UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROELETRÔNICA

OSCAR ELISIO MATTIA NETO

NanoWatt Resistorless CMOS Voltage References for Sub-1 V Applications

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Microeletronics

Advisor: Prof. Dr. Hamilton Klimach
Coadvisor: Prof. Dr. Sergio Bampi

Porto Alegre
October 2014

CIP – CATALOGING-IN-PUBLICATION

Mattia Neto, Oscar Elisio

NanoWatt Resistorless CMOS Voltage References for Sub-1 V Applications / Oscar Elisio Mattia Neto. – Porto Alegre: PGMICRO da UFRGS, 2014.

87 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Microeletrônica, Porto Alegre, BR–RS, 2014. Advisor: Hamilton Klimach; Coadvisor: Sergio Bampi.

1. CMOS analog design. 2. Bandgap voltage reference. 3. Curvature compensation. 4. Resistorless. 5. Ultra-low-power. 6. Low voltage design. I. Klimach, Hamilton. II. Bampi, Sergio. III. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. Carlos Alexandre Netto
Vice-Reitor: Prof. Rui Vicente Oppermann
Pró-Reitor de Pós-Graduação: Prof. Vladimir Pinheiro do Nascimento
Diretor do Instituto de Informática: Prof. Luis da Cunha Lamb
Coordenador do PGMICRO: Prof. Gilson Inácio Wirth
Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

"I get a kick out of thinking about these things."
— RICHARD FEYNMAN

ACKNOWLEDGEMENTS

This work would not be possible without the influence of many people, and I feel grateful to have worked and lived with them for the past year. A BIG thank you to my advisor prof. Hamilton, for teaching me most of what I know about integrated circuit design. For patiently reading and correcting my manuscripts. For the lengthy discussions we've had after the day was done. Also to my co-advisor prof. Bampi, for being always available and willing to help with matters academic and bureaucratic. And as an example of seriousness and commitment to work.

To prof. Marcio Schneider from UFSC, for receiving me in Florianópolis and helping me develop an important part of this thesis. And other circuits. To prof. André Aita from UFSM, for the early work feedback, availability and recommendations. To prof. Luis Fernando Ferreira, for teaching me how to use and maintain a micro-probe station.

To my CI-BRASIL and laboratory colleagues, many of whom became good friends. I learned much and had a great time with all of you. I'll write down the names already knowing that I'll forget a lot of people - John, Helga, Jeffer, Pablo, Alberto, Paulo, Dalton, Valmiro, Diogo, Pedro, Alonso, German, David, Julio, Israel, Moacir, Arthur, Sandro, Leandro, Marlon, Egas, Paulo, Renato.

To my friends for reminding me that there are other stuff than electronics out there. And to my family, for everything.

ABSTRACT

Integrated voltage references have always been a fundamental block of any electronic system, and an important research topic that has been extensively studied in the past 50 years. A voltage reference is a circuit that provides a stable voltage with low sensitivity to variations in temperature, supply, load, process characteristics and packaging stresses. They are usually implemented through the weighted sum of two independent physical phenomena with opposite temperature dependencies. Usually the thermal voltage, related to the Boltzmann's constant and the electron charge, provides a positive temperature dependence, while the silicon bandgap voltage or a MOSFET's threshold voltage provide the complementary term. An auxiliary biasing block is sometimes necessary to provide the necessary currents for the circuit to work, and additional blocks implement the weighted sum. The scaling of process technologies is the main driving factor for low voltage operation, while the emergence of portable battery-operated, implantable biomedical and energy harvesting devices mandate that every circuit consume as little power as possible. Therefore, sub-1 V supplies and nanoWatt power have become key characteristics for these kind of circuits, but there are several challenges when designing high accuracy voltage references in modern CMOS technologies under these conditions. The traditional topologies are not suitable because they provide a reference voltage above 1 V, and to achieve such power consumption levels would require GΩ resistances, that occupy a huge silicon area. Recent advances have achieved these levels of power consumption but with limited accuracy, expensive calibration procedures and large silicon area. In this thesis we present two novel circuit topologies, a resistorless, self-biased and curvature compensated bipolar emitter voltage, and a resistorless bandgap voltage reference for sub-1 V supply (also called sub-bandgap). Both circuits operate in the nanoWatt range and occupy small silicon areas. Simulation results from two different processes, 180 nm and 130 nm, and experimental results from one fabrication run in 130 nm show improvements over such drawbacks, while maintaining the desired characteristics of being resistorless, ultra-low-power, low-voltage operated and occupying very small areas.

Keywords: CMOS analog design, bandgap voltage reference, curvature compensation, resistorless, ultra-low-power, low voltage design.

Referências de Tensão CMOS em NanoWatts e Sem Resistores para Aplicações em Sub-1 V

RESUMO

Referências de tensão integradas sempre foram um bloco fundamental de qualquer sistema eletrônico e um importante tópico de pesquisa que tem sido estudado extensivamente nos últimos 50 anos. Uma tensão de referência é um circuito que provê uma tensão estável com baixa sensibilidade a variações em temperatura, alimentação, carga, características do processo de fabricação e tensões mecânicas de encapsulamento. Elas são normalmente implementadas através da soma ponderada de dois fenômenos físicos diferentes, com comportamentos em temperatura opostos. Normalmente, a tensão térmica, relacionada à constante de Boltzmann e à carga do elétron, fornece uma dependência positiva com temperatura, enquanto que a tensão base-emissor V_{BE} de um transistor bipolar ou a tensão de limiar de um MOSFET fornece o termo complementar. Um bloco auxiliar é às vezes utilizado para fornecer as correntes de polarização do circuito, e outros blocos adicionais implementam a soma ponderada. A evolução da tecnologia de processos é o principal fator para aplicações em baixa tensão, enquanto que a emergência de dispositivos portáteis operados a bateria, circuitos biomédicos implantáveis e dispostivos de captura de energia do ambiente restringem cada circuito a consumir o mínimo possivel. Portanto, alimentações abaixo de 1 V e consumos na ordem de nanoWatts se tornaram características fundamentais de tais circuitos. Contudo, existem diversos desafios ao projetar referências de tensão de alta exatidão em processos CMOS modernos sob essas condições. As topologias tradicionais não são adequadas pois elas provêm uma referência de tensão acima de 1 V, e requerem resistências da ordem de GΩ para atingir tão baixo consumo de potência, ocupando assim uma grande área de silício. Avanços recentes atingiram tais níveis de consumo de potência, porém com limitada exatidão, custosos procedimentos de calibração e grande área ocupada em silício. Nesta dissertação apresentam-se duas novas topologias de circuitos: uma tensão de junção bipolar com compensação de curvatura que não utiliza resistores e é auto-polarizada; e um circuito de referência *bandgap* sem resistores que opera abaixo de 1 V (também chamado de *sub-bandgap*). Ambos circuitos operam com consumo na ordem de nanoWatts e ocupam pequenas áreas de silício. Resultados de simulação para dois processos diferentes, 180 nm e 130 nm, e resultados experimentais de uma rodada de fabricação em 130 nm apresentam melhorias sobre tais limitações, mantendo as características desejadas de não conter resistores, ultra baixo consumo, baixa tensão de alimentação e áreas muito pequenas.

Palavras-chave: projeto analógico CMOS, referência de tensão bandgap, sem resistoress, ultra baixo consumo, baixa alimentação.

LIST OF FIGURES

1.1 Modern communications System on Chip (SoC) DEVICES (2013). . 14
1.2 Load profile of a duty-cycled wireless micro-sensor TORRES; MILNER; RINCON-MORA (2008). 15
1.3 Basic concept of a bandgap voltage reference. 16
1.4 Output voltage deviation of the reference voltage plotted versus temperature for a conventional circuit and a curvature-corrected circuit with the linearized thermal behavior of $V_{BE}(T)$ MEIJER; SCHMALE; VAN ZALINGE (1982). 17

2.1 Classic Widlar Bandgap Voltage Reference WIDLAR (1971). 20
2.2 Kuijk Bandgap Voltage Reference KUIJK (1973). 21
2.3 Brokaw Bandgap Voltage Reference BROKAW (1974). 21
2.4 Tsividis CMOS Bandgap Voltage Reference TSIVIDIS; ULMER (1978). 22
2.5 Vittoz CMOS Bandgap Voltage Reference Block Diagram VITTOZ; NEYROUD (1979). 22
2.6 Stack of elementary PTAT self-cascode cells VITTOZ; NEYROUD (1979). 22
2.7 A curvature corrected bandgap reference with the linearized thermal behavior of $V_{BE}(T)$ MEIJER; SCHMALE; VAN ZALINGE (1982). . 23
2.8 Curvature-compensation concept (not scaled) SONG; GRAY (1983). 23
2.9 (a) Simplified circuit of a conventional voltage reference based on the difference between the gate-source voltages of two MOS transistors. (b) Simplified circuit of the voltage reference proposed in LEUNG; MOK (2003). 25
2.10 (a) Simplified circuit of the voltage reference proposed in DE VITA; IANNACCONE (2005). (b) Simplified circuit of the voltage reference proposed in DE VITA; IANNACCONE; ANDREANI (2006). . 26
2.11 Proposed voltage reference circuit DE VITA; IANNACCONE (2007). 27
2.12 Proposed voltage reference UENO et al. (2009). 29
2.13 Reference voltage and its TC versus dopant concentration UENO et al. (2009). 30
2.14 Entire circuit of the voltage reference. All MOSFETs are in WI, except for M_{R1} UENO et al. (2009). 31
2.15 Bandgap reference circuit configuration MING et al. (2010). 32
2.16 Scheme of principle of a simple voltage reference MAGNELLI et al. (2011). 33
2.17 Proposed CMOS voltage reference circuit MAGNELLI et al. (2011). 34

