunneling removes electrons from the FG node while injection adds electrons. Both of these processes change the amount of charges stored in C_f, therefore, change the output voltage according to the equation:

$$V_{out} = \frac{\Delta Q}{C_f}. \tag{2.19}$$

In the tunneling mode, VTUN is connected to 7 V, and VDDT is switched from 3 V to 1 V to reduce the FG voltage and increase the gate oxide voltage, V_{ox}. The amount of charge added or removed from the FG is controlled by the pulse width of VDDI and VTUN signals, respectively. In the injection mode, VDDI is switched to 3 V from GND, VTUN is switched to GND and VDDT is switched back to 3 V to prevent tunneling. In the read mode, VTUN and VDDI are switched to 0 and VDDT is switched to 3 V to make sure no tunneling or injection is happening. Upon activation of 'read' signal, the output of the selected cell is connected to a pad and read by off-chip read-out circuitry. Because the programming process utilizes feedback based on V_o, non-linearities, finite-gain effects, etc. are accounted for. Table 2.1 summarizes the FGM operation modes and the control signals. In general, the programming time depends on the number of floating-gate memories and the target values. For this prototype, it takes about one minute in average to program all the floating gates. The RMS error between target and actual values for all of the floating-gate memories was less than 1mV. This was calculated based on the errors measured at the V_o node.
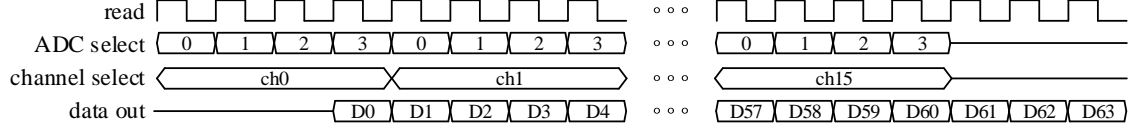
Figure 2.9: The back-end timing diagram: In the read phase 64 analog samples are converted to 8-bit digital data and sent out of the chip using four ADCs operating in TDM scheme.

2.5   The Back-End Interface

An array of 57 current mirrors performs differential to single-ended conversion. The difference currents are then integrated into the capacitors. The analog output voltages are multiplexed into four unity-gain buffers driving four 8-bit SAR ADCs. The timing diagram of Figure 2.9 illustrates the sequence in which ADCs perform the conversion and output digital data.

2.5.1   Current mirrors and I-V Converters

Figure 2.10 shows the schematic of current mirror followed by the I-V converter circuit. An OTA keeps the voltage on the I_{op} node at V_{REF} to improve the accuracy. Using (2.1) and (2.2) the output voltage of the integrator can be written as:

$$V_o(n) \approx \frac{2\,T_{int}}{nU_T C_{int}(L+4)} \sum_{i=-M/2}^{M/2} V_{id}(i) \cdot I_{id}(n-i) \tag{2.20}$$

Where C_{int} is the integration capacitor and T_{int} is the integration time. The CDS technique was implemented on the CFS chip to cancel offset at the array outputs. The CDS is performed in two phases. In the first phase, the off-chip processor writes *reference* input, then updates the analog multipliers and integrates the output current. In the second phase, the processor repeats the same process, this time for *data* input. The data input is subtracted from reference input by switching plates of C_{int} between the two phases.
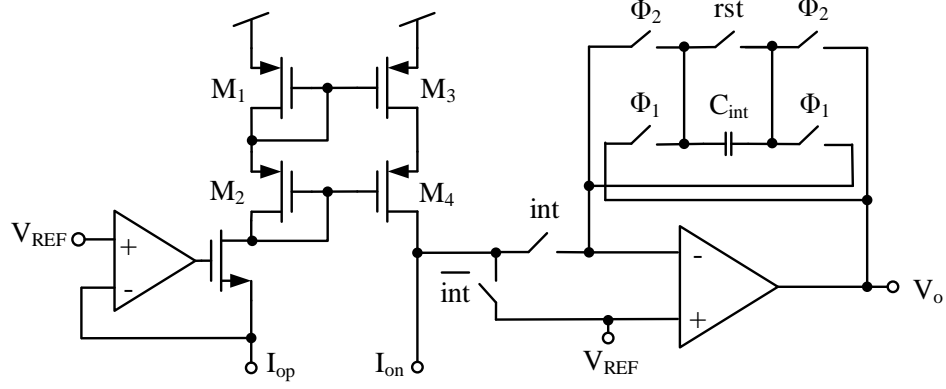
21

Figure 2.10: The current-mirror and the I-V converter.

A correlated double sampling (CDS) offset cancellation technique is implemented [17]. As shown in the Figure 2.11, offset voltage V_{os} and noise voltage V_n are added to the output of I-V converter at any sampling time t. If we apply *reference* input to the system at time t_1, the output signal can be written as:

$$V_o[t_1] = V_{o,ref}[t_1] + V_{os}[t_1] + V_n[t_1] \tag{2.21}$$

Then we apply *data* input to the system at sampling time t_2:

$$V_o[t_2] = V_{o,data}[t_2] + V_{os}[t_2] + V_n[t_2] \tag{2.22}$$

If we subtract these two output voltages and note that $V_{o,ref}$ indeed is the output common-mode voltage with no signal component, we get:

$$
\begin{aligned}
V_{o,CDS} &= V_o[t_2] - V_o[t_1] \\
&= V_{o,data} + (V_n[t_2] - V_n[t_1])
\end{aligned} \tag{2.23}
$$

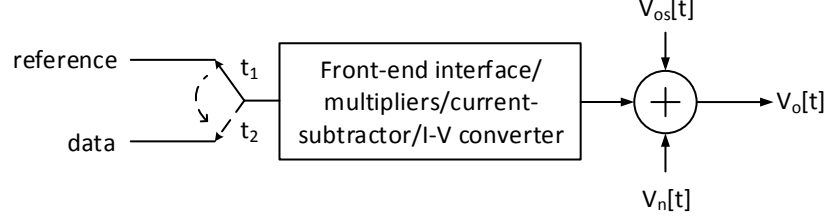if we define $f_s = \frac{1}{t2-t1}$, we can write 2.23 in s-domain as:

22

Figure 2.11: Correlated double sampling technique

$$V_{o,CDS} = V_{o,data} - V_n(\frac{2s}{s + 2f_s}) \tag{2.24}$$

This equation shows that the remaining noise spectrum is shaped by a high-pass filter with a corner frequency at $2f_s$. Thus, CDS not only removes the DC offset voltages but also reduces the low-frequency noise.

Figure 2.12 shows how CDS is implemented in two phases on CF system. In the first phase, Microprocessor writes 'reference' input, then updates the analog multipliers and integrates the output current. This is indeed the first sampling and the resultant voltage is equal to $V_o[t_1]$ from equation 2.21.

In the second phase, Microprocessor repeats the same process, this time for 'data' input (second sampling) and the resultant voltage is equal to $V_o[t_2]$ from equation 2.22.
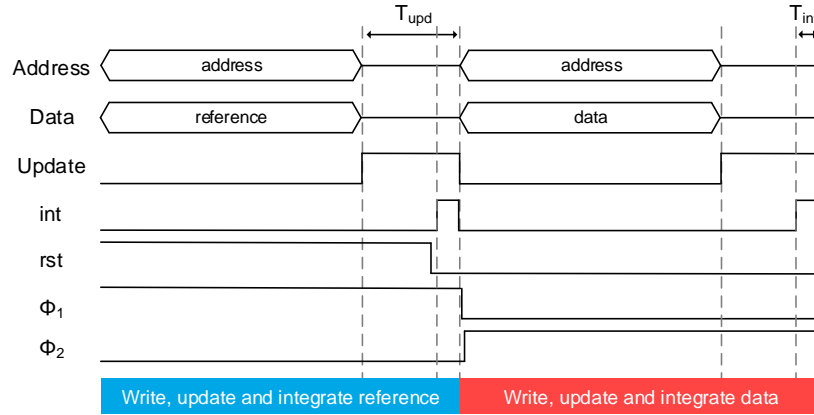

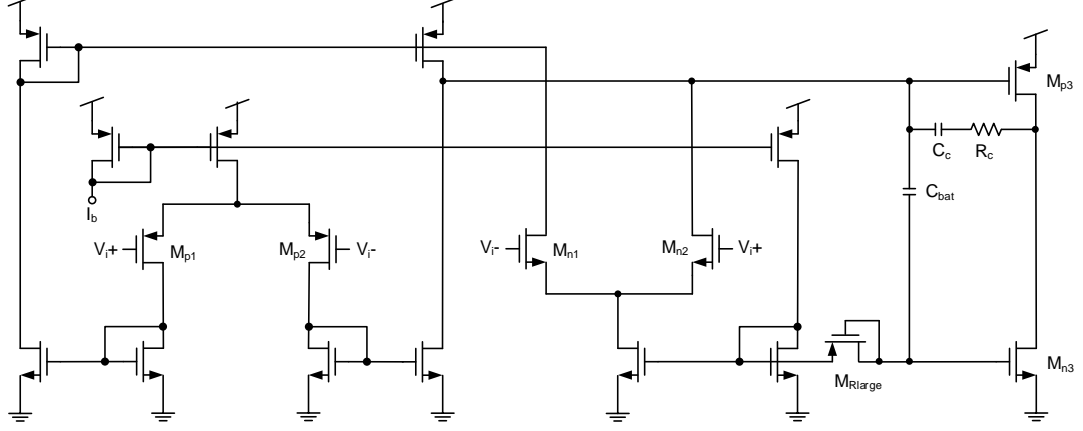Figure 2.12: Timing diagram for CDS implementation on CFS.

23

Figure 2.13: The complementary input two-stage class-AB op-amp circuit schematic

The subtraction in equation 2.23 is simply implemented by switching plates of C_{int}. Two non-overlapping signals ϕ_1 and ϕ_2 control switches S_{4-5} and S_{6-7} respectively. During the first phase, ϕ_1, top plate of the C_{int} is connected to S_1, and the bottom plate is connected to the output. In the beginning of the second phase, ϕ_2, the plates are switched such that the bottom plate is connected to S_1, and the top plate is connected to the output. This process conserves the charge stored in the capacitor and only changes its polarity.

The operational amplifier (op-amp) used for I-V conversion (Figure 2.10) is a class-AB two-stage op-amp designed based on the circuit presented in [35]. The circuit schematic is shown in Figure 2.13.[1]. Compared to [35], the input stage was modified to a complementary stage to extend the input voltage range. The circuit is basically a conventional two-stage class-A op-amp with two additional elements: C_{bat} and M_{Rlarge}. During the quiescent conditions, the circuit operates as a conventional two-stage class-A op-amp because no DC current flows through M_{Rlarge}, however, during dynamic operation, C_{bat} operates as a floating battery and passes the voltage variations to the gate of M_{n3} and enables the class-AB operation with no additional static power dissipation.

---

[1]Design credit goes to Dr. Tan Yang.

24

The 57 analog voltage outputs were multiplexed to four channels. Each channel was buffered and connected to an ADC. The buffered analog voltages were also connected to PADs through a second set of analog buffers for testing and characterization.

2.5.2 Analog Buffers

The circuit shown in Figure 2.13 has been used as a unity gain buffer to drive the ADC input capacitor array during the sampling phase which is as large as $1.66pF$. The class-AB operation helps to reduce power consumption. Based on the simulation results, buffer requires $1\mu A$ of biasing current to settle to half LSB during the ADC sampling time ($312.5ns$ at $24MHz$ ADC clock frequency).
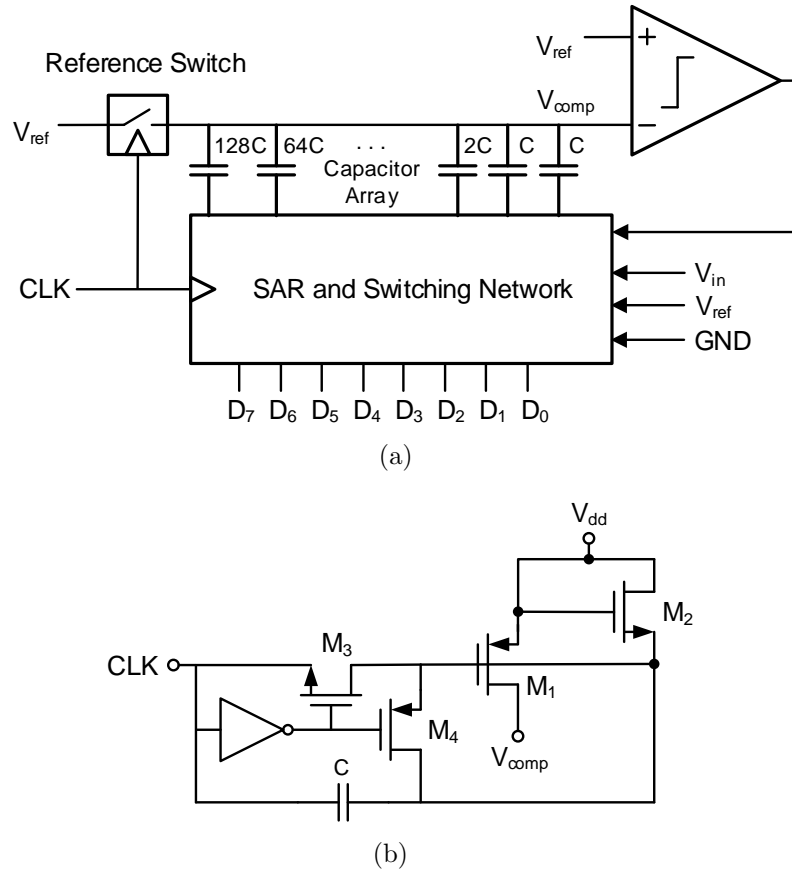

(a)

(b)

Figure 2.14: (a) SAR ADC architecture, and (b) Reference switch implemented using a charge pump circuit.
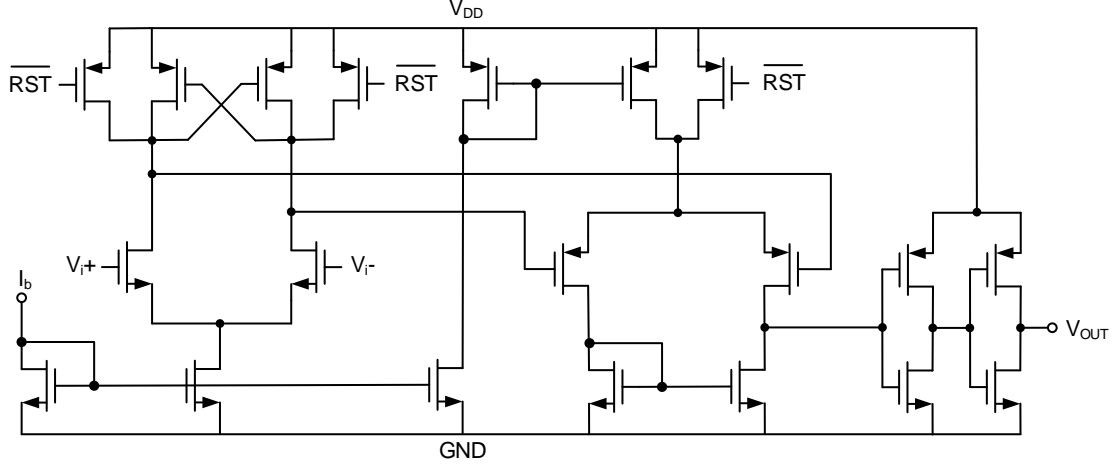
Figure 2.15: Comparator circuit schematic designed for the ADC.

2.5.3 Analog-to-Digital Converter

The SAR ADC is a popular choice for medium resolution/speed applications because of its energy efficiency advantage. A conventional 8-bit SAR ADC was designed[2] to convert analog output vector to digital output vector (Figure 2.14(a)). The 6.5 fF unit capacitors were realized using VNCAP devices and were arranged as an array with a common-centroid configuration to decrease mismatch effects. A single-cycle charge pump circuit [41] was used to boost the gate voltage of the reference switch which made it possible to use the supply voltage as the reference voltage. The comparator circuit (Figure 2.15) comprises a regenerative circuit followed by a differential to single-ended amplifier stage.

2.6 Experimental Results

A prototype 24×57 CFS was designed and fabricated in a 0.13 μm CMOS process. The chip area is 1.7 mm×1.9 mm. Figure 2.16 depicts annotated chip micrograph. The CFS chip was evaluated using a custom test board interfaced to a PC via an FPGA board. The FPGA board controls the write, update, integrate and read timings

---

[2]Design credit goes to Peixing Liu and Dr. Charles Britton.