2.18 Distributions of the most relevant figures of merit for the 40 measured samples: (a) Power consumption @ 25 ○C and V_{DD} = 0.45 V. (b) Line sensitivity @ 25 °C. (c) Temperature coefficient. (d) Generated voltage reference @ 25 °C MAGNELLI et al. (2011). 36
2.19 Power and area comparisons of recently published voltage references SEOK et al. (2012). 36
2.20 Schematic of proposed 2T voltage reference SEOK et al. (2012). . . . 37
2.21 Measurement Summary of the Proposed 2T Voltage Reference and its Variants. 38
2.22 Schematic of trimmable 2T voltage reference (L = 60um is used) SEOK et al. (2012). 38
2.23 Measured (a) TC and (b) output voltage dependency on trim settings SEOK et al. (2012). 39
2.24 Measured reductions of output voltage and temperature coefficient spreads after assisted one temperature point trimming in 0.13um SEOK et al. (2012). 39
2.25 Schematic of 4T voltage reference SEOK et al. (2012). 40
2.26 BGR Reference Circuit OSAKI et al. (2013). 41
2.27 Sub-BGR Reference Circuit OSAKI et al. (2013). 42

3.1 Bandgap voltage versus absolute temperature and its first degree approximation (not to scale) TSIVIDIS (1980). 44
3.2 Different δ values comparison. (a) V_{BE}; (b) first derivative; and (c) I_C versus temperature. 47
3.3 Resulting V_{REF} for different δ values comparison. 47
3.4 Impact of the design parameters variability on V_{REF}. (a) $V_{BE}(T_r)$; (b) δ and (c) n, for $\sigma/\mu = 1\%$. 48

4.1 Schematic of the bias circuit concept. 51
4.2 Schematic of different bias circuit examples. (a) $N = 2$; (b) $N = 3$. 51
4.3 Behavior over temperature. (a) V_{E1} and V_{E2}; (b) $\partial V_{E1}/\partial T$ and $\partial V_{E2}/\partial T$; (c) I_{E1} and I_{E2}. 52
4.4 Self-biased curvature compensation circuit. 53
4.5 Analytical and SPICE results comparison versus temperature: (a) junction voltage; (b) first derivative; (c) collector current 54
4.6 SPICE results comparing V_{REFs} with different junction bias currents. 54
4.7 Schematic of the proposed sub-BGR circuit. 55
4.8 Design space exploration (a) V_E; (b) I_E; and (c) I_{D1}. 57

5.1 Micrograph of the overall chip fabricated in IBM 130 nm with an area of 2.5 x 2.5 mm^2. 58
5.2 Self-biased curvature compensation circuit. 59
5.3 Layout of the proposed bias and curvature compensation circuit in IBM 130nm. 60
5.4 (a) V_E; and (b) I_E over temperature, V_{DD} = 0.8 V. 60
5.5 (a) PSRR versus frequency; (b) V_E and I_E versus power supply, 27°C. 60
5.6 V_E noise versus frequency at 27 °C and V_{DD} = 0.8 V. 61

5.7 V_E Monte Carlo results from 100 runs. Average process variation on (a); Local random mismatch on (b); and Combined analysis on (c). V_{DD} = 0.8 V. 61
5.8 Detail of the fabricated chip and the proposed BJT bias circuit area. . 62
5.9 V_E vs. V_{DD} @ 21°C. 63
5.10 V_E vs. V_{DD} @ 21°C, from 0.7 V to 1.2 V. 63
5.11 V_E vs. V_{DD}, Mean and Std. Dev. @ 21°C. 63
5.12 V_E vs. Temperature, with varying V_{DD}. 64
5.13 I_{TOTAL} vs. Temperature, with varying V_{DD}. 64
5.14 Schematic of the proposed sub-BGR circuit. 65
5.15 Layout floorplanning of the proposed sub-bandgap reference in XFAB 180nm. 65
5.16 (a) V_{REF}; (b) V_E, V_{GS1} and V_{PTAT} voltages; (c) I_{TOTAL}, I_E and I currents over temperature. V_{DD} = 0.9 V. 66
5.17 V_{REF} noise vs. frequency. 66
5.18 (a) PSRR versus frequency; (b) TC_{EFF}; (c) V_{REF} and I_{TOTAL} vs. V_{DD}. 67
5.19 V_{REF} and TC_{EFF} Monte Carlo results. Average process variation on (a) and (b); Local mismatch on (c) and (d); 67
5.20 Fabricated layout of the proposed sub-bandgap reference in IBM 130nm. 68
5.21 Detail of the fabricated chip and the proposed sub-BGR circuit area. 69
5.22 V_E vs. V_{DD} @ 21°C. 69
5.23 V_E vs. V_{DD} @ 21°C, from 0.9 V to 1.2 V. 69
5.24 V_E Mean and Std. Dev. vs. V_{DD} @ 21°C. 70
5.25 V_{REF} vs. V_{DD} @ 21°C. 70
5.26 V_{REF} vs. V_{DD} @ 21°C, from 0.9 V to 1.2 V. 70
5.27 V_{REF} Mean and Std. Dev. vs. V_{DD} @ 21°C. 71
5.28 I_{TOTAL} vs. V_{DD} @ 21°C. 71
5.29 I_{TOTAL} vs. V_{DD} @ 21°C, from 0.9 V to 1.2 V. 71
5.30 I_{TOTAL} Mean and Std. Dev. vs. V_{DD} @ 21°C. 72
5.31 V_{REF} vs. Temperature, with varying V_{DD} - Die # 1. 72
5.32 V_{REF} vs. Temperature, with varying V_{DD} - Die # 2. 72
5.33 V_{REF} vs. Temperature, with varying V_{DD} - Die # 3. 73
5.34 TC_{EFF} vs. V_{DD} - Dies # 1-3. 73
5.35 I_{TOT} vs. Temperature, with varying V_{DD} - Dies # 1-3. 73

LIST OF TABLES

2.1 Reported results summary in DE VITA; IANNACCONE (2007). 28
2.2 Typical condition experimental results summary UENO et al. (2009). 31
2.3 Performance summary MING et al. (2010). 32
2.4 Reported results under typical conditions MAGNELLI et al. (2011). . 35
2.5 Statistical Analysis of Performance MAGNELLI et al. (2011). 35
2.6 Variability Results (49 dies, two runs) 38
2.7 Performance summary of OSAKI et al. (2013). 40
2.8 Summary of recent resistorless CMOS Voltage References. 43

3.1 Linearity comparison for different δs. 48
3.2 Yield (%) for δ = 1. 48
3.3 Yield (%) for δ = 2. 49
3.4 Yield (%) for δ = 3.5. 49

4.1 Linearity performance (TC_{EFF}) of analytical model and SPICE results. 55

5.1 Performance comparison of the BJT Bias simulated circuits. 62
5.2 Performance comparison of the Sub-BGR simulated circuits. 68

6.1 Comparison of recent resistorless CMOS Voltage References. 75

LIST OF ABBREVIATIONS AND ACRONYMS

BGR	Bandgap Voltage Reference
BJT	Bipolar Junction Transistor
CMOS	Complementary Metal Oxide Semiconductor
CTAT	Complementary to Absolute Temperature
DTMOST	Dynamic Threshold MOS Transistor
DUT	Device Under Test
LS	Line Sensitivity
MI	Moderate Inversion
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
PSRR	Power Supply Rejection Ratio
PTAT	Proportional to Absolute Temperature
SI	Strong Inversion
TC	Temperature Coefficient
VR	Voltage Reference
WI	Weak Inversion

CONTENTS

1 INTRODUCTION 14
1.1 Motivation 14
1.2 Low Voltage and Low Power Aspects 15
1.3 Fabrication Process Aspects 15
1.4 Ideal Voltage Reference 16
1.5 Objectives 17
1.6 Organization 18

2 INTEGRATED VOLTAGE REFERENCES OVERVIEW 19
2.1 Classic References 19
2.2 Definitions and Metrics 24
2.3 Recent Developments on Resistorless CMOS Voltage References 25
2.3.1 nMOS Diode in Strong Inversion 25
2.3.2 Self-Cascode Cells and Diode-Connected nMOS 28
2.3.3 Polynomial Current Curvature Compensation Scheme for BGR 31
2.3.4 nMOS Diode in Weak Inversion 32
2.3.5 2-Transistor Voltage Reference 36
2.3.6 Junction Divider and Unbalanced Differential Pair 40
2.3.7 Summary of Recent Advances 41

3 BIPOLAR JUNCTION TRANSISTOR CHARACTERISTICS 44
3.1 BJT Temperature Model 44
3.2 BJT Curvature Correction 46
3.3 Fabrication Variability 47

4 PROPOSED CIRCUITS 50
4.1 BJT Bias and Curvature Compensation 50
4.1.1 Circuit Concept 50
4.1.2 Proposed Circuit 52
4.2 Sub-Bandgap Voltage Reference 54
4.2.1 Self-Cascode PTAT Generator 55
4.2.2 V_E Divider and BJT Bias Circuit 56
4.2.3 Reference Voltage 56
4.2.4 Design Methodology 57

5 SIMULATION AND MEASUREMENT RESULTS 58
5.1 BJT Bias and Curvature Compensation 58
5.1.1 Simulation Results 59
5.1.2 Experimental Results 61
5.2 Sub-Bandgap Voltage Reference 64
5.2.1 Simulation Results 64
5.2.2 Experimental Results 67

6 CONCLUSION 74
6.1 Future Work 75

REFERENCES 77

APPENDIX A LIST OF PUBLICATIONS 80

APPENDIX B ACM MOSFET MODEL 81

APPENDIX C SUMMARY IN PORTUGUESE 82
C.1 Introdução 82
C.1.1 Motivação 82
C.1.2 Aspectos de Baixa Tensão e Baixo Consumo 82
C.1.3 Aspectos do Processo de Fabricação 83
C.1.4 Tensão de Referência Ideal 83
C.1.5 Objetivos 84
C.1.6 Organização 84
C.2 Resumo do Texto 85
C.3 Conclusão 85
C.3.1 Trabalhos Futuros 86

1 INTRODUCTION

In this chapter we contextualize the usage of voltage references in electronics systems, and present the driving necessities of current applications and the restrictions they impose on the circuit design. We define what would be an ideal voltage reference and present the basic concept for temperature compensation, that is the most important performance parameter, while pointing to solutions used to improve the other metrics. We then state the main objectives of the thesis and describe the structure of the work.

1.1 Motivation

Voltage references are a fundamental part of most, if not all, electronic circuits and systems. They are extensively used in analog, mixed-signal, radio-frequency and even digital circuits such as memories, and the required accuracy varies widely across the different application domains. To illustrate their application, Figure 1.1 shows a modern System on Chip (SoC) from Analog Devices that integrates a microprocessor, a precision data acquisition unit, where a voltage reference is explicitly shown, and an RF transceiver front-end. There are other voltage and/or current references - not shown - distributed throughout the system, that provide local references for comparators or biasing, for example.

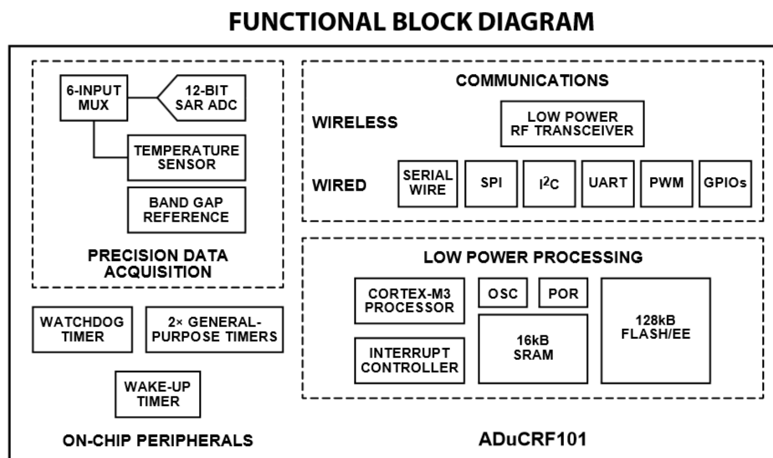


Figure 1.1: Modern communications System on Chip (SoC) DEVICES (2013).

1.2 Low Voltage and Low Power Aspects

The continuing scaling of CMOS technologies is the main driving factor behind low-voltage operation, rapidly achieving sub-1 V supplies for process nodes below 130 nm. Current battery-operated and future self-powered and self-sustaining systems require ultra-low-power operation that ranges from a few nanoamperes to a few microamperes depending on the function being executed, as illustrated by the load profile of duty-cycled wireless micro-sensors in Figure 1.2. Even large SoCs like the one of Figure 1.1 have stand-by power consumption of 280 nA in power-down mode and 1.9 μA with processor and transceiver memory retention. The voltage reference block is usually kept turned on because it provides a switching threshold for power-on reset circuits and an accurate reference for supply regulation.

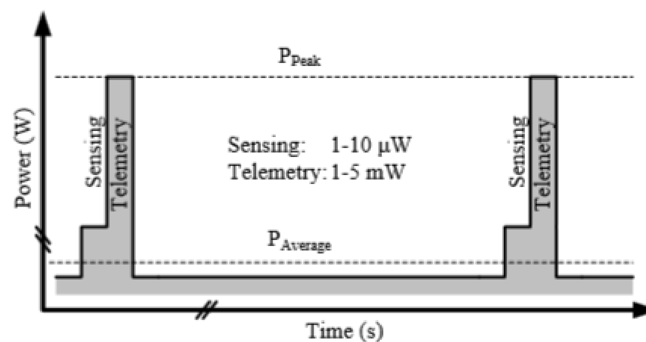


Figure 1.2: Load profile of a duty-cycled wireless micro-sensor TORRES; MILNER; RINCON-MORA (2008).

1.3 Fabrication Process Aspects

These strict supply limitations impose a lower limit on the dynamic range available for signal processing, and the ever increasing system complexity requires challenging specifications to be met. According to the International Technology Roadmap for Semiconductors (ITRS) 2013 report, one of the key challenges of integration is the manufacturing of quality on-chip passive devices, which introduces process complexity and can lead to manufacturing control and cost concerns. This makes resistorless circuits very desirable, but several issues with active device mismatch and average process variations have to be addressed.

The former, device mismatches due to local random variations in the dopant concentrations, for example, can be solved through good design and layout practices. The latter, average process variations that occur from batch to batch and affect entire wafers cannot be solved through design, heavily impacting the performance of every circuit, especially voltage references. That is why the traditional references use physical constants as their constituents, such as the thermal voltage and the silicon bandgap energy. Still, non-idealities degrade the performance of the reference, and for the case of average process variations, they can be addressed by two different but costly approaches: through complex schemes of compensation of opposing process dependences, or through calibration.

As was said before, the use of voltage references is ubiquitous in electronic systems. Examples of recent active research areas where they still require improvements are in

smart sensors, energy harvesters, implantable biomedical devices, digitally assisted analog / RF systems, identification through smart tags and *Internet of Things* devices.

1.4 Ideal Voltage Reference

Ideally a voltage reference is a circuit block that provides a constant voltage that is insensitive to variations in:

- temperature;
- power supply voltage;
- load current;
- packaging stresses;
- fabrication variability.

All of these items are important when specifying a voltage reference. The load current may be neglected when the reference is used to bias a MOSFET gate, for example, but the other parameters usually are universal concerns. Not all circuit topologies address the above mentioned issues, but they all use a common strategy to implement a temperature independent voltage reference. The basic concept is clearly illustrated by the classical bandgap reference (BGR) WIDLAR (1971), that is going to be the main focus of this thesis and shown in Figure 1.3. Other references based on Zener diodes or the Zero Temperature Coefficient (ZTC) point of MOSFETs are similar to the BGR, in the sense that they explore opposite temperature dependent phenomena to compensate each other, but are not the focus of this work.

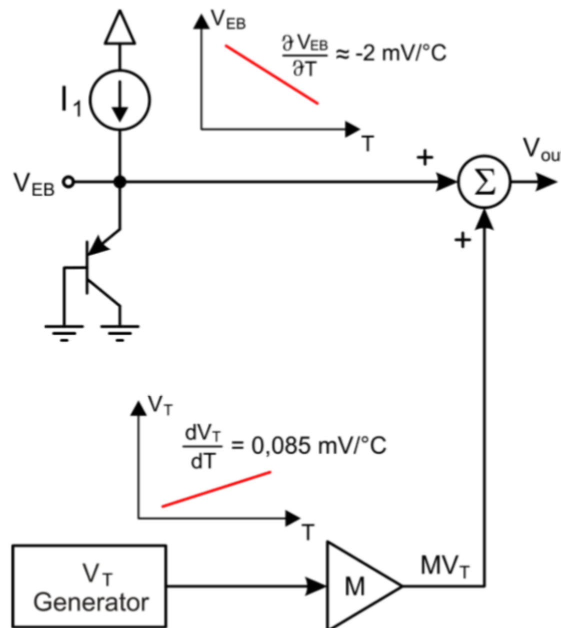


Figure 1.3: Basic concept of a bandgap voltage reference.

In the BGR two different and independent temperature sensitive variables are combined together in a weighted sum that neutralizes the temperature coefficient of the reference in a given temperature range. A linear proportional to absolute temperature (PTAT) term, based on the thermal voltage is amplified by M and counter-balanced by a BJT junction voltage, that is complementary to absolute temperature (CTAT) and slightly non-

linear. It is called a 'bandgap' voltage reference because the reference value V_{OUT} is approximately equal to the silicon bandgap voltage at 0 K. The circuit used to generate the biasing currents and thermal potential are going to be detailed later on.

This ideal behavior is of course degraded by non-idealities in the composing terms. For example, Figure 1.4 shows the non-linear behavior of the junction voltage, that leads to a well known 'curvature' on the final reference voltage. There are schemes that can compensate for this curvature, and they will be detailed on chapter 3.

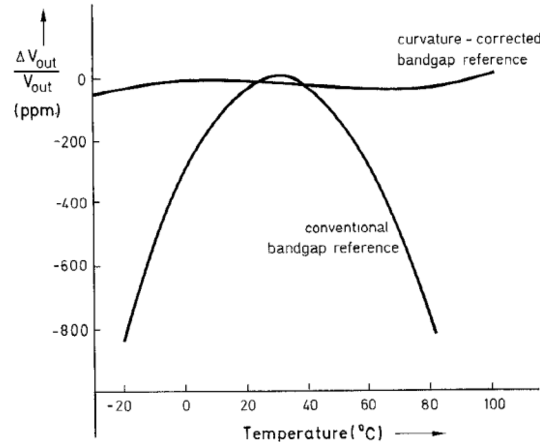


Figure 1.4: Output voltage deviation of the reference voltage plotted versus temperature for a conventional circuit and a curvature-corrected circuit with the linearized thermal behavior of $V_{BE}(T)$ MEIJER; SCHMALE; VAN ZALINGE (1982).

Junction curvature is not the only source of error in the ideal circuit of Figure 1.3, and others such as amplifier offset, device mismatch, bipolar transistor base resistance and finite current gain have been treated extensively in the literature for bandgap references PEASE (1990) and for temperature sensors as well AITA et al. (2013). There are also many ways to improve the DC and AC accuracy of this topology, where some of the solutions (and their respective drawbacks) can be listed as GUPTA; RINCóN-MORA (2010):

- trimming (cost, test time);
- Dynamic Element Matching (complexity, switching noise);
- series buffer (temperature sensitive offset);
- cascode current sources (diminished voltage headrooms).

1.5 Objectives

We recognize the importance of the techniques presented in the previous section, and understand that high performance circuits not restricted by cost or power consumption make use of them to achieve better results. However, the main purpose of this M.Sc. thesis is to address the implementation of novel topologies that achieve ultra-low-power operation with reduced supply voltages and that can be fabricated in standard digital CMOS processes without resistors. We explore proofs of concepts that have different characteristics than the current approaches, assessing their potential to achieve higher performances with the use of such complementary techniques.

1.6 Organization

The thesis is organized as follows: in chapter 2 a summarized bibliographic review of voltage references is presented in chronological order, and a results-oriented synthesis is composed for the most relevant recent advances. In chapter 3 the bipolar junction transistor is explored and its main aspects and limitations are presented. Also a conceptual curvature compensation scheme is derived. In chapter 4 we introduce a novel circuit that implements this concept, as well as a sub-bandgap reference derived from this topology. Thorough simulation results for two different processes are analyzed on chapter 5 together with experimental results of a fabricated chip. The thesis concludes with a comparison of the obtained results against the state of the art, and points out to future directions for the further development of these topologies.

2 INTEGRATED VOLTAGE REFERENCES OVERVIEW

We start this chapter by making a brief chronological review of the first circuits to implement bandgap references in a qualitative way. Then, before we present the most recent developments in resistorless CMOS voltage references, which are the theme of this thesis, we introduce the main performance metrics used to design and compare the different circuits. The last section evaluates recent research activity in low-power voltage references, specifically in resistorless approaches that are adequate to standard CMOS technology. The focus is on design topologies and experimental results that can provide the necessary figures of merit for the development, comparison and/or improvement of the thesis' proposed voltage reference designs. Being so, only recent publications will be considered, spanning at maximum 5 years of research (from 2009 onward), except where earlier publications might provide significant contributions.