26
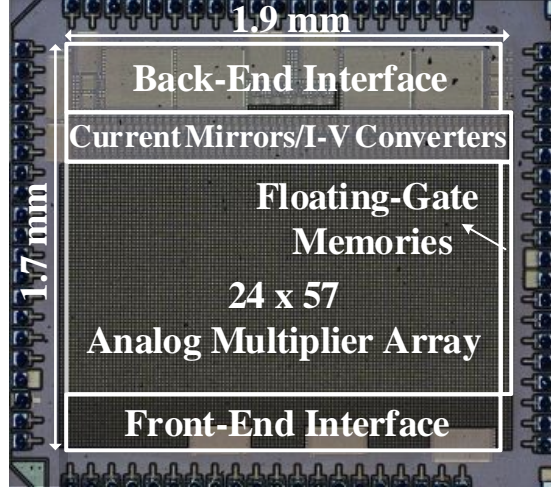
Figure 2.16: The micrograph of the CFS chip fabricated in a 0.13 μm CMOS process.

and communicates with the PC through a USB-UART interface. Figure 2.17 shows the input/output characteristics for a column of multipliers. The output is taken after the charge integrator, which converts the current output to a voltage. The measured worst-case INL and DNL for the inputs were $+4.7/-5.2$ and $+1.8/-1$ 8-bit LSBs, respectively, whereas those of the weights were $+1/-1.2$ and $+0.26/-0.27$ 6-bit LSBs, respectively.
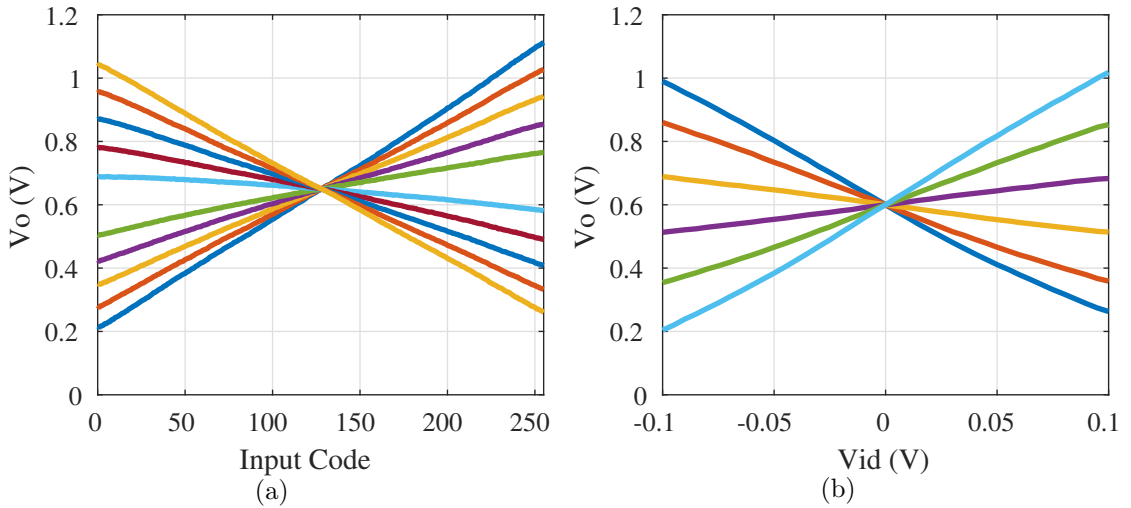
Figure 2.17: The multiplier array characteristics: (a) The output voltage vs. input code for various programmed V_{id} values, and (b) the output voltage vs. V_{id} for various input codes.
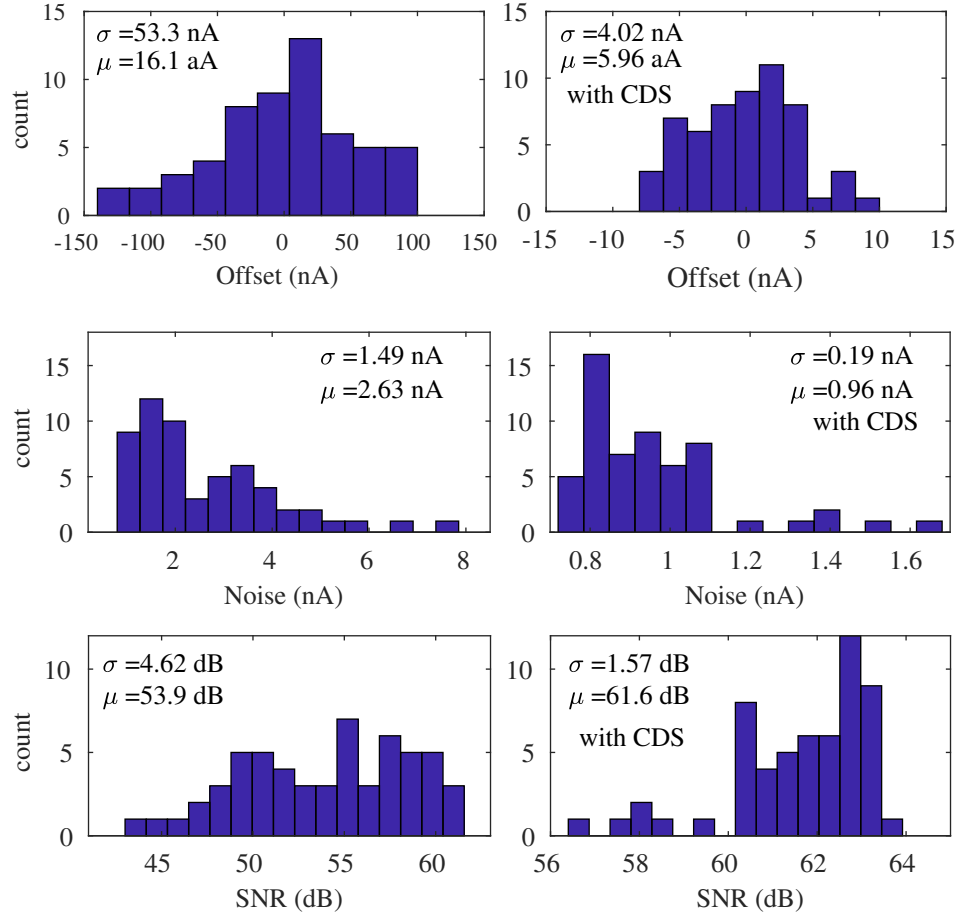
Figure 2.18: The measured input-referred offset, noise, and SNR histograms without and with the CDS.

The input-referred offset, current noise, and SNR with and without performing the CDS offset cancellation technique are depicted in Figure 2.18. Since the measured values are the sum of several independent random processes, a Gaussian distribution is assumed here for calculating the mean and the standard deviation. If the number of samples (i.e. outputs) were sufficiently large the Gaussian distribution would be more evident (the Kolmogorov-Smirnov test also supports the Gaussian assumption). Implementing the CDS technique reduced the offset from 53.3 nA to 4 nA (13.2X reduction). The SNR without the CDS is 53.9 dB with a standard deviation of 4.62 dB which is in agreement with the expected SNR of 60.8 dB. The CDS also

Figure 2.19: The saved power due to the power gating (left) and the power distribution of the CFS chip (right).

reduced low-frequency noise (by a factor of 2.7) and improved the average SNR from 53.9 dB to 61.6 dB.

Operating at the maximum sampling rate of $600kSps$ each ADC consumes an average of $10.1\mu W$ of power. The INL and DNL of the ADCs were measured by applying a slow moving ramp signal to the positive input of all the I-V converters while keeping th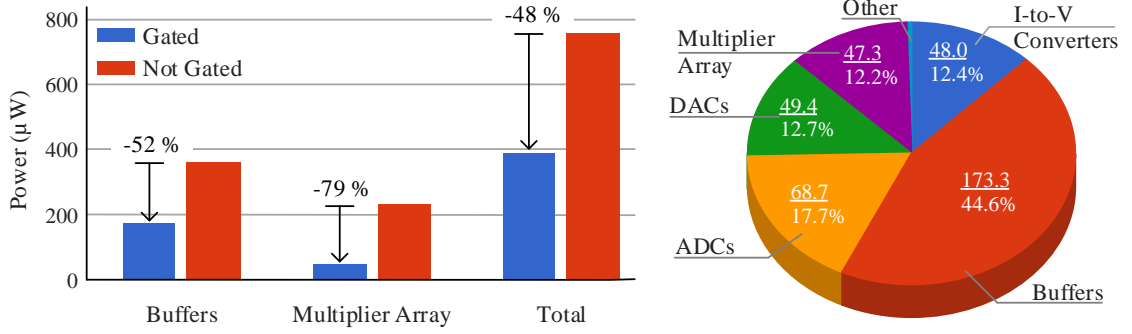e rst switch on. The INL ranges from -2 to 2 LSBs and DNL ranges from -1 to 1.14 LSBs. The ADCs show an SFDR of $45dB$ and an effective number of bits (ENOB) of 6.5 bits.

The comparator was designed to work with $1.5V$ power supply and $150nA$ of bias current with a conversion time of $10ns$ and $5.04\mu W$ power dissipation. However, a mismatch problem caused conversion failure at some input values. To fix that, the biasing current was increased to $1.1\mu A$ and V_{dd} to $1.7V$. This allows for conversion time of $2.8ns$ with the cost of increased power dissipation to $7.55\mu W$.

Table 2.2 summarizes the specifications of the system. Operating at 6 MHz write speed and 2.4 MHz read speed, the entire system achieves 25.2 pJ/MAC of energy efficiency at 11.3 kVec/sec throughput. The multiplier array and the unity-gain buffers were the most power-hungry blocks in the system and therefore they were turned off during the standby intervals. As a result, as shown in Figure 2.19, 48 percent of the power was saved. Table 2.3 compares key specifications of this work with similar systems. The 8-bit CFS chip presented in this paper uses non-volatile

Table 2.2: Summary of Specifications: Gilbert-Multiplier-Based CFS

	Specification	Value
Front-End Interface	Write Speed	6 MHz
	D/A INL/DNL	-0.8/0.54 LSB
	Power Dissipation	49.4 μW
Multipliers	Power Dissipation	47.3 μW
Back-End Interface	Read Speed	2.4 MHz
	A/D SFDR/ENOB	45 dB/6.5 bits
	Power Dissipation	
	OTAs	1.35 μW
	I-V Converters	48 μW
	Buffers	173.3 μW
	A/Ds	68.7 μW
System	Power Supply	1.5 V
	Throughput	11.3 kVec/sec
	SNR	61.6 dB
	Power Dissipation	388.4 μW

floating-gate memories to store filter coefficients and achieves better total energy efficiency compared to other systems reported with similar functionality.

The CFS chip was tested by filter-input vector pairs and the resultant vector was compared to an ideal filtering performed in Matlab. Figure 2.20 shows the results of two experiments. Note that all the values are normalized to ±1. Top plot of Figure 2.20(a) shows the input vector which contains an abrupt transition from -1 to +1 in the middle. As shown in the second plot, the filter coefficients were

Table 2.3: Performance Comparison: Gilbert-Multiplier-Based CFS

| Specification	This work	TCAS II [19]	VLSIC [22]	CICC [12]	VLSIT[34]
Process (μm)	0.13	0.5	0.35	0.5	0.35
Purpose	CF	VMM	Image Filter	VMM	VMM
I/O type	Digital	Digital	Digital	Analog	Analog
No. of Bits	8	1	6	-	-
Memory Type	FG	DRAM	SRAM	FG	FG
Array Size	24×57	512×128	51×80	128×32	100×10
Chip Size (mm^2)	3.23	9	9.8	0.82	25
Settling Time*(μs)	12.5	10	1	0.08	13
Core Energy*(pJ/MAC)	0.43	0.5	22	0.57	0.32
Total Energy (pJ/MAC)	25.2	N/A	68.6	-	-

* Measured for the analog array.
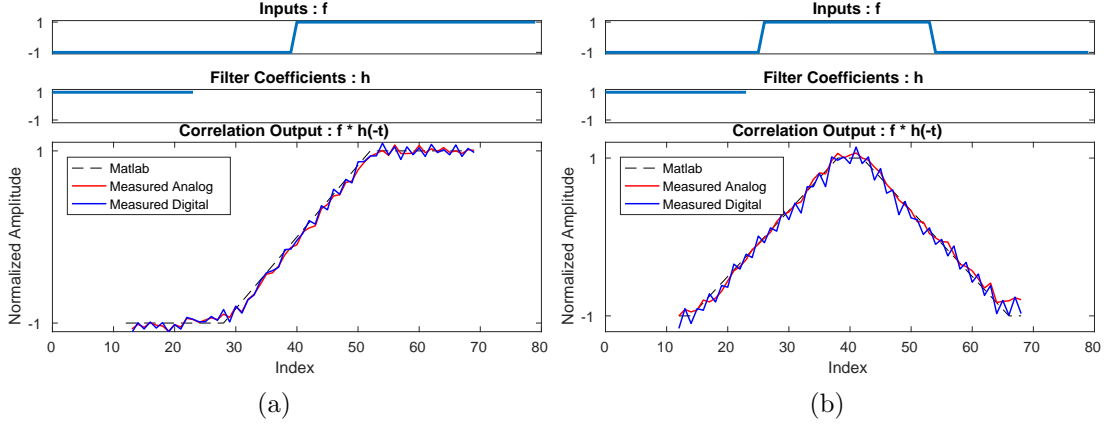
Figure 2.20: Measurement results for correlation between two vectors, coefficients resemble a smoothing filter (a) input signal has one abrupt change (b) input signal has two abrupt changes.

programmed to +1 which resemble a 'smoothing filter' and expected to smooth out abrupt transitions. The digital output vector, as well as the analog output vector, are depicted in the bottom plot. The expected ideal output vector is also plotted for comparison. Figure 2.20(b) shows another experiment, where the input data has two abrupt changes: one from -1 to +1 and the another from +1 to -1. The outputs are again displayed in the bottom plot. The root-mean-square error (RMSE) for the analog output vector is 0.044 or 2.2%. The RMSE is increased to 0.065 corresponding to 3.2% after digitization (CFS output).

To demonstrate the effectiveness of the designed CFS in object tracking, a custom filter was designed to detect vehicles based on the MOSSE algorithm [9]. Figure 2.22 shows the designed filter kernel. As illustrated in Figure 2.21(a) this two-dimensional filter is then decomposed into a *vertical* and a *horizontal* filter (Figure 2.22(b)). The test image went through a few preprocessing steps to reduce shadow and intense lighting effects [10]. The test process is illustrated in Figure 2.21(b). The *preprocessed image* rows were scanned into the chip and correlated with the vertical filter to form a *temporary output*. Afterward, the columns of the temporary output were scanned into the chip and correlated with the horizontal filter to produce the *final output* which is expected to exhibit a sharp peak when there is an authentic match between

31

the filter and the target in the input image. Figure 2.22(d) depicts the expected output from an ideal digital filter implemented in MATLAB for a test image from the DARPA VIVID dataset[14] and Figure 2.22(e) shows the measured results. The final outputs of the analog filter match closely with the simulated digital filter, indicating negligible degradation due to noise, mismatch, etc. The analog filtered image shows a strong peak at the target location which clearly discriminates the target from the background. It worth noting that this prototype only includes one set of floating gate memories; however, the future versions of the chip will include multiple kernels that can be selected at run-time, so that both the horizontal and vertical filters can be executed with no reprogramming. A similar solution could be implemented using two of the current chips, with one programmed for the horizontal filter and one for the vertical filter.
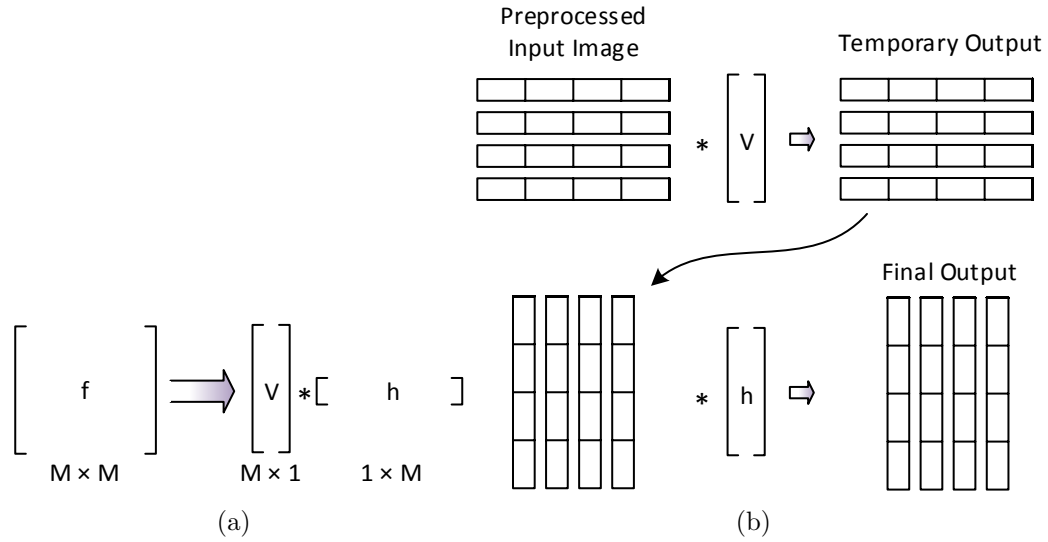
Figure 2.21: (a) A separable two-dimensional filter kernel (f) is decomposed to two vectors, called horizontal (h) and vertical (v) filters (b) Filtering with f is equivalent to filtering first with v and then filtering the result with h.
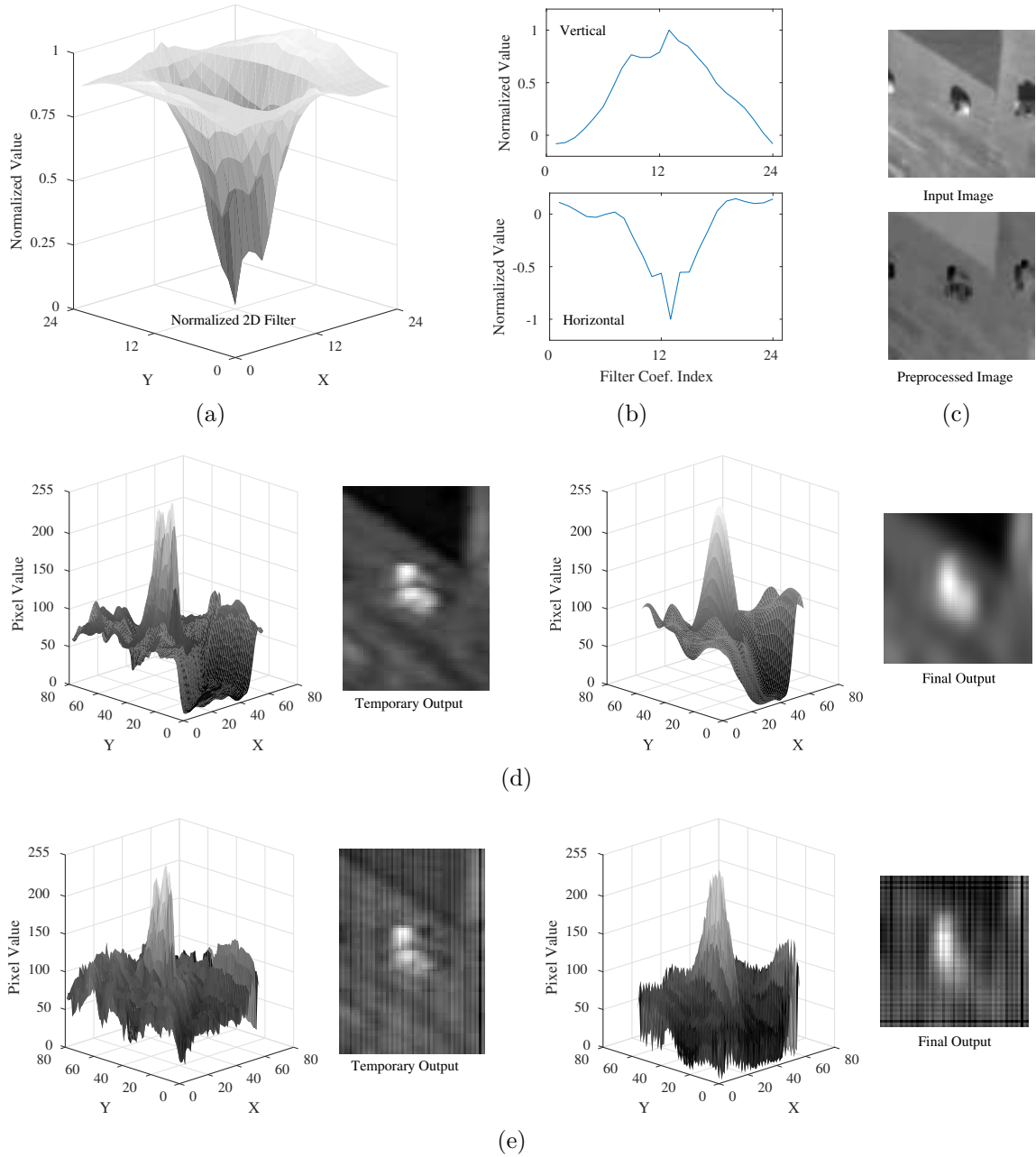
Figure 2.22: The vehicle detection test: (a) normalized two-dimensional filter kernel, (b) decomposed vertical and horizontal filters, (c) original and preprocessed input test image, (d) temporary and final outputs from an ideal digital filter implemented in MATLAB, and (e) temporary and final outputs from the chip.