2.1 Classic References

The concept of the bandgap voltage reference was introduced by Robert Widlar in 1971, in a now classic paper on the Journal of Solid-State Circuits WIDLAR (1971). He also developed more on the regulation of this voltage reference to supply a variable load, but we will focus here on the circuit of the reference voltage only. Shown in Figure 2.1, this reference, implemented in bipolar technology, was called "low voltage" because previous circuits employed zener diodes, that have breakdown voltages of about 5 to 6 V. The circuit proposed by Widlar, instead, used the negative temperature coefficient of the base-emitter junction of a bipolar transistor, added to a positive temperature coefficient voltage derived from the difference of two base-emitter voltages of transistors operating at different current densities. Since these voltages present opposite temperature dependencies, they can be added with the right gain setting to achieve a zero temperature coefficient reference, that is equal to the silicon bandgap voltage, approximately 1.205 V.

In this circuit, Q1 is operated at a relatively high current density. The current density of Q2 is about 10 times lower and the emitter-base voltage difference ΔV_{BE} between the two devices appears across R3, defining the emitter current - and the collector current as well if the transistor gain is high. Q3 then works as a current gain stage, also providing the V_{BE} voltage necessary to form the output. A simplified analysis of the circuit follows.

The collector current I_C of a BJT is given by WIDLAR (1971)

$$I_C = I_0 \left[\exp \left(\frac{V_{BE}}{m\phi_t} - 1 \right) \right] \tag{2.1}$$

where I_0 is a process dependent constant, V_{BE} is the base-emitter voltage, m is the emission coefficient and ϕ_t is the thermal voltage k_bT/q, being k_b the Boltzmann's constant,

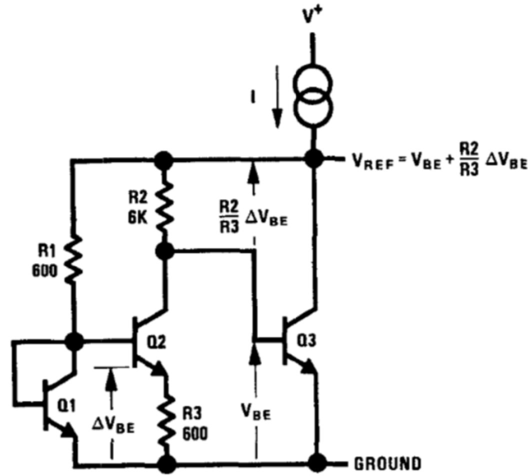


Figure 2.1: Classic Widlar Bandgap Voltage Reference WIDLAR (1971).

T the absolute temperature and q the electron charge. If two transistors are operated with different collector current densities J_i, then the difference of their base-emitter voltages is given by (2.2).

$$\Delta V_{BE} = \phi_t \ln \left(\frac{J_1}{J_2} \right) \tag{2.2}$$

In Figure 2.1 one can see that the output of the BGR proposed by Widlar is then given by the sum of the base-emitter voltage of Q3 and the multiplied PTAT term on R2, as shown by (2.3).

$$V_{REF} = V_{BE} + \frac{R_2}{R_3} \Delta V_{BE} \tag{2.3}$$

For now, we will assume that the base-emitter voltage can be expressed by the linear expression $V_{BE} = V_{G0} - CT$, where V_{G0} is the silicon bandgap voltage at 0 K, C is a constant and T is the absolute temperature. Other higher order effects will be treated extensively in chapter 3. Therefore, R1 and R2 can be scaled together with the current density difference of Q1 and Q2 to achieve $\partial V_{REF} / \partial T = 0$.

In 1973 Kuijk proposed a variation of the bandgap reference using an operational amplifier and BJTs connected as diodes, as shown in Figure 2.2 KUIJK (1973). Here ΔV_{BE} appears across R3 and defines the emitter current, which is PTAT, while R2 provides multiplication of this factor to form the temperature independent output V_O, after being added to D1 junction voltage, which is CTAT.

In 1974 Paul Brokaw introduced what is perhaps the more well known BGR circuit, shown in Figure 2.3 BROKAW (1974). In this circuit, the operational amplifier makes $I_{C1} = I_{C2}$ through resistances R, and the ΔV_{BE} that appears across R2 is then multiplied by 2R1, since both emitter currents of Q1 and Q2 flow into this resistor. The reference output is the sum of $V_{BE,Q1}$ with V_{R1}. He also introduces some compensation schemes for the base currents of Q1-Q2 through the use of auxiliary base resistances, not shown in Figure 2.3.

These architectures provided the first concepts of bandgap reference circuits, and already some non-idealities such as mismatch between devices, base current errors due to finite current gain β and bandgap curvature issues started to show up.

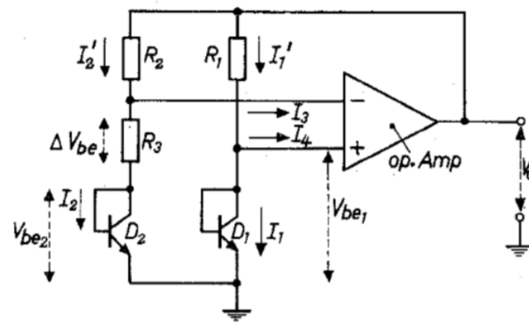


Figure 2.2: Kuijk Bandgap Voltage Reference KUIJK (1973).

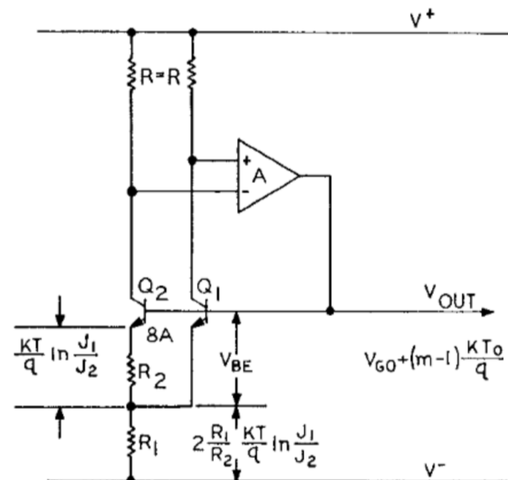


Figure 2.3: Brokaw Bandgap Voltage Reference BROKAW (1974).

Soon later on, Eric Vittoz pioneered the use of CMOS analog ICs operating in weak inversion VITTOZ; FELLRATH (1977). He presents several circuits, including a current and a voltage reference, and already notes that the mismatch of differential pairs are much higher for MOSFETs in weak inversion operation than for BJTs, but that the subthreshold slope n varies little. We will resume these findings later on the thesis.

In 1978 Tsividis presented a CMOS voltage reference that uses an NPN transistor in an N-substrate process, and introduces the generation of a PTAT voltage by a MOSFET unbalanced differential pair operating in weak inversion TSIVIDIS; ULMER (1978). He also cascades several PTAT generators (M1-M2) to achieve the approximate 0.6 V PTAT voltage necessary to counterbalance the CTAT junction voltage (Q), as shown in Figure 2.4.

Vittoz then introduces the self-cascode PTAT generator in 1979 VITTOZ; NEYROUD (1979), as well as a 1.3 V supply and sub-1μA voltage reference. This circuit presented a stability of 3 mV over 100 °C without trimming, which is an expressive result even for today circuits. The concept of the topology used is shown in Figure 2.5, while the PTAT generator is made of a cascade of self-cascode cells - Figure 2.6. The working principle of these cells will be detailed later on, when we present the proposed circuits of this thesis.

In 1982 Meijer proposed one of the first curvature compensation schemes through a heavily temperature dependent collector current MEIJER; SCHMALE; VAN ZALINGE (1982). Referring to Figure 2.7, that was implemented by balancing four series-connected

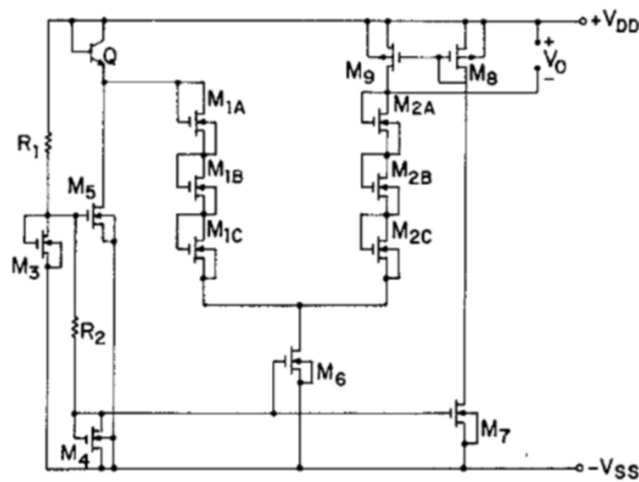


Figure 2.4: Tsividis CMOS Bandgap Voltage Reference TSIVIDIS; ULMER (1978).

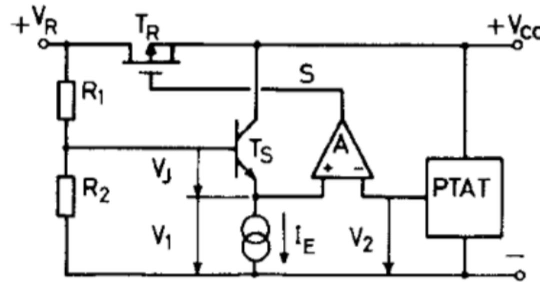


Figure 2.5: Vittoz CMOS Bandgap Voltage Reference Block Diagram VITTOZ; NEYROUD (1979).

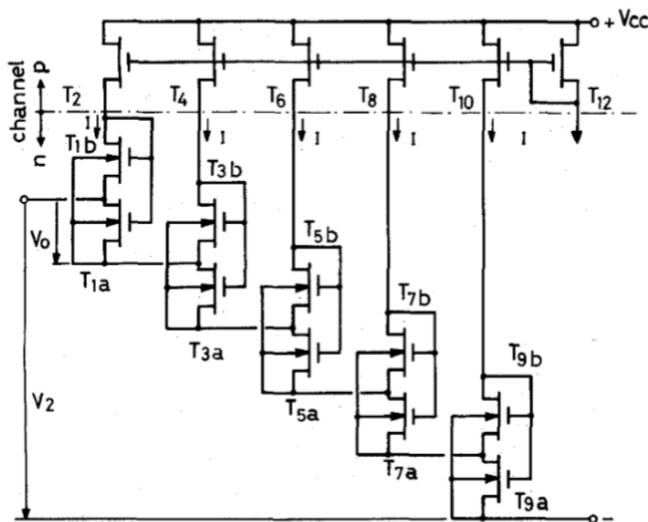


Figure 2.6: Stack of elementary PTAT self-cascode cells VITTOZ; NEYROUD (1979).

base-emitter junctions Q1-4 with a PTAT current against three series-connected base-emitter junctions Q12-14, that have a constant current. The PTAT current that biases Q1-4 makes their base-emitter voltage approximately 25% more linear against temperature, as

shown in the paper and later on in this thesis. Subtracting the three base-emitter voltages with higher non-linearity from the four with 25 percent lower non-linearity yields a voltage V'_{BE} which changes linearly with temperature. This linear portion of V'_{BE} is canceled by R1, which makes a voltage reference available at the emitter of Q14. The temperature independent current is automatically obtained when making R2 temperature independent. The output is thus regulated by the error amplifier and the bypass transistor Q23.

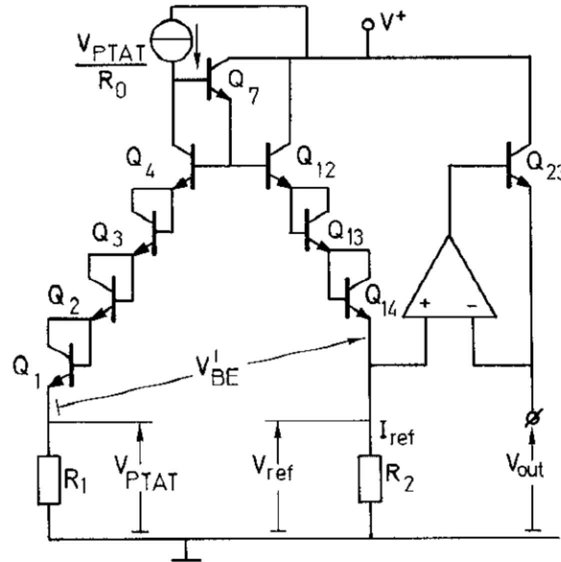


Figure 2.7: A curvature corrected bandgap reference with the linearized thermal behavior of $V_{BE}(T)$ MEIJER; SCHMALE; VAN ZALINGE (1982).

Another possible curvature correction technique is to add a PTAT^2 term to the linear PTAT term to compensate for the junction curvature. This was done in 1983 by SONG; GRAY (1983), and the concept is clearly demonstrated on Figure 2.8.

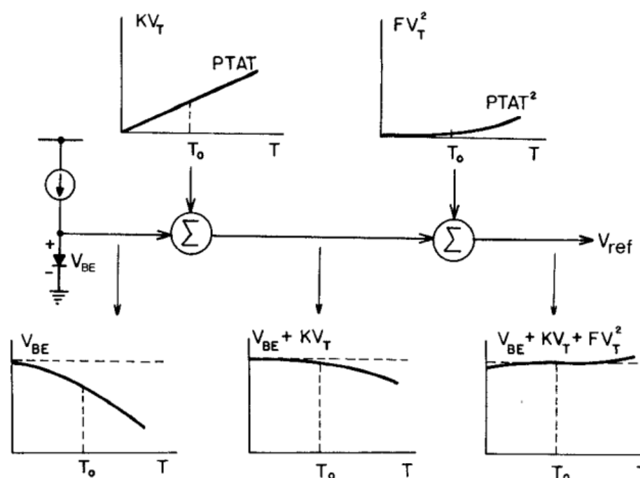


Figure 2.8: Curvature-compensation concept (not scaled) SONG; GRAY (1983).

These are some of the classic circuits that implement temperature independent voltages. The literature is very broad, but the topologies can be generally categorized as DE VITA; IANNACCONE (2007).

- Bandgap voltage reference (parasitic BJTs):
  - resistive TSIVIDIS; ULMER (1978);
  - resistorless BUCK et al. (2002);
  - switched-capacitors KLIMACH et al. (2013).
- Sub-bandgap voltage references:
  - resistive BANBA et al. (1999), LEUNG; MOK (2002);
  - resistorless OSAKI et al. (2013), MATTIA; KLIMACH; BAMPI (2014a), MATTIA; KLIMACH; BAMPI (2014b).
- Single threshold voltage devices:
  - pMOS and nMOS V_{TH} difference LEUNG; MOK (2003);
  - regular nFET based MATTIA; KLIMACH; BAMPI (2014c);
  - DTMOST based ANNEMA (1999);
- Different threshold voltages in the same technology DE VITA; IANNACCONE (2007), SEOK et al. (2012).

2.2 Definitions and Metrics

There are several different performance metrics used to specify and compare voltage references. These are listed below as:

1. Effective Temperature Coefficient (TC_{EFF});
2. Minimum Supply Voltage ($V_{DD,min}$);
3. Power Consumption;
4. Noise;
5. Line Sensitivity (LS);
6. Power Supply Rejection Ratio (PSRR);
7. Silicon Area.

Some of these parameters like area, power consumption and minimum supply voltage are pretty straightforward, while others deserve a better explanation. Of course, all of them are susceptible to the characteristics and variations of the fabrication process, which makes a comparison fair only when looking at the same or similar processes.

Historically, the temperature coefficient of a voltage reference has been presented in many ways. For example, it can be defined as the derivative of the reference voltage versus temperature at the reference temperature, which is essentially zero for the classic BGR. It is more common to use the effective temperature coefficient TC_{EFF} instead, which is given by (2.4). It measures the maximum variation of the voltage reference against a determined temperature range ΔT, and normalizes this value against the nominal V_{REF} at room temperature (also called box method). Its unit is in ppm/°C.

$$TC_{EFF} = \frac{V_{REF,max} - V_{REF,min}}{V_{REF,nom}\Delta T} \tag{2.4}$$

Another important metric is the ability of the voltage reference to reject supply variations. This measurement at DC is called line sensitivity (LS), given by (2.5) and measured in mV/V. Its frequency behavior is given by (2.6), called Power Supply Rejection Ratio

or PSRR, and measured in dB.

$$LS = \frac{V_{REF,max} - V_{REF,min}}{\Delta V_{DD}} \tag{2.5}$$

$$PSRR(jw) = \frac{v_{REF}(jw)}{v_{DD}(jw)} \tag{2.6}$$

As with every electronic circuit, the intrinsic device noise is the ultimate limiting accuracy factor. For recent voltage references, it's power spectral density is usually specified at 100 Hz for comparison purposes, while the total RMS noise depends on the bandwidth chosen.

These are the main design metrics used to specify and compare voltage references. We will use them extensively in the review of recent circuits, as well as on the evaluation of our proposed designs.

2.3 Recent Developments on Resistorless CMOS Voltage References

2.3.1 nMOS Diode in Strong Inversion

In DE VITA; IANNACCONE (2007) the authors' proposed voltage reference scheme consists of a resistorless nanoampere current source used to bias an nMOS diode in strong inversion. This current is proportional to $\mu(T)T^2$, obtained by using different threshold voltage transistors operating in weak (WI) and strong inversion (SI). The paper starts with a review of previous operating principles, that we decided to summarize here as well.

2.3.1.1 Previous Operating Principles

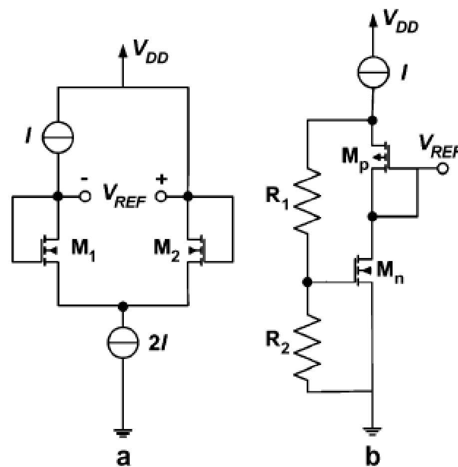


Figure 2.9: (a) Simplified circuit of a conventional voltage reference based on the difference between the gate-source voltages of two MOS transistors. (b) Simplified circuit of the voltage reference proposed in LEUNG; MOK (2003).

Referring to Figure 2.9 (a), considering both transistors in SI, and neglecting channel length modulation and body effect, the MOSFET drain current is given by:

$$I_D = \frac{1}{2}k(V_G - V_{TH})^2 \tag{2.7}$$

where V_G is the gate-source voltage and V_{TH} is the MOSFET threshold voltage, $k_i = \mu C_{OX} W_i / L_i$ being μ the mobility, C_{OX} the oxide capacitance, W the device width and L the device length. While the reference voltage of Figure 2.9(a) is given by:

$$V_{ref} = V_{GS2} - V_{GS1} = V_{th2} - V_{th1} + \sqrt{2I}\left(\frac{1}{\sqrt{k_2}} - \frac{1}{\sqrt{k_1}}\right) \tag{2.8}$$

to reduce the temperature effects, two threshold voltages with the same temperature coefficients are needed. A very low current ensures that $V_{ref} = V_{th2} - V_{th1}$. This approach has obvious issues with process control on the mobility, the gate oxide capacitance, the temperature coefficients of V_{th} and V_{th} itself.

An improvement on this circuit is shown in Figure 2.9 (b), which is based on the cancellation of the threshold temperature dependence of an nMOS and a pMOS device. Again considering saturation, SI, no short-channel and body bias effects:

$$V_{ref} = \left(1 + \frac{R_1}{R_2}\right) V_{GS_n} - |V_{GS_p}| \tag{2.9}$$

A complete derivation is presented in LEUNG; MOK (2003). Basically, this voltage reference is composed of a linear term that depends on the difference between the threshold voltages temperature dependences, and a non-linear term related to the mobility of both devices. The resistance ratio is chosen to minimize the threshold voltage TC, while W and L minimize the mobility temperature dependence *only*. Both effects are compensated only at the reference temperature, resulting in a non-linear temperature-dependent error voltage that appears at the reference output voltage.