2.7 Conclusion

A 24×57 correlation filter system has been p resented. The proposed system performs MAC operations in the analog domain and provides a standard and scalable digital I/O interface for data transfer. An array of non-volatile floating-gate memories are used to store filter coefficients. The CFS chip dissipates 388.4 μW of power at a throughput of 11.3 kVec/s, achieving an energy efficiency of 25.2 pJ/MAC. The fabricated chip occupies 3.23 mm^2 of the silicon area in a 0.13 μm CMOS process. A custom filter based on the MOSSE algorithm to detect vehicles was programmed into the CFS chip. The result of applying the filter on an image from DARPA VIVID dataset was presented, showing a strong peak at the location of the target.

Chapter 3

A PWM-Based Correlation Filter System

In the previous chapter, a Gilbert-multiplier-based architecture for CFS was presented and it was shown that the output of the system matches closely with the expected output. Nevertheless, here is a summary of limitations regarding the Gilbert-multiplier-based CFS:

- Operating Speed

  The speed for the multiplier array is limited by the input stage and is directly proportional to input current amplitude and inversely proportional to the number of rows in the array. Input current amplitude is limited to a few hundred nanoamps because of the fact that the multiplier is linear only in the sub-threshold region. Adding auxiliary circuits, e.g. OTAs help reduce the input capacitance but at the cost of additional power and circuit complexity. In addition to the multiplier array, digital I/O interface can also limit the overall speed. The low power DACs and the S/H circuits operating in the sub-threshold region at the input interface can operate up to $6MHz$ and the output interface operates at a maximum speed of $2.4MHz$.

- Power Consumption

Although the proposed Gilbert-multiplier-based CFS demonstrated a better energy-efficiency compared to similar works reported in the literature, it was clear that energy consumption of the system was dominated by the I/O interfaces. The input and output interfaces consume ∼13% and ∼62% of the power, respectively. The power consumption for the output interface was dominated by the analog buffers (∼45%) which are required to drive the A/D converters input capacitance.

- Input linearity range

  The multiplier in the Gilbert-multiplier-based CFS was linearized by applying degeneration techniques. However, the linear range depends on the input current level. An optimum design that could cover the entire input current level results in a narrower linear range than a linearized differential pair operating at a certain bias current.

In this chapter, an alternative architecture is proposed to address the issues listed above. The proposed architecture is capable of operating at a higher speed and a lower power consumption.

3.1 PWM-Based Four-Quadrant Multiplication

The search for a better way of implementing multiplication lead to using the *time* as the second factor (the first factor being the voltage stored on floating-gate memories). In this method, the image pixel values are encoded into pulse-width modulated (PWM) signals. Figure 3.1 shows four-quadrant multiplier implementation using the PWM modulation technique. The transconductance amplifier (TC) block converts the filter coefficient (stored as voltage) to a differential current signal and the switch network controls the direction of current flow. By defining the clock period, T_u, as the time unit, the PWM signals for a $M - bit$ multiplication can be written as:
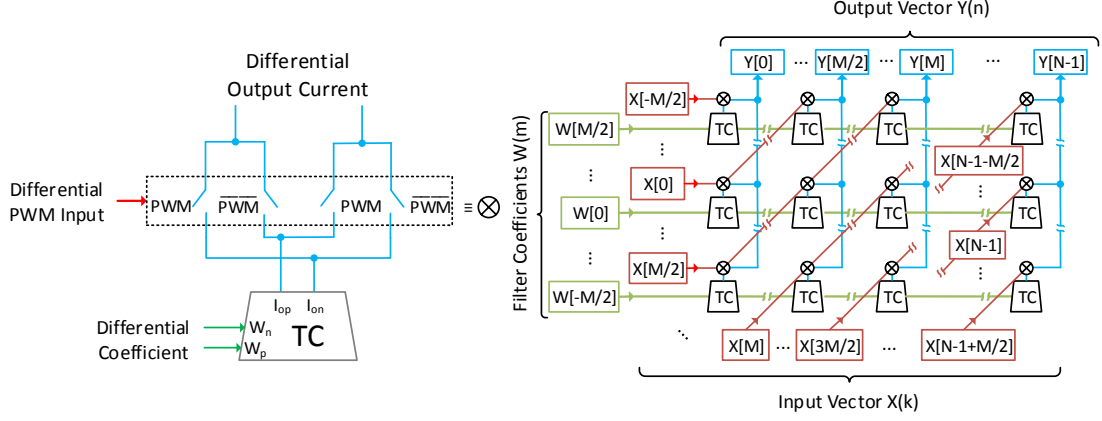
Figure 3.1: Left: The PWM-based multiplication. The image pixels are encoded to PWM signals and the filter coefficients are converted to current signals. Right: A fully-parallel time-domain multiplier array. The wire-summed current at each column will be integrated to a capacitor, forming the product voltage.

$$PWM = d_{in}T_u$$
$$\overline{PWM} = (2^M - d_{in})T_u \tag{3.1}$$

where d_{in} is the 8-bit image pixel value. The output current is integrated into a capacitor C_{int} (Figure 3.5), resulting in a voltage that is proportional to the product of the image pixel and filter coefficient values:

$$V_o = \frac{g_m T_u}{C_{int}}(W_p - W_n)(d_{in} - 2^{M-1}) \tag{3.2}$$

To convert the product back to digital, the voltage stored on the capacitor is discharged by a constant reference current source I_{dis} (Figure 3.5) and the time that V_o reaches V_{ref} is measured by a counter running at $1/T_u$ clock frequency.

In PWM-based multiplication, the current switching happens at the output of the transconductance amplifier which has much lower parasitic capacitance compared to the input of the multiplier in previous architecture. Therefore, this system can potentially operate at much higher speeds than the Gilbert-multiplier-based CFS.
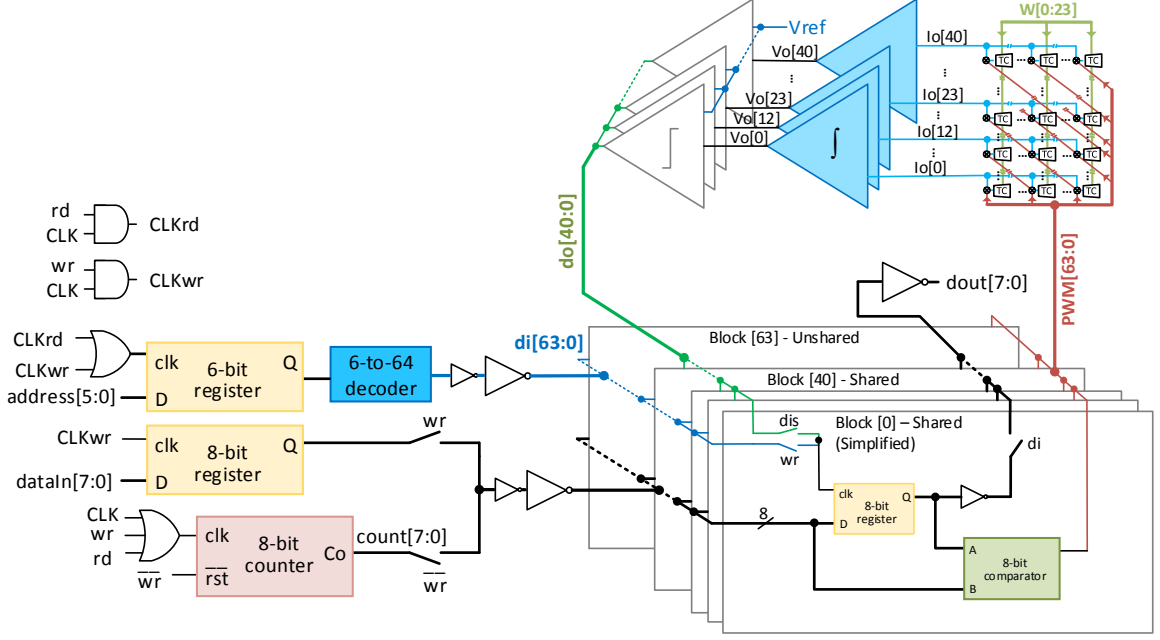
Figure 3.2: Simplified block diagram of the PWM-based CFS.

Furthermore, this architecture can be easily scaled without much decrease in operating speed. Increasing the array size means that the PWM signals need to drive extra minimum size switches and interconnects. This is in contrast with the Gilbert-multiplier-based CFS, where analog currents were driving non-minimum size (analog) devices plus interconnects. Therefore, this technique can be used for large arrays without a significant speed penalty.

3.2 System Description

Figure 3.2 shows block diagram of the proposed architecture. It is clear from this block diagram that, unlike the Gilbert-multiplier-based CFS, the I/O interface does not require any explicit D/A or A/D converters. Instead, this architecture utilizes digital-to-PWM (D/PWM) and PWM-to-digital (PWM/D) conversion all realized in digital (using a counter, a digital comparator, and a dynamic comparator). All-digital I/O interface is indeed an attractive feature for this type of systems especially moving

toward deep sub-micron CMOS technologies. As technology scales, the propagation delay of digital gates reduce, and therefore the I/O interface can operate at higher speeds. Operating at high frequencies reduces the leakage current contribution in total power dissipation, resulting in a better energy-efficiency. Additionally, digital circuits are capable of operating under low supply voltages, making them a suitable choice for I/O interfacing in modern technologies.

The active analog blocks in this architecture include transconductance amplifiers and integrators. These blocks perform the main processing task of multiply-accumulation which, as discussed before, is costly in terms of area and power consumption if implemented in digital.

In summary, this architecture takes full advantage of all-digital design over analog or even mixed-signal design (A/D or D/A as in Gilbert-multiplier-based CFS) in I/O interfacing, and, on the other hand, benefits from the efficiency of analog design in signal processing.

The system building blocks are described in the following sections.

3.3   Digital I/O

In this architecture, the energy efficiency of digital I/o interface is increased by exploiting several techniques. Firstly, low-threshold devices are employed instead of regular devices. It should be noted that using low-threshold devices has a caveat: increased leakage power consumption of the digital blocks in low-frequencies.
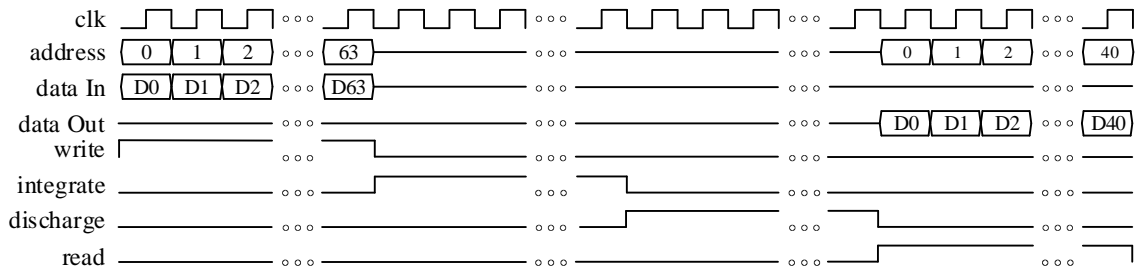

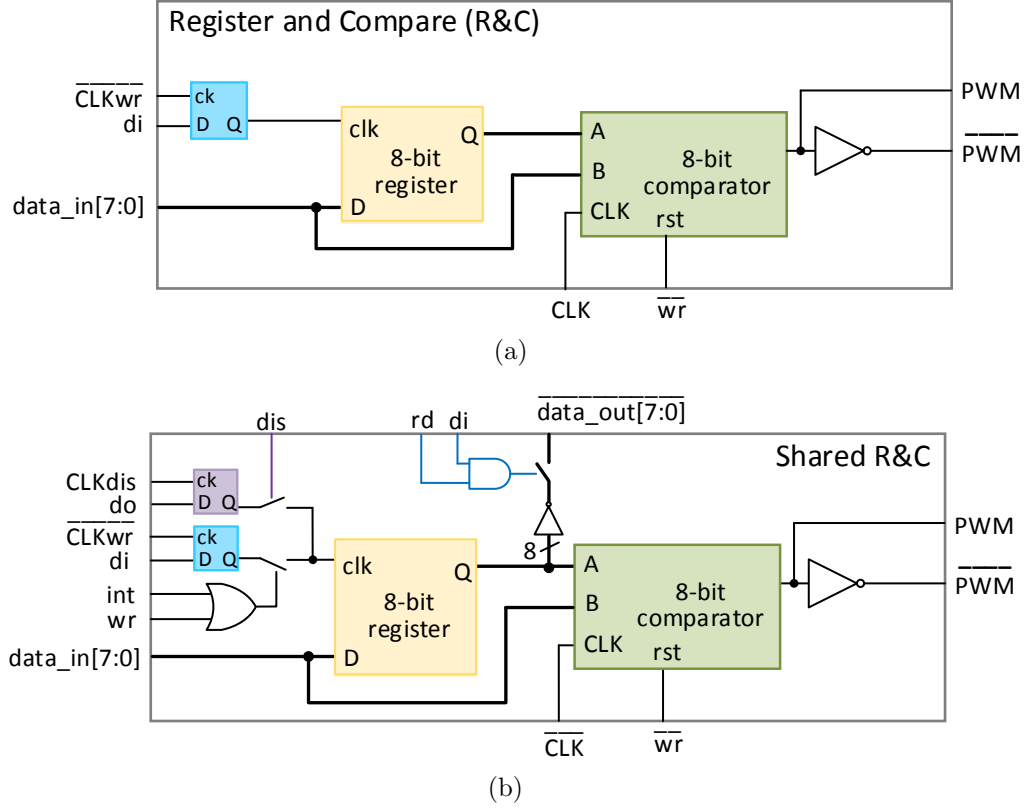Figure 3.3: Timing diagram of the PWM-based CFS.

Figure 3.4: (a) The register and compare (R&C) block stores the input data and compares it with the counter value to generate the PWM signals. (b) The shared R&C block not only generates the PWM signals in the *integration* phase but also stores the counter value during the *discharge* phase to be read as the digital output data.

However, this system is expected to operate at clock frequencies higher than 100 MHz with a 1 V power supply voltage which reduces the leakage current. Secondly, unnecessary switching activities are minimized. This was done by carefully exploiting clock gating techniques throughout the design. Finally, power- and area-hungry resources are shared. For example, an array of M × N multipliers have N outputs and M+N-1 inputs. Instead of assigning one counter per input and one per output, they could share one 8-bit counter for both digital-to-PWM (D/PWM) and PWM-to-digital (PWM/D) conversions. Moreover, each input and each output need an 8-bit register. However, since those registers are not utilized during the same time periods, they are shared to save even more area and power.
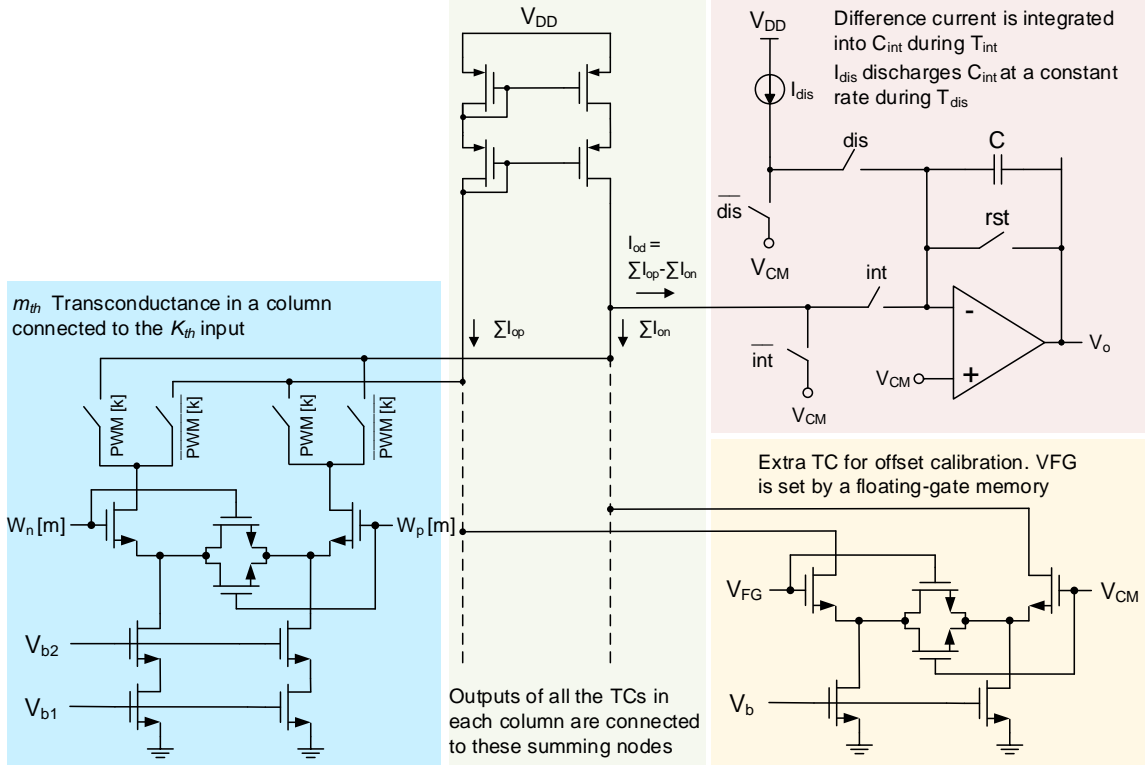
Figure 3.5: Detailed diagram illustrating the analog processing blocks and offset calibration circuit.