A simplified circuit of a voltage reference, based on a difference between the gate-source voltages of two MOS transistors, that allows a perfect suppression of the temperature dependence of the mobility is shown in Figure 2.10 (a) DE VITA; IANNACCONE (2005).

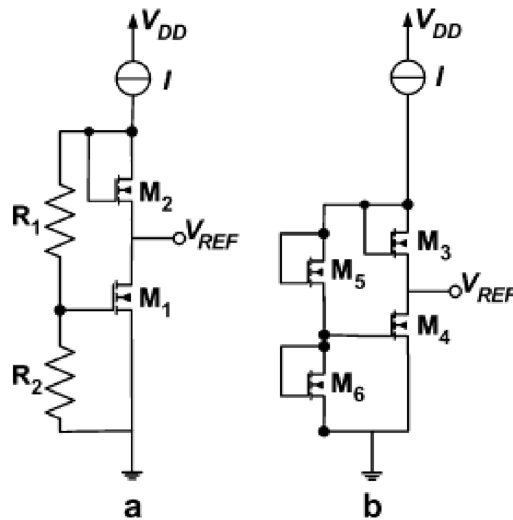


Figure 2.10: (a) Simplified circuit of the voltage reference proposed in DE VITA; IANNACCONE (2005). (b) Simplified circuit of the voltage reference proposed in DE VITA; IANNACCONE; ANDREANI (2006).

Referring to 2.10 (a), the reference voltage is given by:

$$V_{ref} = \left(1 + \frac{R_1}{R_2}\right) V_{GS1} - V_{GS2} \quad (2.10)$$

$$V_{ref} = \frac{R_1}{R_2} V_{th} + \sqrt{2I}\left[\left(1 + \frac{R_1}{R_2}\right)\frac{1}{\sqrt{k_2}} - \frac{1}{\sqrt{k_1}}\right] \quad (2.11)$$

If $I \; \alpha \; \mu(T)T^2$, the temperature dependence of $\mu(T)$ and $V_{th}(T)$ can be cancelled. Non idealities such as channel length modulation and body effect will degrade this performance. The circuit of figure 2.11 (b) solves these issues. Again considering SI, saturation and no short-channel effects:

$$V_{ref} = V_{th} + \left[\frac{1}{\sqrt{k_4}}\left(1 + \sqrt{\frac{S_6}{S_5}}\right) - \frac{1}{\sqrt{k_3}}\right]\sqrt{2I} \quad (2.12)$$

Again, if $I \; \alpha \; \mu(T)T^2$, one can obtain S_6/S_5 that sets $\partial V_{ref}/\partial T = 0$. Now, considering channel modulation effect, if the previous condition is satisfied, then:

$$\frac{\partial V_{ref}}{\partial T} = \frac{\lambda}{4} K_{t1}\left(2V_{th} - V_{ref}\right) \quad (2.13)$$

where K_{t1} is the first order temperature coefficient of V_{th} and λ is the short-channel effect coefficient. Setting $V_{ref} = 2V_{th}$ solves the channel length modulation issue, while keeping the sources of M3 and M5 at the same potential solves the body effect temperature dependence. Again, process spread issues with V_{th}, λ and maybe μ and K_{t1} will increase the circuit's variability, even though the transistors are in SI.

2.3.1.2 Proposed Operating Principle

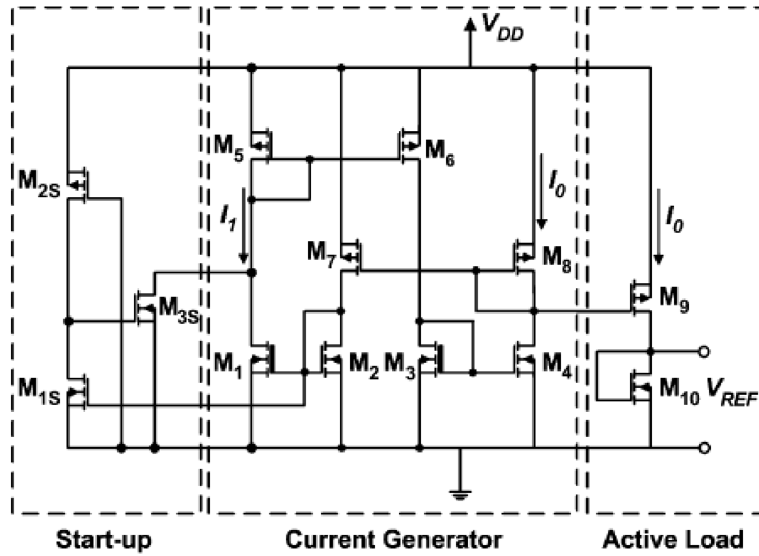


Figure 2.11: Proposed voltage reference circuit DE VITA; IANNACCONE (2007).

Referring to Figure 2.11, M1 and M3 are thick-oxide devices working in WI and saturation. M2, M4 and M10 are standard devices, working in SI and saturation. Assuming

Table 2.1: Reported results summary in DE VITA; IANNACCONE (2007).

Technology	0.35 μm CMOS
Supply Voltage (V)	0.9 to 4
Supply Current (μA)	0.04 @ 0.9 V, 0.055 @ 4 V
V_{REF}	670 mV
TC (ppm/°C)	10
Line Sensitivity	0.27 %/V
PSRR @ 100 Hz	-47 dB, V_{DD} = 0.9 V
Die area (mm^2)	0.045

no channel length modulation effect and that $V_{GS1} = V_{GS2}$, $V_{GS3} = V_{GS4}$, $V_{th1} = V_{th3}$ and $V_{th2} = V_{th4}$, the output current I_o is given by:

$$I_o = \frac{\mu C_{OX} S_4}{2(N-1)^2} n^2 \phi_t^2 \ln\left(\frac{S_3}{S_4}\right) \tag{2.14}$$

Where $N = \sqrt{S_4/S_2}$. M10 is the active load, and if it operates in strong inversion and saturation, the classic quadratic equation shows that:

$$V_{ref} = V_{th10} + \sqrt{\frac{2I_o}{k_{10}}} \tag{2.15}$$

Substituting the output current (2.14) into (2.15) produces the temperature compensated output voltage.

$$V_{ref} = V_{th10} + \frac{n\phi_t}{N-1}\sqrt{\frac{S_4}{S_{10}}} \ln\left(\frac{S_3}{S_1}\right) \tag{2.16}$$

This is a better relationship than the previous ones, since the authors have canceled the dependences on the gate oxide capacitance, the mobility and the threshold voltage temperature coefficients. Still, it shows a linear dependence on the threshold voltage of the output device. The authors point out that "*by neglecting matching errors on W/L ratios in (2.16), the sensitivity of the reference voltage is mainly due to the accuracy of the threshold voltage of the diode-connected nMOS transistor M10*". [...] The effect of process variations on the threshold voltage are not compensated in the present circuit, while those due to mismatch are dominated by random dopant fluctuations and lead to a standard deviation of the threshold voltage inversely proportional to $\sqrt{WL}$.

2.3.1.3 Measured Results

They measured 20 different samples from the same batch, and the results summary can be seen in Table 2.1.

2.3.2 Self-Cascode Cells and Diode-Connected nMOS

The proposed reference of UENO et al. (2009) is composed of a β multiplier bias circuit with an nMOS in triode region, followed by two self-cascode PTAT generators and diode-connected nMOS as the CTAT voltage component. The output voltage is fed

back into the gate of the nMOS resistor, making it less sensitive to process variations, as will be shown. The PTAT and CTAT terms are scaled to provide an output voltage of V_{TH0}, or the threshold voltage at 0 K. Since it has a fairly process independent TC, it can be used as a threshold voltage monitor as well. The proposed scheme is shown in figure 2.12.

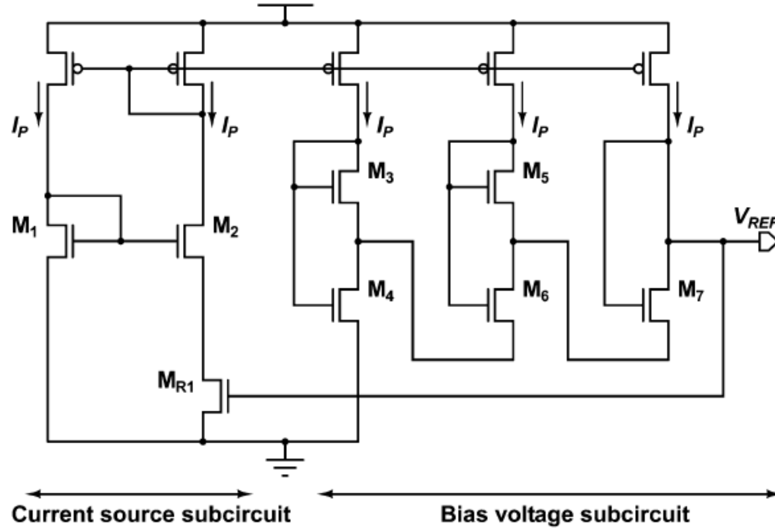


Figure 2.12: Proposed voltage reference UENO et al. (2009).

All transistors are in WI, except for M_{R1}, that is in strong inversion and deep triode region. For the bias circuit, the derivation is as follows:

$$V_{GS1} = V_{GS2} + V_{DS(MR1)} \tag{2.17}$$

Where $V_{DS(MR1)} = n\phi_t \ln(S_2/S_1)$. The bias current I_P is then given by:

$$I_P = \frac{V_{DS(R1)}}{R_{MR1}} = S_{R1}\mu C_{OX}(V_{REF} - V_{TH})n\phi_t \ln(S_2/S_1) \tag{2.18}$$

On the other hand, V_{REF} is given by:

$$V_{REF} = V_{GS4} + n\phi_t \ln\left(\frac{2S_3S_5}{S_6S_7}\right) = V_{TH} + n\phi_t \ln\left(\frac{3I_P}{S_4I_0}\right) + n\phi_t \ln\left(\frac{2S_3S_5}{S_6S_7}\right) \tag{2.19}$$

Where $I_0 = \mu C_{OX}(n-1)\phi_t^2$. The authors then show that $\partial V_{REF}/\partial T = 0$ if $V_{REF} = V_{TH0}$, being V_{TH0} the threshold voltage at 0 K. Assuming a linear temperature behavior for the threshold voltage $V_{TH} = V_{TH0} - K_1T$ results in

$$I_P = S_{R1}\mu C_{OX}K_1Tn\phi_t \ln(S_2/S_1) \tag{2.20}$$

This current is supposed to be less sensitive to process variations because it depends on the first thermal coefficient of the threshold voltage K_1 instead of V_{TH} itself. Still depends on μC_{OX} though. Also, they show that I_P is PTAT.