Figure 3.4(a) shows the block diagram of the register and compare (R&C) block. The R&C blocks store the input data during the *write* phase and generate the PWM signal by comparing it to the counter value during the *integration* phase. From a total of 64 R&C blocks, 41 of them are shared between the inputs and outputs. The block diagram of the shared R&C block is displayed in Figure 3.4(b). The main difference between the two blocks is that in addition to generating PWM signals during the *integration* phase, the shared R&C block stores the counter value when the analog comparator makes a decision during the *discharge* phase. This stored data is indeed the output data which is read out during the *read* phase. The timing diagram of the proposed CFS is depicted in Figure 3.3.

3.4 Analog Processing and Offset Calibration

Figure 3.5 illustrates the analog processing blocks. The transconductance amplifier is a degenerated differential-pair operating in the sub-threshold region. As was mentioned before, a PWM signal controls the current flow to the summing nodes. In this design, there are 24 transconductance amplifiers at each column with their outputs connected to the summing nodes.

To compensate for the mismatch induced offset errors an extra transconductance amplifier is incorporated at each column. This transconductance amplifier is adjusted by a floating-gate memory. A cascode current mirror subtracts ΣI_{op} from ΣI_{on} (differential to single-ended conversion). The difference current I_{od} is then integrated into capacitor C_{int} during the *integration* phase. A current source discharges C_{int} at a constant rate at the *discharge* phase.

Figure 3.6 depicts the schematic of a class-AB OTA designed for integrator[1]. The class-AB operation allows the integrator to operate effectively in association with other high-speed blocks in the system while having a low quiescent bias current. In particular, it helps the integrator to settle faster in *reset* phase (reset happens during the *write* phase) by drawing instant high current from the power supply. To explain the operation of this OTA let's consider the end of a computation cycle where the C_{int} is discharged and the V_{OUT} value is close to ground and V_{IN+} is at V_{CM}. At this condition reset switch in Figure 3.5 closes as part of the preparedness phase for the next cycle. At this moment, transistor M_3 turns off causing a voltage increase at node A. This increases the V_{GS} of M_5, drawing more current from M_2 thus M_8 and M_{10}. This considerable current flow instantly pulls the $V_{OUT} = V_{IN-}$ voltage up to V_{CM}.

During the *discharge* phase and right after that the following comparator makes a decision, a network of switches deactivate the OTA to reduce the average power consumption. This reduction is noticeable at the bottom plot of Figure 3.27 which

---

[1]Design credit goes to Tan Yang.
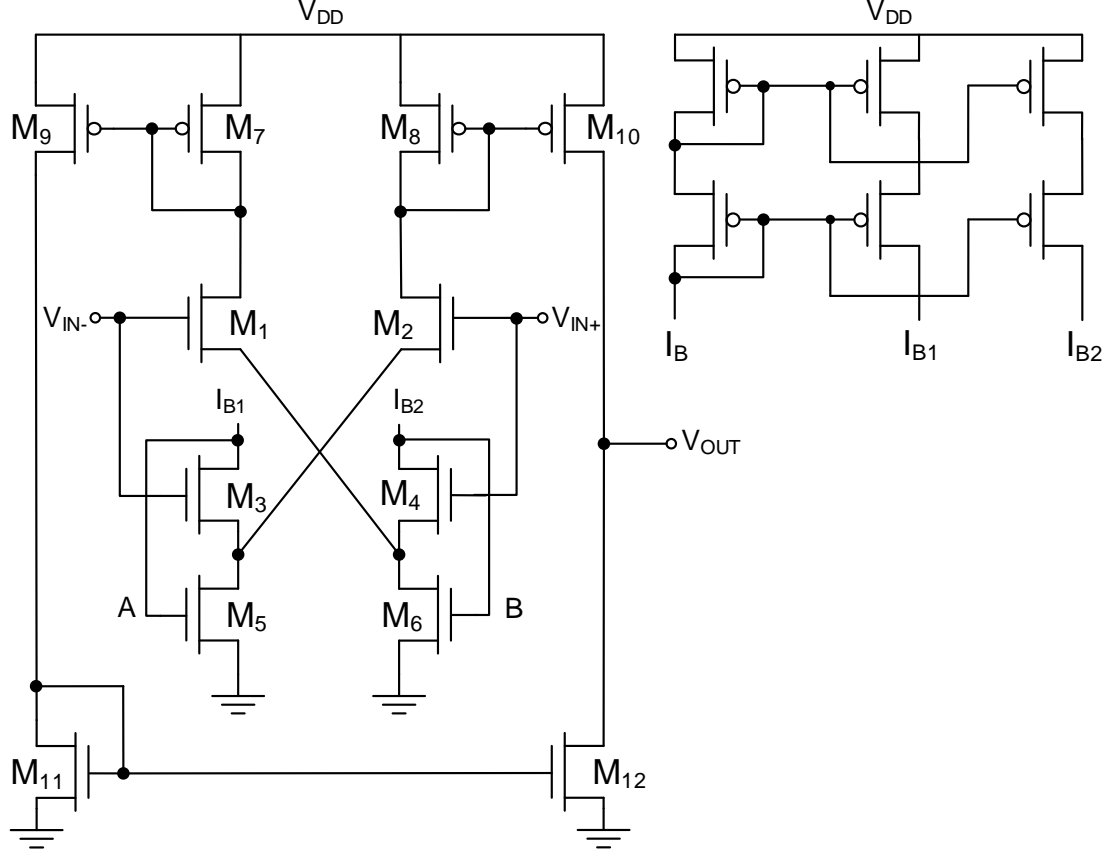
Figure 3.6: Schematic of the integrator OTA.

shows the power supply current profile of the integrators based on a post-layout simulation. To account for process variations and also to provide a flexible range for choosing the clock frequency of the system, the integration capacitor is made tunable to 8 different values. Table 3.1 shows the nominal value of C_{int} versus control bits.

Table 3.1: PWM-based CFS C_{int} Setting

C_{int} (fF)	S_2	S_1	S_0
167	0	0	0
334	0	0	1
467	0	1	0
634	0	1	1
768	1	0	0
935	1	0	1
1068	1	1	0
1235	1	1	1

3.5 Energy-Efficient Dynamic Comparator with a Fast Convergence Rate

The comparator is an essential building block in many applications that require analog to digital (A/D) conversion. The comparator plays a crucial role in the final performance of the systems. Although technology scaling has provided necessary means for low-power and high-speed comparator design, the offset voltage remains the main concern because by moving to deep submicron technology nodes not only the transistor mismatch does not improve but also degrades significantly [21]. To overcome this problem, several offset cancellation techniques have been reported in the literature. The conventional method utilizes a pre-amplifier with substantial gain to decrease the input-referred offset voltage which sometimes combines with auto-zeroing techniques implemented during negative feedback [36]. However, the maximum achievable gain for a single-stage amplifier has decreased in advanced technology nodes and adding more stages comes at the cost of extra power dissipation.

On the other hand, a dynamic comparator is a more attractive option since it does not consume any static power. Nevertheless, dynamic comparators suffer from even larger input referred offset compared to linear ones due to the extra mismatch caused by parasitic capacitors at the internal nodes. This indeed raises the necessity of implementing an effective offset calibration technique for dynamic comparators. In [43] offset calibration was implemented by digitally controlling the load capacitance of the dynamic latch. Yet the additional capacitance lowers the comparator speed and increases the chip area. Additionally, since in this method the calibration time is inversely proportional to the residual offset, it can take a significant amount of time before reaching the convergence for high-resolution A/D converters. In [2] a D/A converter controls the body voltage of input transistors in order to reduce the offset voltage. This technique realizes a 7-bit D/A converter with additional complex digital circuitry and takes 128 clock cycles to complete the calibration process. The offset cancellation method in [27], senses the offset by measuring the phase difference
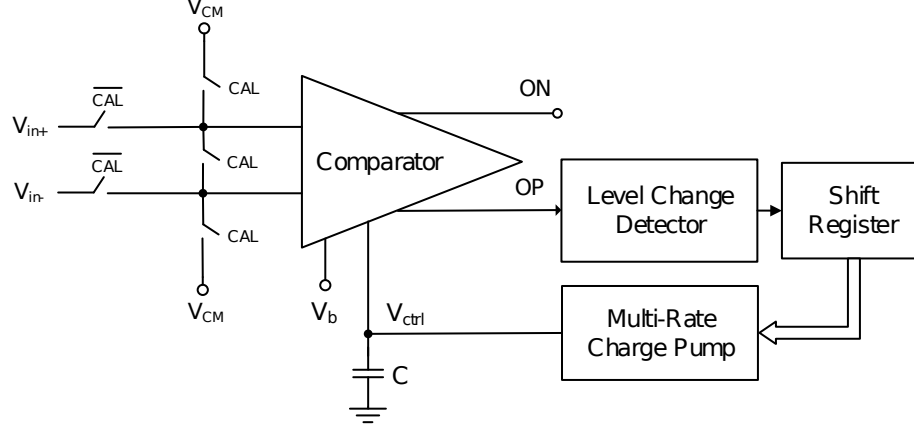
Figure 3.7: Block diagram of the proposed offset calibration technique.

between the outputs and adjusts the body voltage of the input transistors. While this technique needs only a few clock cycles for calibration, the comparator speed is limited to under 7 MHz because of the phase detector resolution. An offset cancellation method using voltage-controlled current sources is presented in [31]. A charge pump circuitry changes the gate voltage of the current source which is in parallel with input transistor and compensates for the offset. Although this technique uses very few logic circuits to control the charge pump, it requires a long calibration time if a low residual offset is desired.

Here, a novel calibration technique is proposed with no static power dissipation which is able to converge rapidly while maintaining a high precision. The circuit uses a multi-rate charge pump circuitry controlled with simple logic circuits in order to achieve a short calibration time without sacrificing the accuracy and results in a better energy efficiency.

3.5.1 Offset Calibration

The block diagram of the proposed offset calibration technique is shown in Figure 3.7. In the calibration mode, comparator inputs are connected to a common-mode voltage. A digital block generates a pulse whenever a level change occurs at the comparator output (OP). This pulse goes into a shift register which generates a

Figure 3.8: (a) The digital block that generates CN and CP signals for the charge pump and $\overline{EN}$ signal for the comparator. (b) The offset calibration flowchart.

digital sequence to control the charging/discharging rate in a multi-rate charge pump circuitry.

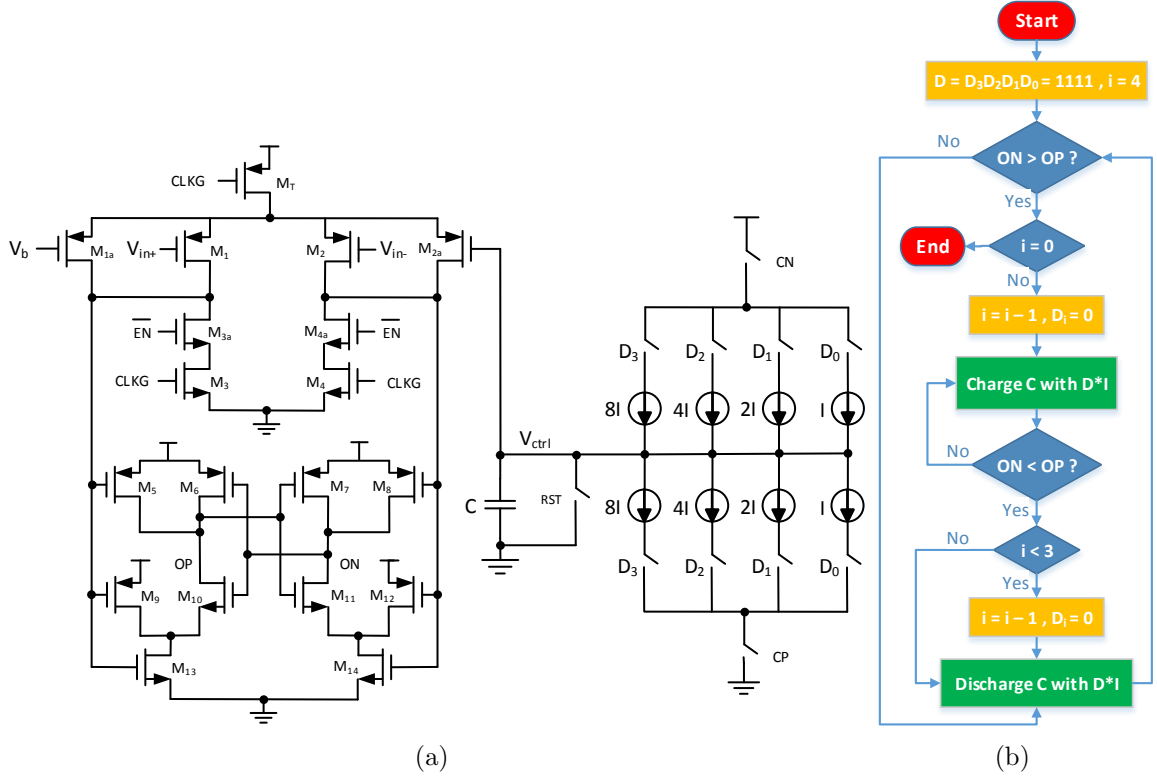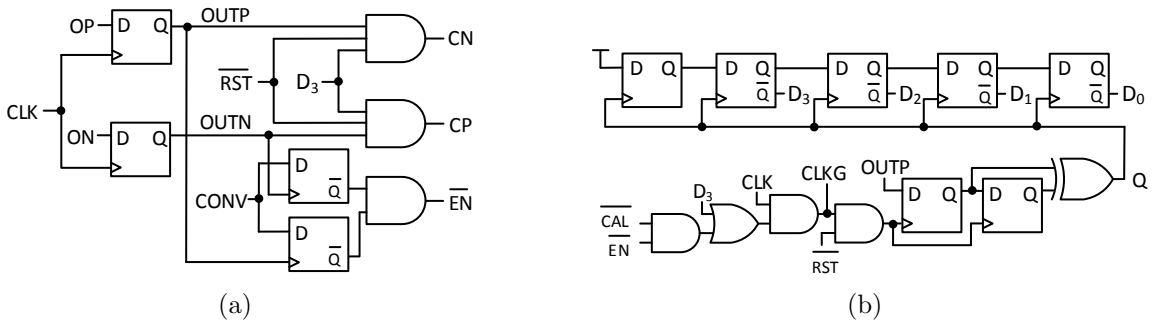Figure 3.9: (a) The schematic diagram of the double-tail latched comparator and multi-rate charge pump circuitry, (b) The digital block that generates CN and CP signals for the charge pump and $\overline{EN}$ signal for the comparator. (c) The level change detector and the following shift register that generates D_3 to D_0 bits.

The schematic diagram of the comparator and the charge pump circuits are shown in Figure 3.8(a). The circuit is designed based on a double-tail latched comparator in [31]. The comparator is clocked with $CLKG$ which is a gated version of the main CLK signal. The $CLKG$ is only active during calibration and conversion phases. The circuit shuts down the $CLKG$ after reaching the convergence in the calibration phase. Therefore the amount of time that the $CLKG$ remains active during the calibration phase depends on the comparator initial offset. Moreover, transistors M_{3a} and M_{4a} disable the comparator when a decision has been made to save power. The clock gating approach combined with the fast convergence rate during the calibration phase, as well as, deactivating the comparator after making a decision during the conversion phase can have a significant impact on energy efficiency of the VLSI systems, where arrays of such comparators operate in parallel.

Transistors M_{1a} and M_{2a} are voltage controlled current sources. The gate of M_{1a} is connected to a reference voltage, V_b, and the gate of the M_{2a} is connected to capacitor C which holds the control voltage, V_{ctrl}. The V_{ctrl} which is regulated by the charge pump circuitry controls the gate of M_{1a} in order to balance out the mismatch caused by process variations by injecting current into internal nodes. The charge pump rate is controlled by 4 bits. The current sources and sinks are arranged in a binary weighted form as $8I$, $4I$, $2I$, and I. Figure 3.10 illustrates the timing diagram and the typical waveforms during the offset calibration phase. At the beginning of the calibration phase, the capacitor C is initialized to the V_b and all the D_3 to D_0 bits are set to logic '1' and thus, all the current sources and sinks are connected to the V_{ctrl} node. The charge pump starts with the highest current of $15I$ which leads to the highest pump rate and then decreases the rate to $7I$, $3I$, and I as V_{ctrl}, approaches the final value. The digital sequence managing these transitions is generated by the level change detector and the shift register. This sequence is simply implemented by turning off D_3 to D_0 switches, one at a time when the OP changes from logic '1' to logic '0' or vice versa. This process continues until the last bit, i.e. D_0 goes from logic
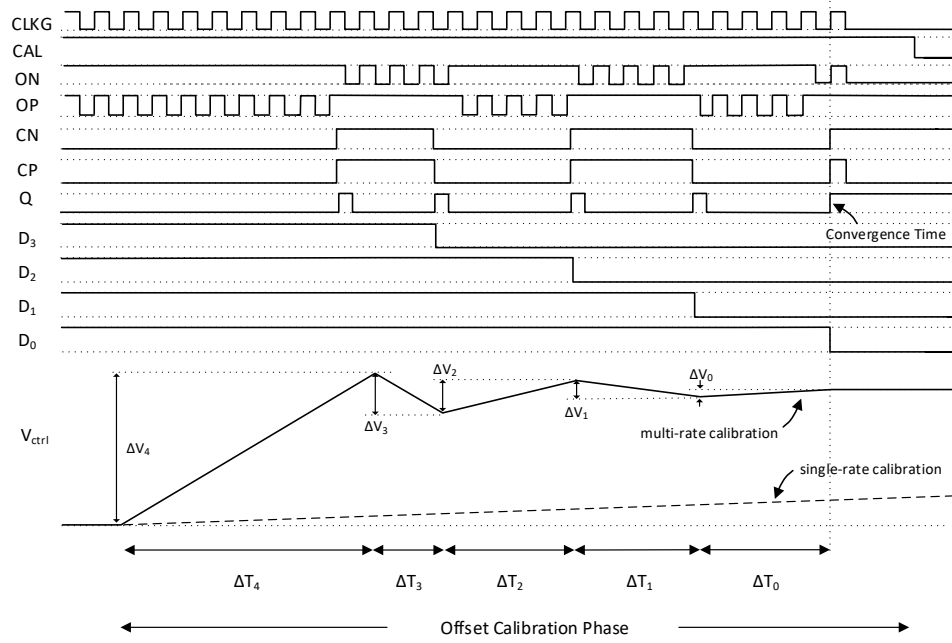
Figure 3.10: The timing diagram and typical waveforms during offset calibration phase.

'1' to logic '0', disconnecting the last current source and sink from the V_{ctrl} node, and ending the calibration phase.

It is worth mentioning that unlike the conventional method, in this method the charge pump is controlled by CN and CP signals rather than by comparator outputs (ON and OP). As shown in Figure 3.9(a) the CN and CP signals are generated by sampling comparator outputs at the rising edge of the CLK, and therefore, they remain at their current state until comparator outputs change. The level change detector and the shift register circuits are depicted in Figure 3.9(b). One of the outputs of the comparator is sampled at the rising edges of two consecutive $CLKG$ signal in order to capture any logic level change. The indicator signal, Q is then used to clock a shift register circuit which simply shifts a logic '1' to the right and therefore generates D_3 to D_0 bits as shown in the timing diagram of Figure 3.10.

48

3.5.2 Convergence Time

Thanks to the utilization of the multi-rate charge pump circuit, the proposed calibration scheme needs a very short time to reach the convergence. The dashed line in Figure 3.10 represents the V_{ctrl} node voltage in case of a conventional single-rate method with a current source I [31]. For the single-rate case, the calibration time for an initial offset value of V_{os} is equal to:

$$T_{cal,s} = V_{os} \cdot \frac{C}{I} \tag{3.3}$$

Based on this equation, for a short calibration time, one should use large I and small C. On the other hand, the residual offset voltage after calibration can be written as:

$$V_{res} = T_{clk} \cdot \frac{I}{C} \tag{3.4}$$

Clearly, the residual offset has an opposite proportionality to I and C, meaning that reaching low residual offset requires long calibration time. The total calibration time in the case of proposed multi-rate calibration, for the same residual offset, can be written as:

$$T_{cal,m} = \Delta T_0 + \Delta T_1 + ... + \Delta T_4 \tag{3.5}$$

where ΔT_i are the time intervals depicted in Figure 3.10. A simple calculation leads to:

$$T_{cal,m} = N \cdot T_{clk} + V_{os} \cdot \frac{C}{15I} \tag{3.6}$$

where N can be calculated from the current source ratios. If the current sources are binary weighted, N is close to 12. Thus, the first term on the right-hand side of (3.6) is constant for a given clock frequency. However, the second term depends on the V_{os} and has a 15X faster rate compared to the conventional single-rate method. As illustrated in the Figure 3.10, the total calibration time is significantly reduced compared to the single-rate approach, with the same goal for the residual offset.
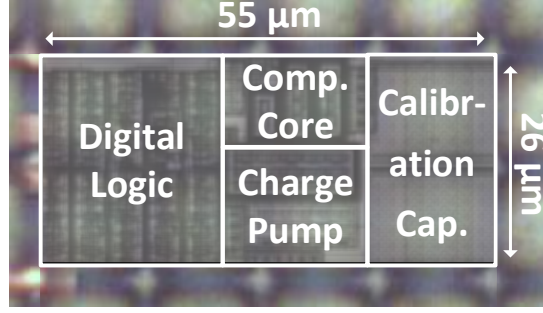
49

Figure 3.11: Chip micrograph of the comparator and offset calibration circuit.

3.5.3 Simulation and Experimental Results

The comparator circuit was designed and fabricated in a 0.13μm CMOS process. The chip micrograph is shown in Figure 3.11. To verify the effectiveness of the proposed technique over process variation and device mismatch, a statistical analysis was performed using the Monte-Carlo simulation. One hundred iterations were performed for the comparator with and without offset calibration. The circuit was
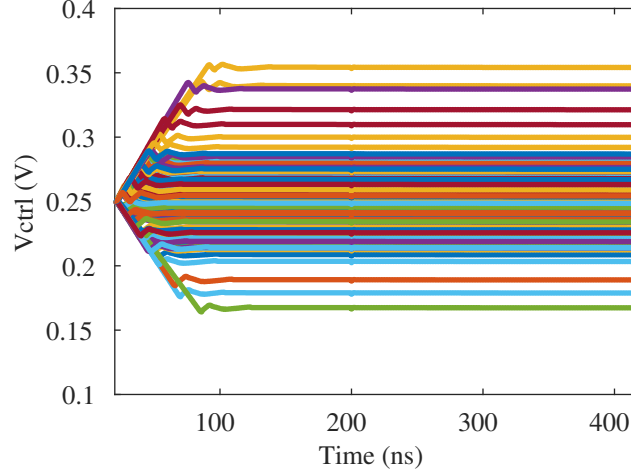

Figure 3.12: Convergence of the offset calibration for 100 Monte-Carlo iterations.

Table 3.2: Convergence Time

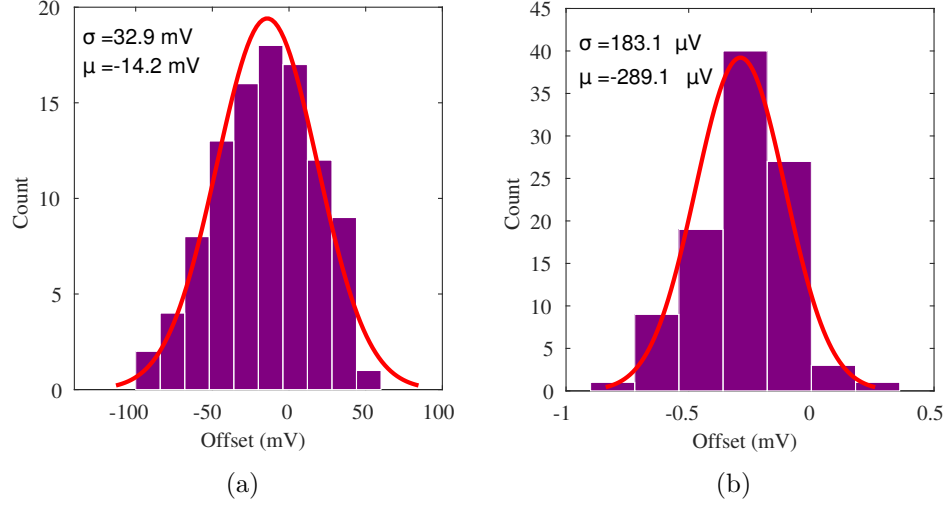| Specification	multi-rate (this work)	single-rate
Mean (μ)	82 ns	286 ns
Standard deviation (σ)	15 ns	680 ns

50

Figure 3.13: Simulation results from 100 Monte-Carlo iterations (a) without calibration (b) with calibration.

operating at 1 V power supply and 500 MHz clock frequency. Figure 3.12 shows the V_{ctrl} voltage during the time for the aforementioned Monte-Carlo simulation. While the maximum convergence time was only 139.6 ns, the average time was 82 ns with a standard deviation of 15 ns which is a significant improvement over the single-rate method, simulated at the same conditions, with an average of 286 ns and a standard deviation of 680.6 ns. Figure 3.13 shows the residual offset distribution of the comparator with and without calibration. As it is clear, the offset voltage without
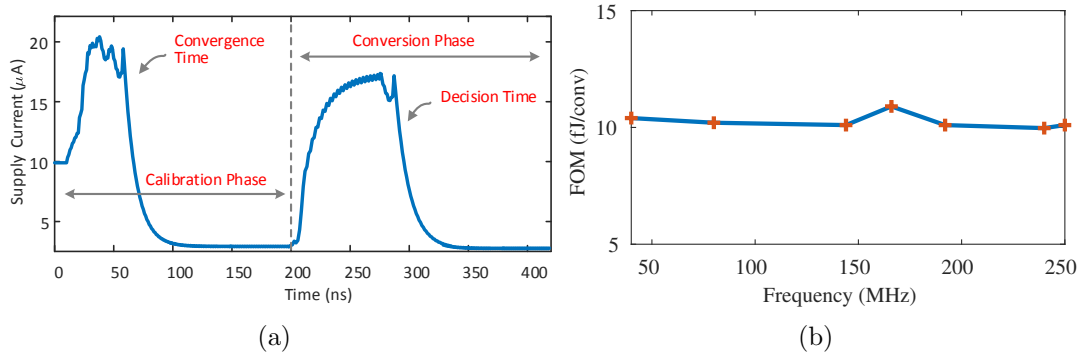

Figure 3.14: (a) Typical power supply current profile during offset calibration and conversion phases. (b) Measured FOM over clock frequencies from 40 MHz to 250 MHz.

51

Table 3.3: Comparator Summary of Performance

Specification	Value
CMOS Process	0.13 μm
Supply Voltage	1 V
Average Power Dissipation	5.1 μW
Operating Frequency	250 MHz
FOM	10.2 fJ/conv
Area	26×55 μm^2
Residual offset (σ_r)	183.1 μV
Offset Reduction Ratio (σ_i/σ_r)	179 X
Average Convergence Time	82 ns
Maximum Convergence Time ($\pm$100mV)	139.6 ns

calibration varies in a range of $\pm$100mV. The standard deviation of comparator offset without offset calibration is 32.9 mV which is reduced to 183.1μV with calibration, indicating a 179 times improvement. Figure 3.14(a) shows the power supply current profile during calibration (0-200ns) and conversion (200-450ns). This current profile shows how shorter calibration time reduces the average power dissipation and leads to a better FOM for the comparator.

The measured average power consumption at 250 MHz clock frequency was 5.1 μW, achieving a FOM of 10.2 fJ/conv. Figure 3.14(b) plots the measured FOM over a wide range of operating frequencies from 40 to 250 MHz. A summary of circuit

Table 3.4: Comparator Performance Comparison

Specification	This work	[31]
CMOS Process	130 nm	90 nm
Supply Voltage	1 V	1 V
Average Power Dissipation	5.1 μW	40 μW
Operating Frequency	250 MHz	1 GHz
FOM	10.2 fJ/conv	20 fJ/conv
Residual offset (σ_r)	183.1 μV	1.3 mV

performances is given in Table 3.3 and Table 3.4 compares the performance of this work with [31]. By fast convergence, the proposed circuit reduces the time required for offset calibration and thus, achieves two times better FOM compared to work reported in [31].

3.6   Experimental Results (First Prototype)

The proposed PWM-based system was designed and fabricated in a 130 nm CMOS process. Figure 3.16(a) shows the chip micrograph. The 24×41 PWM-based CFS occupies an area of $0.65 \times 1.2 mm^2$. Figure 3.16(a) shows measured output code versus input code for all of the 41 outputs. All coefficients were programmed to 50 mV for this test. The counter clock frequency was 20 MHz and the write/read clock frequency was 5 MHz.

The measurement results revealed that there were a limited number of problematic codes which were identified as: [0x80,0xC0,0xE0,0xF0,0xF8,0xFC,0xFE] The source of the problem was found to be the propagation delay of the asynchronous counter at the time of reset which could initiate a false comparison result at the output of the digital comparator and consequently could lead to firing a wrong PWM signal. This problem was fixed in the revised design by making sure that the digital comparator observes the valid data during the reset period.
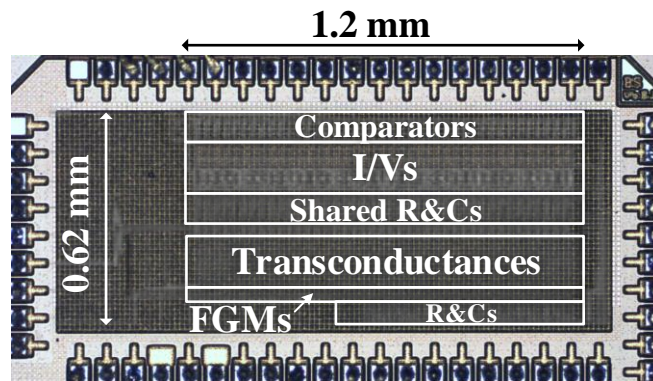

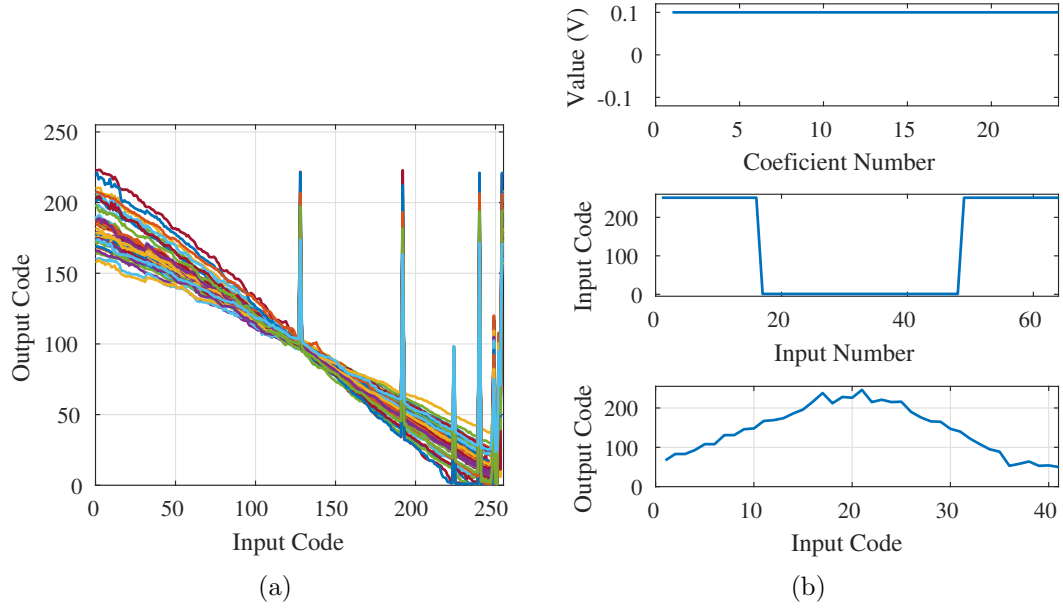Figure 3.15: Micrograph of the PWM-based CFS chip.

Figure 3.16: Measurement results: (a) Output code versus input code for all of the outputs. (b) Correlation output between two vectors: coefficients resemble a smoothing filter and the input signal has two abrupt changes.

The first prototype of the chip did not include digital buffers for the digital outputs as well as for the $di[63:0]$, $do[40:0]$, and $din[7:0]$ signals which drive long on-chip interconnect wires. The output data $Dout[7:0]$ which drives the PADs were not buffered either. For that reason, the write/read speed was dramatically compromised. Digital buffers were added to the critical nodes in the revised prototype.

Figure 3.16(b) shows the correlation output between two vectors (similar to test in Figure 2.20). For this test, all of the filter coefficients were programmed to 0.1 V while the input vector had two abrupt changes. The filter smooths out abrupt transitions of the input vector. The output is displayed at the bottom plot. This test shows that the large signal response of the system is similar to the Gilbert-multiplier-based CFS. Table 3.5 summarizes the measured performance of the PWM-based CFS.

Table 3.5: PWM-based CFS Measurement Results

Specification	Value
Supply Voltage	1 V
Array Size	24×41
Active Area (mm^2)	1.2×0.62
Tclk-Counter (ns)	6.25
Tclk-Write (ns)	12.5
Tclk-Read (ns)	50
Total time (μ s)	6.05
Throughput ($kVec/s$)	165
Total average power diss. (mW)	1.16
Energy Efficiency (pJ/MAC)	7.1

3.7 Revised Prototype

To address issues observed in testing the first prototype, a revised version was designed and submitted for fabrication. the revised chip was under fabrication at the time of writing this dissertation, and therefore no measurement results are presented. However, the performance of the revised version is verified by extensive nominal, statistical and post-layout simulations which will be presented shortly. But first, I will discuss a few important changes that have further improved the performance of the PWM-based CFS.

3.7.1 Asynchronous vs. Synchronous Counter

In addition to the problem during the *reset* phase caused by the propagation delay of the asynchronous counter, it also limits the maximum operating frequency of the system. Based on post-layout simulations the worst-case delay (MSB transition) was ∼2.8 ns which limits the maximum operating speed of the counter to ∼178 MHz. On the contrary, the estimated propagation delay for a synchronous counter is less than 115 ps (∼24X faster). However, the penalty paid is a ∼20X increase of the power consumption ($6.1\mu W$ to $51.5\mu W$). It should be noted that since the counter is shared for all the blocks, this power increase only accounts for about 4% of the total power

dissipation. Therefore, the asynchronous counter was replaced with a synchronous one in the revised prototype to achieve higher operating speed and better reliability at the expense of a slight increase in the power dissipation.

3.7.2 A Linearized Pseudo-differential OTA

A linear transconductance block, or operational transconductance amplifier (OTA), is an essential building block in analog signal processing systems. Achieving a wide linear range for the OTA in modern CMOS technologies is difficult because of the reduced power supply. A high linear range at low power supply not only helps to implement energy-efficient systems but also improves the signal-to-noise ratio (SNR) of the system simply by allowing for a larger signal amplitude to be processed.

Furthermore, demand for low-power operation has motivated the analog designers to exploit circuits operating in moderate and weak inversion regions. It is relatively easy to achieve a wide linear range in strong inversion due to the square-law relationship between the input voltage and output current. And the linear range can be extended by increasing the current level. However, the exponential relationship in sub-threshold region limits the linear range to less than a thermal voltage, $U_T \approx 26\text{mV}$ and it does not depend on the current level.

Various linearization techniques have been reported in the literature, including source, gate, and bulk degeneration of a differential pair [25, 13, 40, 23, 18], bulk-driven transistors [25, 44], floating and quasi-floating gate transistors [44, 29, 30], triode input transistors [20, 1], capacitive division [37], current division [42, 4, 44, 15, 3], and pseudo-differential structures [32, 16, 6].

Among the techniques listed above, source, gate, and bulk degeneration is the most favorable one since it can be implemented on a simple differential pair without increasing the power consumption, noise, mismatch offset, and area of the fully-differential (FD) OTAs. Nonetheless, the input (and linearity) range of the FD OTA is limited due to the voltage drop across the tail current source.

The pseudo-differential (PD) OTAs are designed by removing the tail transistor and thus they can operate under a low power supply voltage and wider input ranges. However, the common mode gain (A_{CM}) of a PD OTA is equal to the differential mode gain (A_D) and hence, common-mode rejection ratio, $CMRR = A_D/A_{CM} = 1$. Therefore, one of the shortcomings of the PD structure is the sensitivity to input common-mode voltage variations which could cause several problems for some systems including significant swings at high-impedance nodes, changes in transconductance value, and sensitivity to the common-mode noise.

However, in the proposed CFS architecture these problems are dealt with in architectural level. First of all, the common-mode voltage at high impedance nodes (output nodes) is set by the following stages (integrators). Secondly, the V_{id} is set by the programming two FGMs and therefore, the common-mode voltage is constant for each set of filter coefficients and there are no common-mode variations. And thirdly, the proposed CFS utilizes a voltage domain subtraction which effectively rejects any common-mode signal components at the output of the system. This will be discussed in detail in Section 3.7.3.