The derivation presented by the authors assumed that the lower transistors M4 and M6 of the self-cascode cell are in saturation, which is not valid. Equation 2.21 presents the

correct output voltage, using the ACM MOSFET model for weak inversion and triode operation (see equation 4.5 and Appendix B for the complete model equations, and Chapter 4 equation 4.6 for the derivation of the self-cascode V_{DS} voltage).

$$V_{REF} = V_{GS7} + V_{PTAT} = V_{T0} + n\phi_t \ln\left(\frac{I_P}{2eI_{SQ}S_7}\right) + \phi_t \ln\left[\left(\frac{2S_5}{S_6}+1\right)\left(\frac{3S_3}{S_4}+1\right)\right] \tag{2.21}$$

The first term of (2.21) is due to V_{GS7}, which is the diode connected device, while the second term is due to the PTAT voltage of the self-cascode cells.

2.3.2.1 Process Variation

Even though the output voltage depends on the threshold voltage, its TC depends on K_1, that is proportional to $\ln(1/N_A)$, so it varies little with process. This is illustrated by a simulation with first-order models that show that in the typical dopant concentration for CMOS technologies V_{TH} varies $\pm$ 20%, while the TC would only vary $\pm$ 2%.

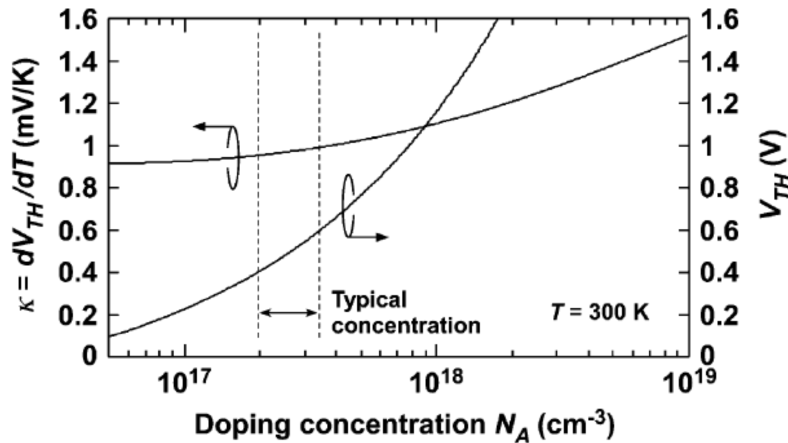


Figure 2.13: Reference voltage and its TC versus dopant concentration UENO et al. (2009).

The final circuit, shown in Figure 2.14, includes some extra circuitry to compensate for poor line sensitivity and noise of the β multiplier circuit, as well as frequency compensation of the inserted op-amp.

2.3.2.2 Simulation Results

They have done a within-die (gaussian distribution) and batch-to-batch (uniform distribution) MC simulation with 300 runs. On the -20 to 80 °C temperature range, $\mu(V_{REF}) = 840$ mV, with a standard deviation $\sigma = 60$ mV, which gives $\sigma/\mu = 7$ %.

2.3.2.3 Experimental Results

Table 2.2 presents a summary of the typical condition results.

They also measured 17 samples of the same die, reporting a total V_{REF} variation of 25 mV, a $\sigma/\mu(V_{REF}) = 0.87$ %, a mean TC = 15 ppm/°C, with minimum of 7 ppm/°C and maximum of 45 ppm/°C.

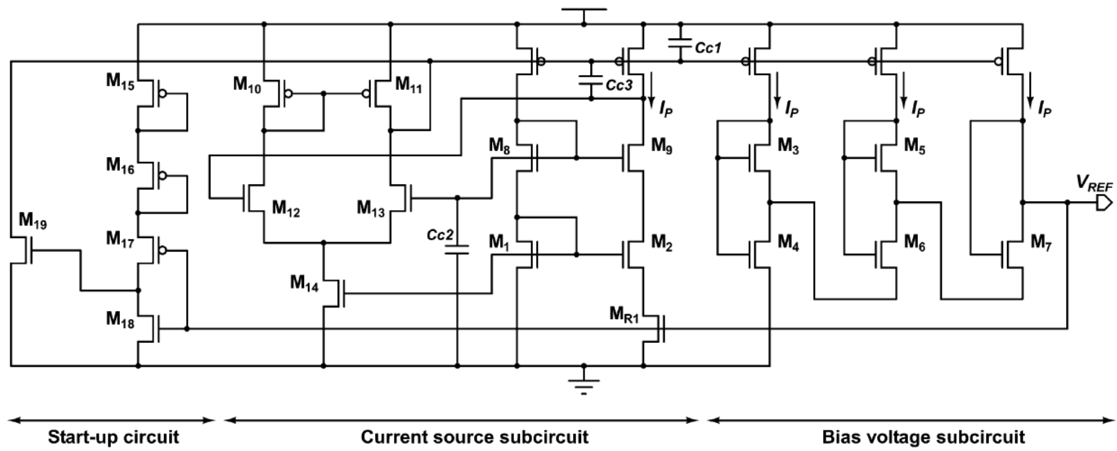


Figure 2.14: Entire circuit of the voltage reference. All MOSFETs are in WI, except for M_{R1} UENO et al. (2009).

Table 2.2: Typical condition experimental results summary UENO et al. (2009).

	This work
Process	0.35 μm CMOS
Temperature Range	- 20 to 80 °C
V_{DD}	1.4 to 3 V
V_{REF}	745 mV
Power	0.3 μW (@1.4V) Room Temp.
TC	7 ppm/°C
Line Sensitivity	20 ppm/V
PSRR @ 100 Hz	-45 dB
Die area	0.055 mm^2

2.3.3 Polynomial Current Curvature Compensation Scheme for BGR

In MING et al. (2010) a polynomial current source is used to compensate for non-linearities in the BJT junction. That is summed with a PTAT voltage provided by the difference between two BJT junctions, multiplied by a transconductance gain, a current gain and a transresistance gain. MOS transistors are not in weak inversion.

2.3.3.1 Operating Principle

Referring to Figure 2.15, *the difference between the forward-bias voltages across two diodes (Q_1 and Q_2) ΔV_D is applied across the differential pair $M_{16} - M_{17}$, which acts as a transconductance. The resulting current is multiplied using a current mirror $M_{13} - M_{20}$ and is delivered to differential pair $M_{18} - M_{19}$, which operates as a transresistance.*

With the transistor aspect ratios given in Figure 2.15, the output voltage is derived from two gate-source loops and is given by (2.22).

$$V_{REF} = V_{D3}(T) + \sqrt{AG}(V_{D1} - V_{D2}) \tag{2.22}$$

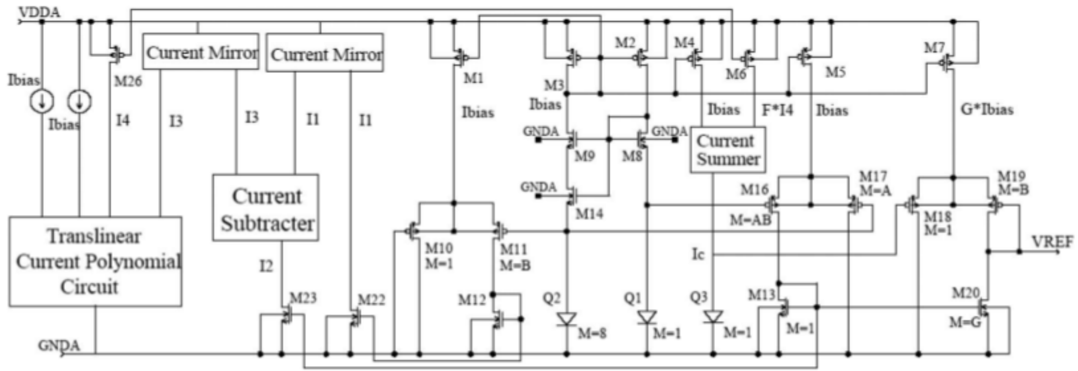


Figure 2.15: Bandgap reference circuit configuration MING et al. (2010).

Table 2.3: Performance summary MING et al. (2010).

	This work
Process	0.5 μm CMOS
Threshold Voltage	0.85 V
Supply Voltage	3.6 V
Power dissipation	0.648 mW
V_{REF}	1.23 V $\pm$ 1.2 mV
TC	11.8 ppm/°C
PSRR @ 100 Hz	-45 dB

The authors emphasize that *the important difference from* [the work of 2002-Buck] *is that an additional diode Q_3 is adopted for the base-emitter voltage instead of Q_1 or Q_2. The reason why this is done is that it can be used for curvature compensation more freely without affecting bias of the circuit.*

2.3.3.2 Measured Results

Five devices from the same wafer were fabricated in CMOS 0.5 μm technology, resulting in a silicon area of 500 μm x 200 μm. Process variability results are not reliable, but they are expected to be reasonably low, since the directly-process dependent parameters on the output voltage are related to bipolar junctions. The MOSFETs contributions are only through sizing ratios, that are fairly independent on process and, furthermore, can be trimmed digitally. The authors trim the reference value first, then trim the curvature compensation.

Table 2.3 summarizes these results for the five chips, that were digitally trimmed before the measurements. They achieve $\pm$ 3mV considering average absolute value and temperature variations.