Here, a linearized PD OTA is presented which exploits the bulk degeneration along with the source degeneration. Moreover, the tail transistors which operate in the linear region, are utilized to further improve the linearity.

Figure 3.17 depicts four linearized OTA architectures based on source and bulk degeneration technique. Figure 3.17(a) (circuit A) is an FD structure, where the input transistors have low threshold voltages and tail transistors operate in the saturation region as current sources. Figure 3.17(b) (circuit B) is a PD structure, where the input transistors have high threshold voltages and tail transistors operate in the linear region. These two architectures use resistors for degeneration. Figure 3.17(c) (circuit C) is similar to Figure 3.17(b), however the only difference is that it utilizes voltage-controlled degeneration. Figure 3.17(d) (circuit D) is the proposed PD OTA which in addition to utilizing voltage-controlled degeneration, takes advantage of tail transistors to further improve the linearity.
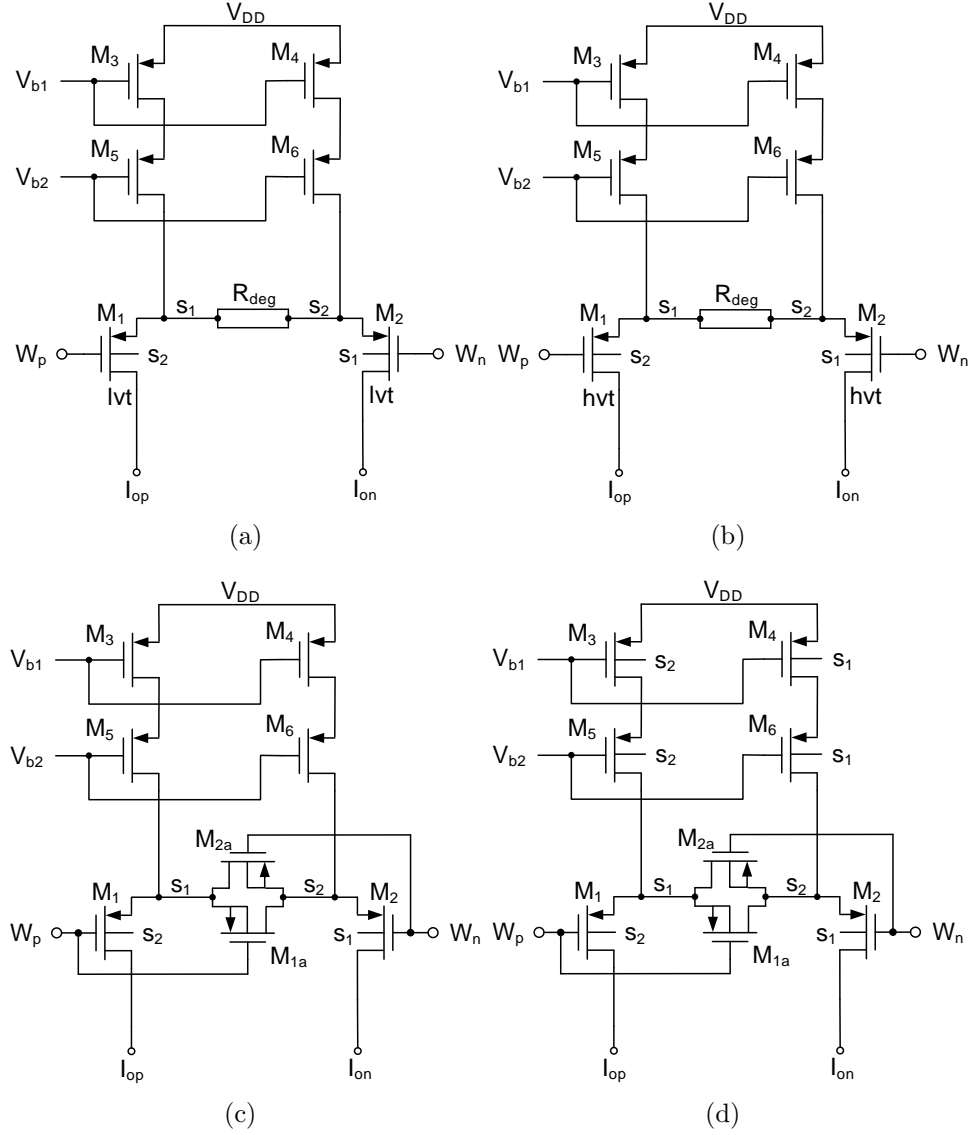
Figure 3.17: Four OTA linearization technique based on bulk and source degeneration. (a) OTA-A is an FD structure, where the input transistors have low threshold voltages and tail transistors operate in the saturation region as current sources. (b) OTA-B is a PD structure, where the input transistors have high threshold voltages and tail transistors operate in the linear region. (c) OTA-C is similar to OTA-B, however, the only difference is that it utilizes voltage-controlled degeneration. (d) OTA-D is the proposed PD OTA which in addition to utilizing voltage-controlled degeneration, takes advantage of tail transistors to further improve the linearity.

Using the small-signal model of Figure 3.18 it can be shown that the effective transconductance can be calculated from:
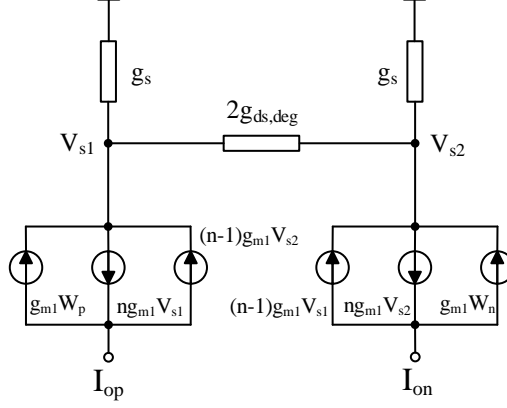
58

Figure 3.18: A small-signal model of the proposed PD OTA shown in Figure 3.17(d).

$$G_{m,eff} = \frac{g_{m1}(g_s + 2g_{ds,deg})}{g_s + 2g_{ds,deg} + (2n-1)g_{m1}} \tag{3.7}$$

Where gm_1 is the transconductance of the input transistor, gs is the transconductance of the tail transistor, gds_{deg} is the degeneration transistor transconductance, and n is the sub-threshold slope factor. The tail transistors and the degeneration transistors operate in the linear region and their conductance can be written as:

$$g_{ds} = \mu C_{ox} \frac{W}{L} (V_{SG} - |V_{TH}| - V_{SD}) \tag{3.8}$$

The V_{SG} of the tail transistors are fixed and their conductance is modulated by modifying their threshold voltage with body-effect. On the other hand, the V_{TH} of the degeneration transistors is fixed and their conductance is modulated by their V_{SG}. When V_{s1} goes down (as shown in Figure 3.19(d)) it moves V_{s2} down which reduces the threshold voltage of the tail transistor M_3 and hence, pushes more current. It worse noting that the input transistors operate in weak inversion region up to around $V_{id} = \pm 0.45V$ and then move toward moderate and strong inversion regions where their g_m decreases by increased V_{id} as implied by $I_D/(V_{GS} - V_{TH})$.

Figure 3.19 compares main parameters of the OTA architectures shown in Figure 3.17, including the input, tail, and degeneration transconductances as well as the output currents and differential transconductance.
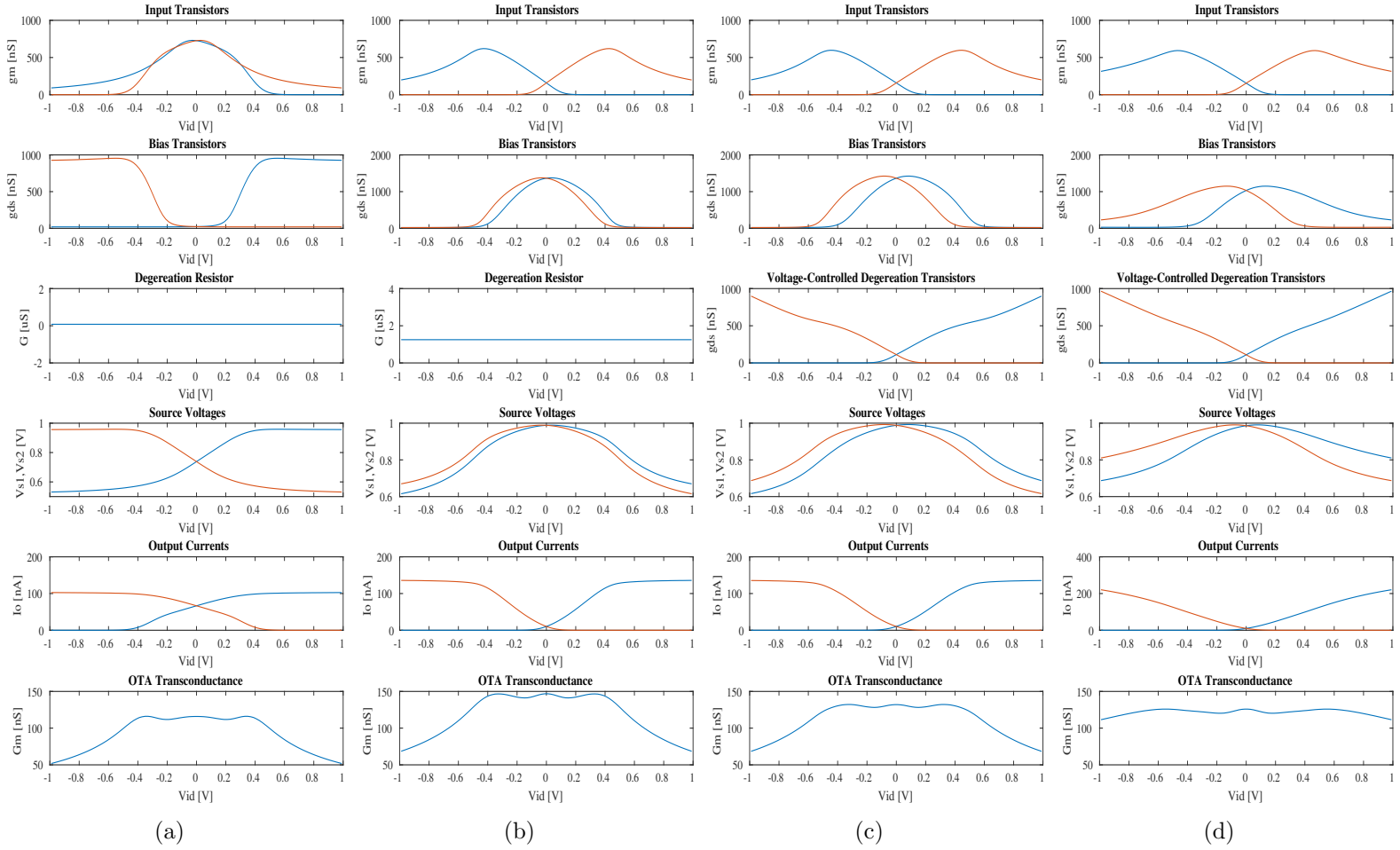
59

Figure 3.19: Main parameters of the OTA architectures shown in Figure 3.17, including their input, tail, and degeneration transconductances as well as their output currents and differential transconductances.
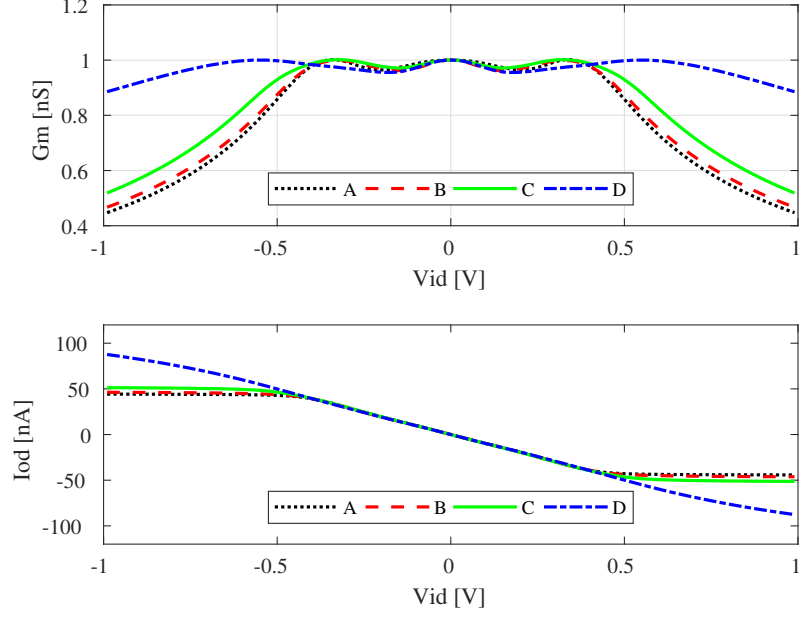
Figure 3.20: Normalized OTA transconductance versus differential input voltage for the OTAs A, B, C, and D. The linear range of the proposed OTA exceeds ±800mV.

Figure 3.20 compares normalized $G_{m,eff}$ of the OTAs A, B, C, and D. The proposed technique further widens the linearity range to more than ±800 mV. Figure 3.21 compares the transconductance error for the previously mentioned OTAs and the proposed one. The transconductance error for the proposed OTA is ±2.5% in the linear range. Figure 3.22 shows the responses of the OTAs A, B, C, and D to ±10 mV change in their input common-mode voltage. As expected, OTA-A shows the smallest change in $G_{m,eff}$, while the rest of them show less than 5% change. As mentioned in the previous chapter, the programming RMS error is less than 1 mV and therefore, the common-mode changes experienced by the OTAs are minuscule.

A linearity figure of merit (FoM) can be defined as:

$$FoM = \frac{G_{m,max}}{G_{m,eff}} \cdot \frac{V_{DD}}{V_{LR}} \cdot \%e \tag{3.9}$$

Where V_{DD} is the power supply, $\%e$ is the percentage error, V_{LR} is the linear range, and $G_{m,max}$ is the equivalent maximum transconductance reached in weak inversion for a MOS transistor:
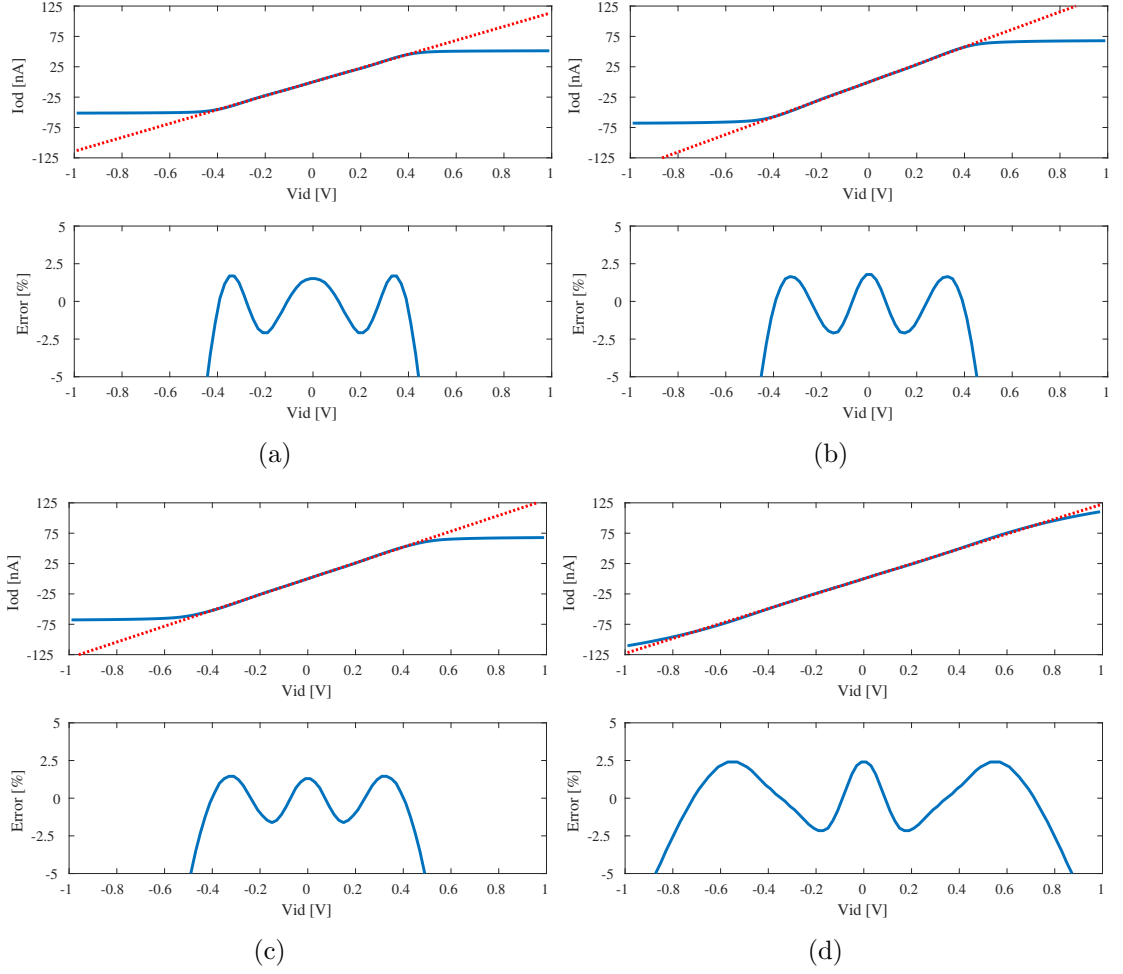
Figure 3.21: A comparison of the transconductance error between OTAs A, B, C, and D.

$$G_{m,max} = \frac{I}{n \cdot U_T} \qquad (3.10)$$

Where I is the OTA current, n is the sub-threshold slope factor, and U_T is the thermal voltage. The proposed FoM is a unitless quantity and a lower FoM value indicates a better performance of the OTA. Table 3.6 compares the performance of this work with other OTAs reported in the literature.
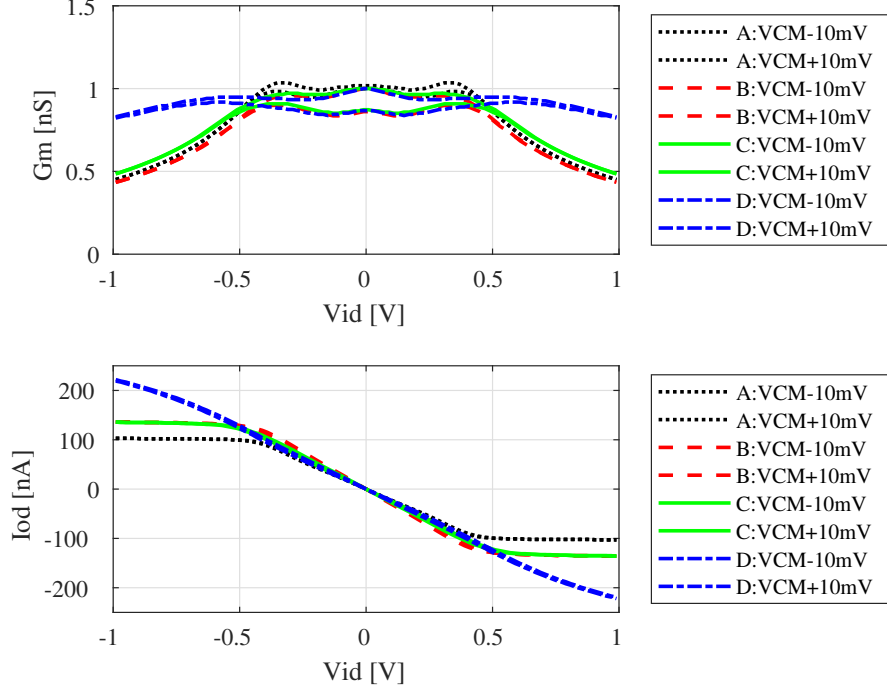
Figure 3.22: The response of the OTAs A, B, C, and D to $\pm10mV$ change in their input common-mode voltage.

3.7.3 Voltage-Domain Subtraction

One of the dominant sources of error in the first prototype was the mismatch of the current mirrors performing the current subtraction (differential to single-ended conversion). Although the current mirror was using relatively large devices and regulated cascodes, the total RMS error was still ~1.5% according to statistical simulations. Besides, large devices increased the parasitic capacitance resulting in a systematic offset at high operating speeds. Moreover, using regulated cascodes

Table 3.6: OTA Performance Comparison

Specification	This Work	ISCAS'15	TCAS-I'14	ISOCC'08	TCAS-II'07
Supply (V)	1	3.3	2.1	1.8	2.5
Current (A)	20 n	0.1 n	80 n	5 m	1 m
$G_{m,eff}$ (A/V)	125 n	0.491 n	160 n	2 m	2.4 m
V_{LR} (V)	1.6	0.26	1	1.8	1
Error ($e\%$)	2.5	1	5	1	3
FoM	0.068	0.708	1.440	0.856	0.685

63

increased the power consumption. This issue was addressed in the revised prototype by performing the subtraction in the voltage domain. To implement this, two integrators are utilized as shown in Figure 3.23. At the end of the *integration* phase the outputs of the both integrators are nothing but the integrated ΣI_{op} and ΣI_{on} during the *integration* phase:
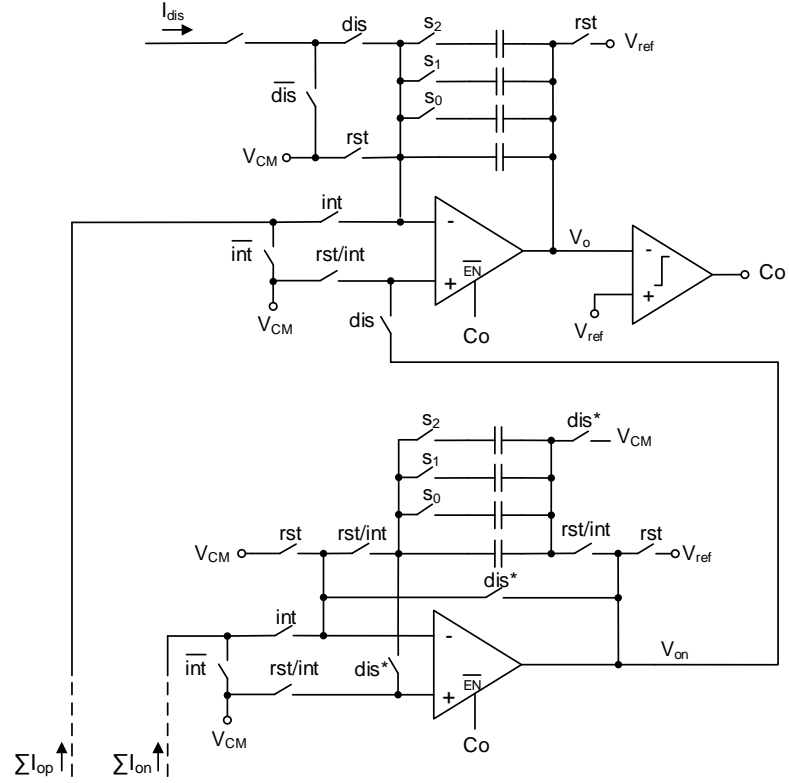

Figure 3.23: Schematic of the integrators. The integrated voltages are subtracted at the end of *integration* phase.

$$V_o(t_{int}) = V_{ref} - \frac{T_{int}}{C_{int}} \Sigma I_{op} \tag{3.11}$$

$$V_{on}(t_{int}) = V_{ref} - \frac{T_{int}}{C_{int}} \Sigma I_{on} \tag{3.12}$$

64

The dis^* signal is an advanced version of dis signal. Thus, right before starting the $discharge$ phase the integrator at the bottom of Figure 3.23 is reconfigured as a unity-gain buffer and V_{on} is changed to:

$$V_{on}(t_{int}{}^+) = 2V_{CM} - V_{on}(t_{int}) \tag{3.13}$$

and therefore, when dis signal is activated V_o becomes:

$$V_o(t_{int}{}^+) = V_o(t_{int}) + V_{on}(t_{int}{}^+) = V_{CM} + V_o(t_{int}) - V_{on}(t_{int}) \tag{3.14}$$

This design reduced the RMS Error by a factor of $\sim$2X (to 0.28%) without compromising the operation speed or energy-efficiency of the system.

3.7.4 Gain Error Calibration

As mentioned before, the first prototype included a set of extra transconductance amplifiers and floating-gate memories for output offset calibration. Figure 3.24(a) shows a modified offset calibration circuit for the revised prototype which utilizes the redesigned transconductance amplifier. In addition, it has been shown that being able to correct the gain error (scaling error) could further improve the accuracy [8] of the system. Figure 3.24(b) shows a circuit schematic implemented in the revised prototype for gain error calibration. A current source modulated by a floating-gate memory adds an extra component (I_g) to the discharge current (I_d). As a result, the actual discharge current becomes: $I_{dis} = I_d + I_g$.

3.8 Simulation Results (Revised Prototype)

The revised prototype of PWM-based CFS was designed and simulated in a 130 nm standard CMOS process. Figure 3.25 shows the layout of the system which is currently under fabrication. The layout occupies an active area of $0.975mm^2$. According to the post-layout simulations, PWM-based CFS dissipates an average power of 1.12 mW
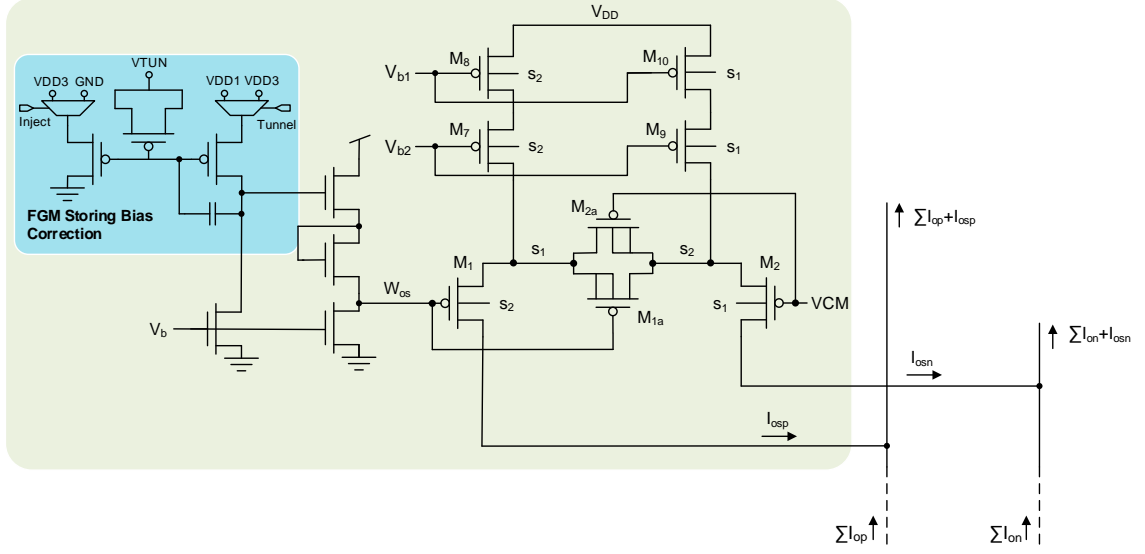
(a)
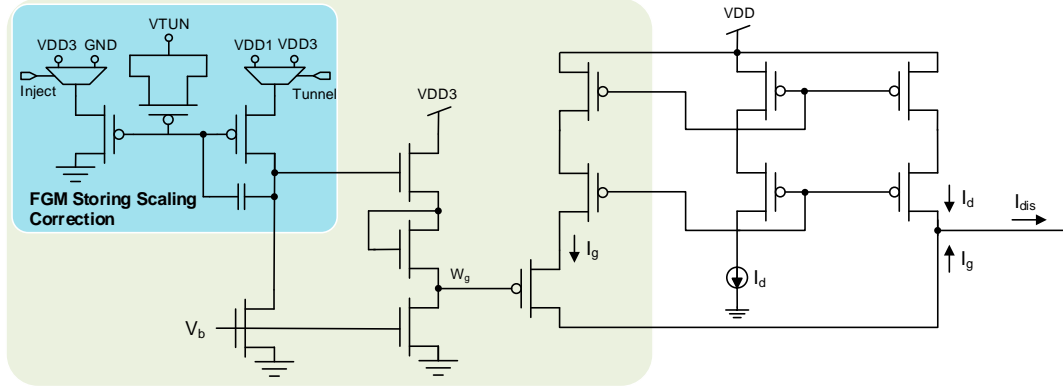
(b)

Figure 3.24: (a) Output offset (bias) calibration circuitry. The FGM modulates I_{osp} which is added to ΣI_{op}.(b) Output gain (scale) calibration circuitry. The FGM modulates I_g which is added to I_d to form I_{dis}.

at 200 MHz clock frequency. The system achieves 319 kVec/s throughput with an average energy efficiency of 3.6 pJ/MAC. Compared to Gilbert-multiplier-based CFS, the PWM-based CFS is more than 28 times faster and 7 times more energy efficient. Table 3.7 summarizes post-layout results of the revised PWM-based CFS. Figure 3.26 shows the distribution of the power consumption. The Digital I/O interface and the dynamic comparators consume 41.2% and 35.8% of the total power while the analog

processing blocks, i.e. the TC array and the integrators only dissipate 23% of the total power.

The typical current profile of the dynamic comparators and the integrators are presented in Figure 3.27. The comparators operate at full power during the calibration and conversion phases until they reach a convergence or decision which at that time they go to a standby mode with almost no power dissipation until the next cycle. The integrators also are shutdown following the lead of their associated comparators during the *discharge* phase.

In order to verify the functionality, the PWM-based CFS chip was simulated using actual filter coefficients and image pixel data. For this simulation, two vectors were chosen from the 2-D image array. The first one is chosen from the *background* area and the other one is selected from the *object* area (Figure 3.28). The filter coefficients are the vertical filter coefficients shown in Figure 2.22(b). The expected output for the vector containing the object is a valley while for the vector containing the background is zero (No significant peak or valley). The system was evaluated in Cadence ADE using the statistical Monte Carlo simulations. Figure 3.29(a) presents the simulation results of five Monte Carlo iterations. The system was operating at 1 V power supply and 250 MHz clock frequency. A post-layout simulation response at 1 V power supply and 200 MHz clock frequency for the same set of inputs are plotted in Figure 3.29(b). These simulation results substantiate the effectiveness of the system in discrimination
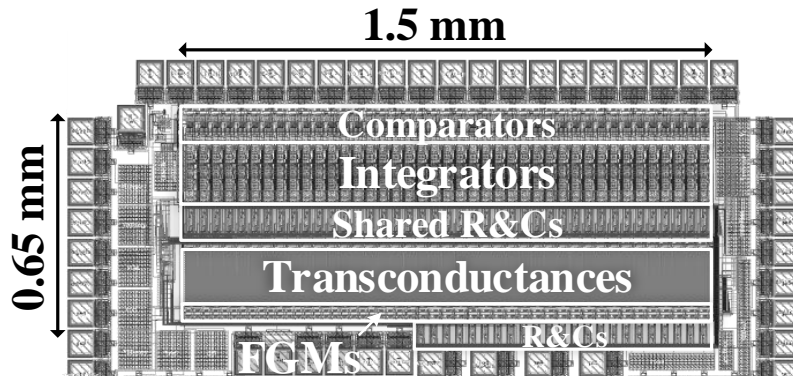

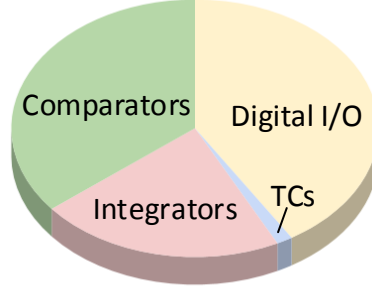Figure 3.25: The revised PWM-based CFS chip layout.

Figure 3.26: The power distribution of the revised PWM-based CFS chip.

of the object from the background. The system response appears to be resilient to process variations and layout parasitics (up to 200 MHz).

To further evaluate the system, a nominal simulation was performed on a 64×64 test image. The simulation procedure is the same as the test procedure explained in Section 2.6 for the Gilbert-multiplier-based CFS. Figure 3.30(a) shows
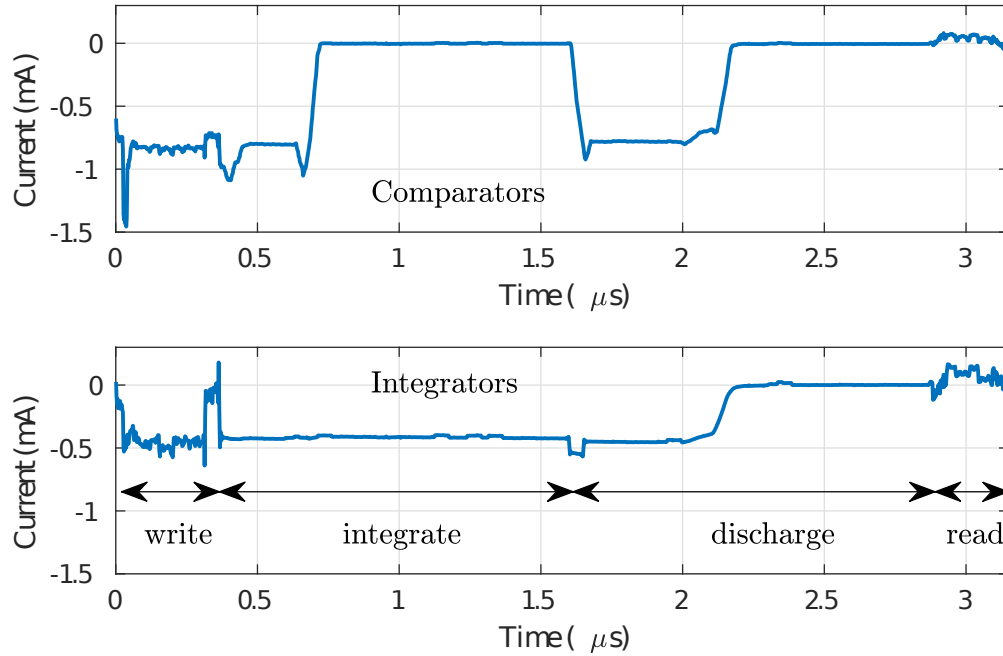

Figure 3.27: Power supply current profile of integrators and comparators during system operation.
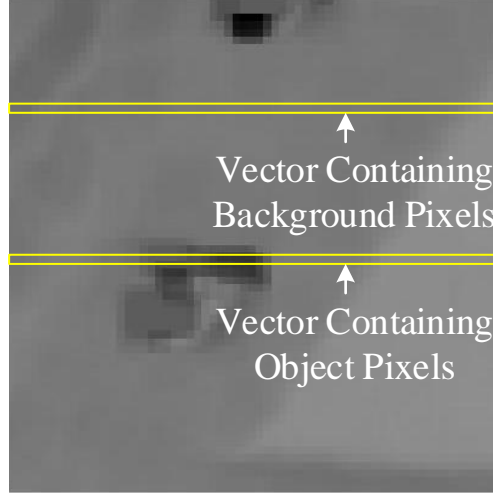
Figure 3.28: Selection of vectors from the input image for simulation.

the mathematical response from a CF implemented in MATLAB and Figure 3.30(b) displays the simulated response of PWM-based CFS chip. This response verifies the effectiveness of the system in the differentiation of the object from the background. This is in spite of the fact that the system output is saturated in this simulation. The

Figure 3.29: The output response of the PWM-based CFS to presence and absence of the target object is presented: (a) Five iterations of Monte Carlo simulation. the system operates at 250 MHz clock frequency. (b) Post-layout simulation results. the system operates at 200 MHz clock frequency.

Table 3.7: PWM-based CFS Post-Layout and Monte Carlo Simulation Results
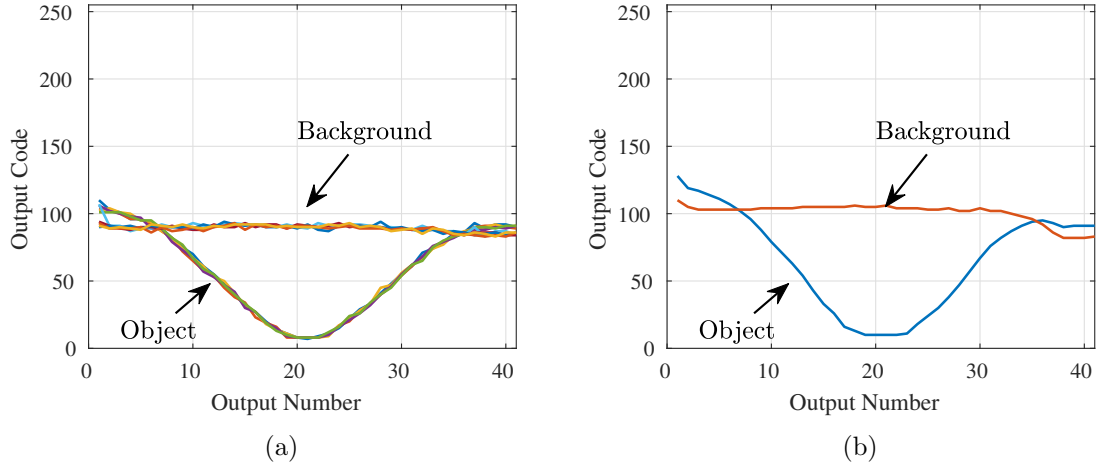
Specification	Value
Supply Voltage	1 V
Array Size	24×41
Layout Size (mm^2)	1.5×0.65
Tclk (ns)	5
Total time (μs)	3.13
Throughput ($kVec/s$)	319
Total average power diss. (mW)	1.12
Energy Efficiency (pJ/MAC)	3.6
Energy Efficiency (GOPS/W)	559
RMS Error (object/background)	0.71 LSBs (0.28%)

system parameters could be adjusted to prevent output saturation and to increase the discrimination even more.

3.9   Conclusion

A 24×41 PWM-based CFS is presented. Benefiting from a time-domain approach to multiplication, this system eliminates the need for any explicit D/A and A/D converters and takes advantage of an all-digital I/O interface. Careful utilization of the clock and available hardware resources in the digital I/O interface, along with the application of power management techniques have significantly reduced the circuit complexity and energy consumption of the system. Additionally, programmable transconductance amplifiers are incorporated at the output of the analog array for offset and gain error calibration. The prototype system occupies an area of $0.98mm^2$ and is expected to achieve an outstanding energy-efficiency of 3.6pJ/MAC at 319kVec/s with 0.28% RMS error.
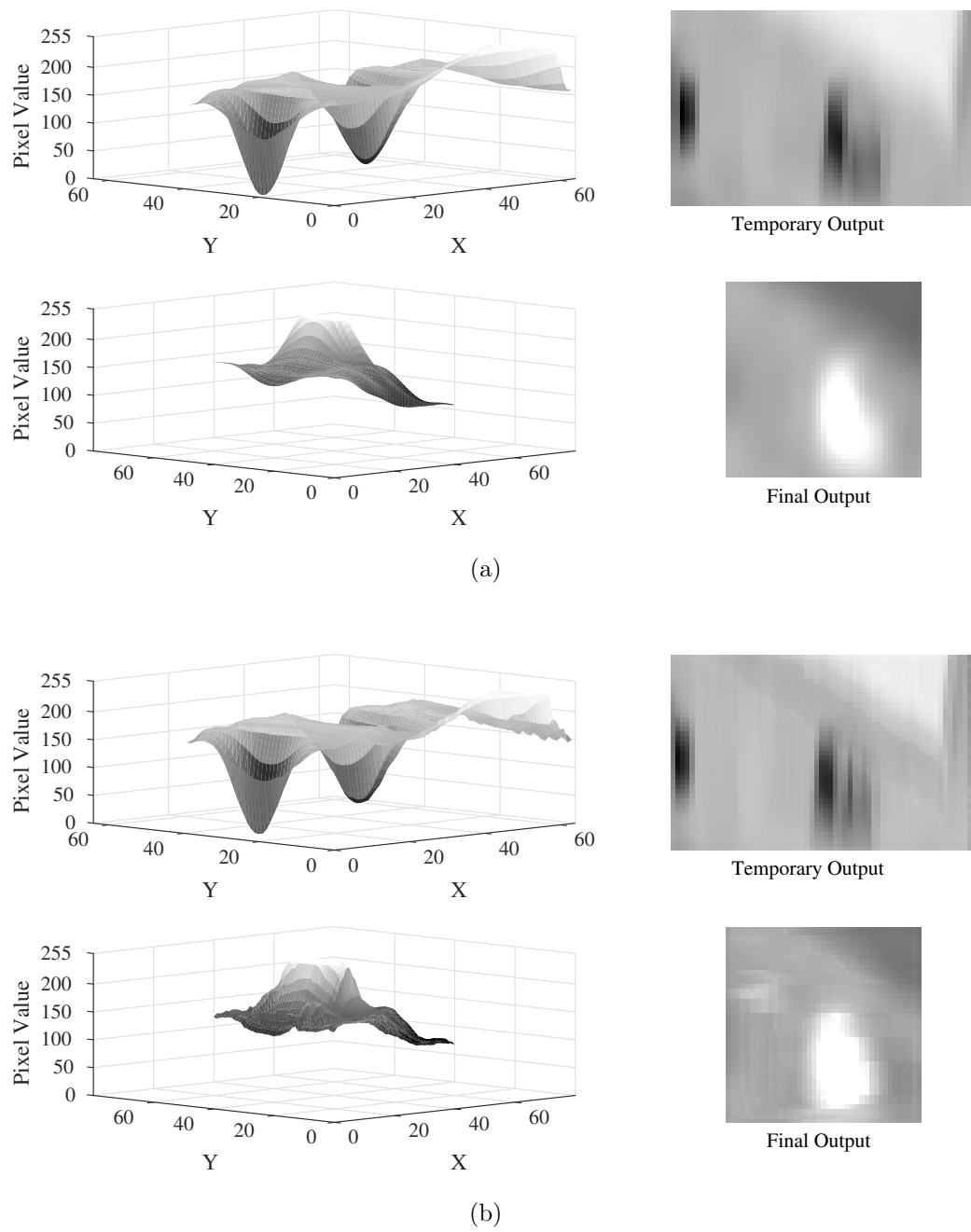
(a)

(b)

Figure 3.30: Nominal simulation of a 64×64 test image. (a) MATLAB implementation, and (b) simulated chip results. The correlation output is saturated at the object location.

71

Chapter 4

Conclusions and Future Work

This chapter summarizes this dissertation and proposes future work in this research area.

4.1   Conclusions

This dissertation investigates the effective implementation of analog signal processing systems with digital interfacing. In particular two architectures are proposed for a digitally-interfaced analog correlation filter system. While digital interfacing provided a standard and scalable way of communication with pre- and post-processing blocks without undermining the energy efficiency of the system, the multiply-accumulate operations were performed in analog. Moreover, non-volatile floating-gate memories are utilized as storage for coefficients. The proposed systems incorporate techniques to reduce the effects of analog circuit imperfections.

The first system implements a 24×57 Gilbert-multiplier-based correlation filter. The I/O interface is implemented with low-power D/A and A/D converters, and a correlated double sampling technique is implemented to reduce offset and low-frequency noise at the output of the analog array. The prototype chip occupies an area of 3.23 mm^2 and demonstrates a 25.2 pJ/MAC energy-efficiency at 11.3 kVec/s and 3.2% RMSE.

The second system realizes a 24×41 PWM-based correlation filter. Benefiting from a time-domain approach to multiplication, this system eliminates the need for explicit D/A and A/D converters. Careful utilization of clock and available hardware resources in the digital I/O interface, along with the application of power management techniques have significantly reduced the circuit complexity and energy consumption of the system. Additionally, programmable transconductance amplifiers are incorporated at the output of the analog array for offset and gain error calibration. The prototype system occupies an area of 0.98 mm^2 and is expected to achieve an outstanding energy-efficiency of 3.6 pJ/MAC at 319 kVec/s with 0.28% RMSE.

4.2 Future Work

Based on this dissertation, the following areas can be considered for future research.

- The energy efficiency can be further improved by considering the signal nature and statistics.

  First, if the correlation filter output fluctuates around a certain range for most of the time, the system could be designed to have its best energy performance around that range. For instance, by adjusting the design such that the dynamic comparator processes that range first.

  Second, if the vectors are highly correlated in the input data, it would be beneficial to redesign the system to process the difference between the consecutive vectors. To get the most out of this approach, known as Δ modulation, it would be necessary to make sure the post-processing blocks are fully compatible with this scheme.

- The proposed system could be integrated with pre- and post-processing blocks on a chip to form a real-time object tracking system.

- The proposed architecture could be applied to other signal processing applications. For example vector-matrix multiplication (VMM) systems and classifiers.

Bibliography

[1] Peterson R Agostinho, Sandro AP Haddad, Jader A De Lima, and Wouter A Serdijn. An ultra low power CMOS pA/V transconductor and its application to wavelet filters. *Analog Integrated Circuits and Signal Processing*, 57(1-2):19–27, 2008. 56

[2] Erkan Alpman, Hasnain Lakdawala, L Richard Carley, and Krishnamurthy Soumyanath. A 1.1 v 50mw 2.5 GS/s 7b Time-Interleaved C-2C SAR ADC in 45nm lp Digital CMOS. In *2009 IEEE International Solid-State Circuits Conference-Digest of Technical Papers*, pages 76–77. IEEE, 2009. 44

[3] Alfredo Arnaud, Rafaella Fiorelli, and Carlos Galup-Montoro. Nanowatt, sub-nS OTAs, with sub-10-mV input offset, using series-parallel current mirrors. *IEEE Journal of Solid-State Circuits*, 41(9):2009–2018, 2006. 56

[4] Alfredo Arnaud and Carlos Galup-Montoro. Fully integrated signal conditioning of an accelerometer for implantable pacemakers. *Analog Integrated Circuits and Signal Processing*, 49(3):313–321, 2006. 56

[5] Amer Aslam-Siddiqi, Werner Brockherde, and Bedrich J Hosticka. A 16× 16 nonvolatile programmable analog vector-matrix multiplier. *IEEE Journal of Solid-State Circuits*, 33(10):1502–1509, 1998. 4

[6] Faramarz Bahmani and Edgar Sánchez-Sinencio. A highly linear pseudo-differential transconductance [CMOS OTA]. In *Solid-State Circuits Conference, 2004. ESSCIRC 2004. Proceeding of the 30th European*, pages 111–114. IEEE, 2004. 56

[7] BBC. Meet the world's most powerful computer, 2017. 2

[8] D. Bolme, A. Mikkilineni, D. Rose, S. Yoginath, M. Judy, and J. Holleman. Deep modeling: Circuit characterization using theory based models in a data driven framework. In *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–4, May 2017. 65

[9] David S Bolme, J Ross Beveridge, Bruce A Draper, and Yui Man Lui. Visual object tracking using adaptive correlation filters. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2544–2550. IEEE, 2010. 1, 2, 31

[10] David S Bolme, Bruce A Draper, and J Ross Beveridge. Average of synthetic exact filters. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2105–2112. IEEE, 2009. 1, 31

[11] Shantanu Chakrabartty and Gert Cauwenberghs. Sub-microwatt analog VLSI trainable pattern classifier. 42(5):1169–1179, 2007. 4

[12] Ravi Chawla, Abhishek Bandyopadhyay, Venkatesh Srinivasan, and Paul Hasler. A 531 nW/MHz, 128× 32 current-mode programmable analog vector-matrix multiplier with over two decades of linearity. In *Custom Integrated Circuits Conference, 2004. Proceedings of the IEEE 2004*, pages 651–654. IEEE, 2004. 30

[13] Zhiming Chen, Yuanjin Zheng, Foo Chung Choong, and Minkyu Je. A low-power variable-gain amplifier with improved linearity: Analysis and design. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 59(10):2176–2185, 2012. 56

[14] Robert Collins, Xuhui Zhou, and Seng Keat Teh. An open source tracking testbed and evaluation web site. In *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, volume 35, 2005. 32

[15] Z El-Khatib, L MacEachern, and SA Mahmoud. Highly-linear CMOS cross-coupled compensator transconductor with enhanced tunability. *Electronics letters*, 46(24):1597–1598, 2010. 56

[16] Ahmed A Emira and Edgar Sánchez-Sinencio. A pseudo differential complex filter for Bluetooth with frequency tuning. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 50(10):742–754, 2003. 56

77

[17] Christian C Enz and Gabor C Temes. Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization. *Proceedings of the IEEE*, 84(11):1584–1614, 1996. 22

[18] Joel Gak, Matias R Miguez, and Alfredo Arnaud. Nanopower OTAs with improved linearity and low input offset using bulk degeneration. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 61(3):689–698, 2014. 56

[19] Roman Genov, Gert Cauwenberghs, et al. Charge-mode parallel architecture for vector-matrix multiplication. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 48(10):930–936, 2001. 5, 30

[20] Joël MH Karel, Sandro AP Haddad, Senad Hiseni, Ronald L Westra, Wouter A Serdijn, and Ralf LM Peeters. Implementing wavelets in continuous-time analog circuits with dynamic range optimization. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 59(2):229–242, 2012. 56

[21] Peter R Kinget. Device mismatch and tradeoffs in the design of analog circuits. *IEEE Journal of Solid-State Circuits*, 40(6):1212–1224, 2005. 44

[22] Keisuke Korekado, Takashi Morie, Osamu Nomura, Teppei Nakano, Masakazu Matsugu, and Atsushi Iwata. An image filtering processor for face/object recognition using merged/mixed analog-digital architecture. In *Digest of Technical Papers. 2005 Symposium on VLSI Circuits, 2005.*, pages 220–223. IEEE, 2005. 6, 30

[23] Francois Krummenacher and Norbert Joehl. A 4-MHz CMOS continuous-time filter with on-chip automatic tuning. *IEEE Journal of Solid-State Circuits*, 23(3):750–758, 1988. 13, 56

[24] BVK Vijaya Kumar, Joseph A Fernandez, Andres Rodriguez, and Vishnu Naresh Boddeti. Recent advances in correlation filter theory and application. In *SPIE*

Defense+ Security, pages 909404–909404. International Society for Optics and Photonics, 2014. 1

[25] Yen-Ting Liu, Donald YC Lie, Weibo Hu, and Tam Nguyen. An ultralow-power CMOS transconductor design with wide input linear range for biomedical applications. In *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pages 2211–2214. IEEE, 2012. 56

[26] Junjie Lu and Jeremy Holleman. A floating-gate analog memory with bidirectional sigmoid updates in a standard digital process. In *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, pages 1600–1603. IEEE, 2013. 19

[27] Junjie Lu and Jeremy Holleman. A low-power high-precision comparator with time-domain bulk-tuned offset cancellation. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(5):1158–1167, 2013. 44

[28] Abhijit Mahalanobis, Robert R Muise, S Robert Stanfill, and ALAN Van Nevel. Design and application of quadratic correlation filters for target detection. *IEEE Transactions on Aerospace and Electronic Systems*, 40(3):837–850, 2004. 1

[29] JM Algueta Miguel, CA De La Cruz Blas, and AJ Lopez-Martin. CMOS triode transconductor based on quasi-floating-gate transistors. *Electronics letters*, 46(17):1190–1191, 2010. 56

[30] Jose Maria Algueta Miguel, Antonio J Lopez-Martin, Lucia Acosta, Jaime Ramirez-Angulo, and Ramón Gonzalez Carvajal. Using floating gate and quasi-floating gate techniques for rail-to-rail tunable CMOS transconductor design. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 58(7):1604–1614, 2011. 56

[31] Masaya Miyahara, Yusuke Asada, Daehwa Paik, and Akira Matsuzawa. A low-noise self-calibrating dynamic comparator for high-speed ADCs. In *Solid-State*

Circuits Conference, 2008. A-SSCC'08. IEEE Asian, pages 269–272. IEEE, 2008. 45, 47, 49, 52, 53

[32] A Nader Mohieldin, Edgar Sánchez-Sinencio, and José Silva-Martínez. A fully balanced pseudo-differential OTA with common-mode feedforward and inherent common-mode feedback detector. *IEEE journal of solid-state circuits*, 38(4):663–668, 2003. 56

[33] Nicholas Conley Poore. Digital-to-analog converter interface for computer assisted biologically inspired systems. Master's thesis, University of Tennessee, Knoxville, aug 2014. 13

[34] Shubha Ramakrishnan, Jennifer Hasler, et al. Vector-matrix multiply and winner-take-all as an analog classifier. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 22(2):353–361, 2014. 4, 30

[35] Jaime Ramirez-Angulo, Ramón González Carvajal, Juan A Galán, and Antonio López-Martín. A free but efficient low-voltage class-AB two-stage operational amplifier. *IEEE Trans. on Circuits and Systems*, 53(7):568–571, 2006. 24

[36] Behzad Razavi. *Principles of data conversion system design*, volume 126. IEEE press New York, 1995. 44

[37] Christopher D Salthouse and Rahul Sarpeshkar. A practical micropower programmable bandpass filter for use in bionic ears. *IEEE Journal of Solid-State Circuits*, 38(1):63–70, 2003. 56

[38] R. Sarpeshkar, T. Delbruck, and C. A. Mead. White noise in MOS transistors and resistors. *IEEE Circuits and Devices Magazine*, 9(6):23–29, Nov 1993. 16

[39] Rahul Sarpeshkar. Analog versus digital: extrapolating from electronics to neurobiology. *Neural computation*, 10(7):1601–1638, 1998. 4

[40] Rahul Sarpeshkar, Richard F Lyon, and Carver Mead. A low-power wide-linear-range transconductance amplifier. *Analog Integrated Circuits and Signal Processing*, 13(1):123–151, 1997. 56

[41] Michael D Scott, Bernhard E Boser, and Kristofer SJ Pister. An ultralow-energy ADC for smart dust. *IEEE Journal of Solid-State Circuits*, 38(7):1123–1129, 2003. 26

[42] Sergio Solis-Bustos, José Silva-Martínez, Franco Maloberti, and Edgar Sánchez-Sinencio. A 60-dB dynamic-range CMOS sixth-order 2.4-Hz low-pass filter for medical applications. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 47(12):1391–1398, 2000. 56

[43] Geert Van der Plas, Stefaan Decoutere, and Stephane Donnay. A 0.16 pJ/conversion-step 2.5 mW 1.25 GS/s 4b ADC in a 90nm digital CMOS process. In *2006 IEEE International Solid State Circuits Conference-Digest of Technical Papers*, 2006. 44

[44] Anand Veeravalli, Edgar Sánchez-Sinencio, and José Silva-Martínez. Transconductance amplifier structures with very small transconductances: A comparative design approach. *IEEE Journal of Solid-State Circuits*, 37(6):770–775, 2002. 56

[45] Alice Wang, Benton H Calhoun, and Anantha P Chandrakasan. *Sub-threshold design for ultra low-power systems*, volume 95. Springer, 2006. 13

Vita

Mohsen Judy received his B.S. and M.S. degrees from the K. N. Toosi University of Technology, Tehran, Iran, in 2008 and 2011, respectively, all in Electrical Engineering. He was with Yekta-Fanavar-Samaneh Company during 2011-12 and with Maharan Engineering Corporation between 2012-13, as a Design Engineer. He started his Ph.D. studies in the Department of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville in 2013. His research interests include low-power and high-performance analog and mixed-signal integrated circuits and biomedical circuits and systems.