2.3.4 nMOS Diode in Weak Inversion

The proposed voltage reference consists of a current source with polinomial temperature dependance, that cancels out the temperature dependance of the threshold voltage of

an NMOS diode acting as an active load. The author's derivation is as follows:

2.3.4.1 Operating Principle

The I-V characteristics of an nMOS operating in subthreshold region is given by

$$I_D = S\mu C_{OX}\phi_T^2 \exp\left(\frac{V_{GS} - V_{th}}{n\phi_T}\right)\left[1 - \exp\left(-\frac{V_{DS}}{\phi_T}\right)\right] \tag{2.23}$$

From (2.23), for $V_{DS} >= 4\phi_T$, I_D becomes almost independent of V_{DS}, thus V_{REF} can be approximated by

$$V_{REF} = V_{GS} = V_{th} + n\phi_t \ln\left(\frac{I_D}{SC_{OX}\mu\phi_T^2}\right) \tag{2.24}$$

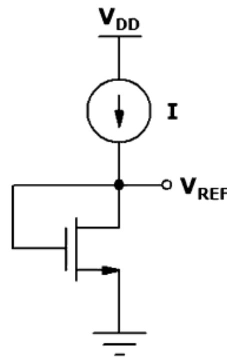


Figure 2.16: Scheme of principle of a simple voltage reference MAGNELLI et al. (2011).

The authors also consider that V_{th} decreases linearly with temperature, according to:

$$V_{th} = V_{th}(T_0) + (k_{t1} + k_{t2}V_{BS})\left(\frac{T}{T_0} - 1\right) \tag{2.25}$$

where coefficients k_{t1} and k_{t2} have negative values. In order to have a temperature compensated output reference voltage, the following condition must be satisfied:

$$\frac{\partial V_{REF}}{\partial T} = 0 \tag{2.26}$$

Based on (2.24), a simple solution of (2.26) can be obtained by generating the current I_D with the following temperature dependence:

$$I_D(T) = \alpha\mu T^2 \tag{2.27}$$

Therefore, imposing (2.26) a value for α (constant with temperature) can be found. Substituting (**??**) in (2.24) and recalling that $V_{BS} = 0$ V in the scheme of Figure 2.16, the reference voltage becomes:

$$V_{REF} = V_{th}(T_0) + |k_{t1}| \tag{2.28}$$

However, this solution does not ensure subthreshold operation for the active load. Another solution, which can ensure subthreshold operation, can be written as

$$I_D = \alpha\mu T^2 \exp\left(\frac{AT + B}{CT}\right) \tag{2.29}$$

where A, B and C are constant with temperature. Substituting (2.29) in (2.24), from (2.26) an expression for α can be found which leads to the temperature-compensated reference voltage:

$$V_{REFopt} = V_{th}(T_0) + |k_{t1}| + n\frac{k_b}{q}\frac{B}{C} \tag{2.30}$$

In this case, if the term B/C is negative, the load transistor can work in subthreshold region.

2.3.4.2 Electrical Configuration

In Figure 2.17, the proposed circuit is shown, where M1S and M2 are thick-oxide MOSFETs.

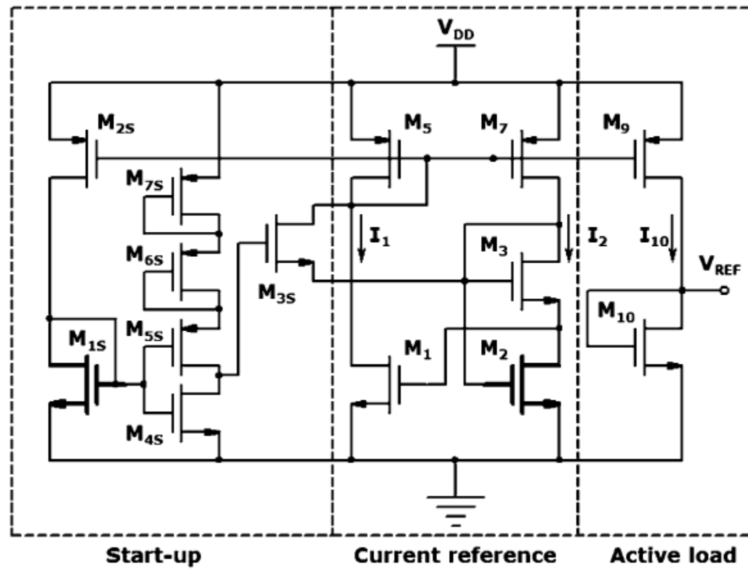


Figure 2.17: Proposed CMOS voltage reference circuit MAGNELLI et al. (2011).

Following the authors derivation, the load current is:

$$I_{10} = CI_1 = CQ^{1/\sum n}\mu\phi_T^2 \exp\left(-\frac{\partial V_{th}}{\phi_T \sum n}\right) \tag{2.31}$$

where $Q = \alpha^{n_2 - n_3}(S_3^{n_3} S_1^{n_1}/S_2^{n_2})$, $C = S_9/S_5$ and $\sum n = n_1 + n_3 - n_2$. Setting $\partial V_{REF}/\partial T = 0$, a value for S_{10} can be found and V_{REF} becomes

$$\begin{aligned} V_{REFopt} = V_{th10}(T_0) + |k_{t1,10}| + \frac{n_{10}}{\sum n}(V_{th2}(T_0) + |k_{t1,2}|) \\ - \left[\frac{n_{10}}{\sum n}(V_{th1}(T_0) + V_{th3}(T_0) + |k_{t1,1}| + |k_{t1,3}| + |k_{t2,3}|V_{BS3})\right] \end{aligned} \tag{2.32}$$

showing that the reference voltage value depends linearly on process parameters variations such as $\sigma_{V_{th}}$. Assuming no correlation between standard and thick-oxide MOSFETs, the output voltage variability can be predicted by (2.33).

$$\sigma_{VREF} = \sqrt{\sigma_{HVT}^2 + \sigma_{SVT}^2} \tag{2.33}$$

Table 2.4: Reported results under typical conditions MAGNELLI et al. (2011).

	Proposed Configuration
Process	0.18 μm CMOS
Supply Voltage	0.45 to 2 V
Supply Current	0.007 μA at 0.45 V and 27 °C
V_{REF}	263.5 mV
Temperature Range	0 to 125 °C
TC	142 ppm/°C
Line Sensitivity	0.44 %/V
PSRR @ 100 Hz	-45 dB, V_{DD} = 0.45 V
Die Area	0.043 mm^2

Table 2.5: Statistical Analysis of Performance MAGNELLI et al. (2011).

Number of samples	40	(three runs)
	μ	σ
TC (ppm/°C)	165	100
LS (%/V) @ 25 °C	0.444	0.058
Power (nW) @ 0.45 V & 25°C	2.6	0.7
V_{REF} (mV) @ 25 °C	257.5	10

Where σ_{HVT} and σ_{SVT} correspond to the variability of the threshold voltage of thick oxide and standard MOSFETs.

2.3.4.3 Measurement Results

The authors have measured the threshold voltage variation σ_{SVT} for 40 dies on three different runs, and reached the average and standard deviation values of 318.5 mV and 6.7 mV, respectively, resulting in $\frac{\sigma}{\mu}(V_{th}) = 2.1\%$. The typical reported values are shown in table 2.4.

The histograms of the measured results are presented in Figure 2.18. The variability results reported must be analyzed carefully, since they are more prone to evaluate mismatch effects than process variations (only three different wafers). Considering that the output voltage depends linearly on the threshold voltage variation, greater spreads would be expected if a higher number of wafers were to be measured. The authors mention that *Die to Die variations influence the absolute accuracy of transistor parameters and their effects are not compensated in the proposed configuration.*

Also, according to the authors, *To our knowledge, detailed statistical information about key performance indicators like TC, LS or power consumption, are not given in works concerning low power voltage references. For that reason it was not possible to compare the statistical results in detail.*

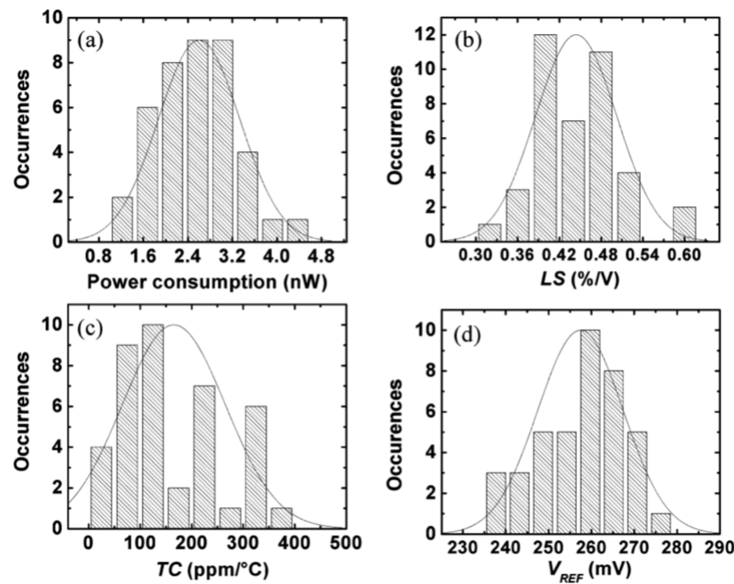


Figure 2.18: Distributions of the most relevant figures of merit for the 40 measured samples: (a) Power consumption @ 25 ∘C and V_{DD} = 0.45 V. (b) Line sensitivity @ 25 °C. (c) Temperature coefficient. (d) Generated voltage reference @ 25 °C MAGNELLI et al. (2011).

2.3.5 2-Transistor Voltage Reference

In SEOK et al. (2012) a thorough bibliographical research is presented, and an interesting comparison graph (Figure 2.19) demonstrates the performance improvement of the proposed circuit against prior art.

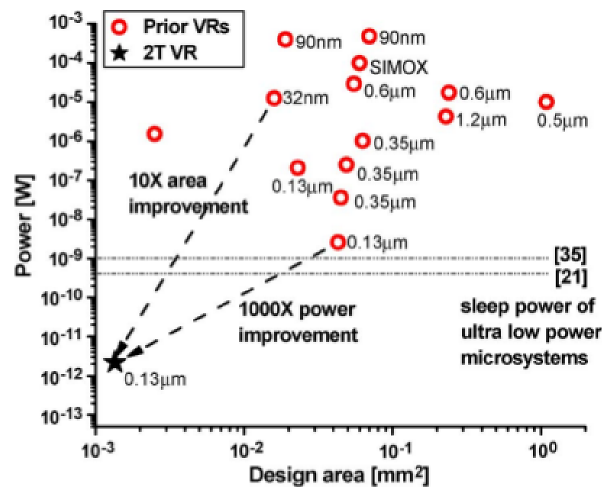
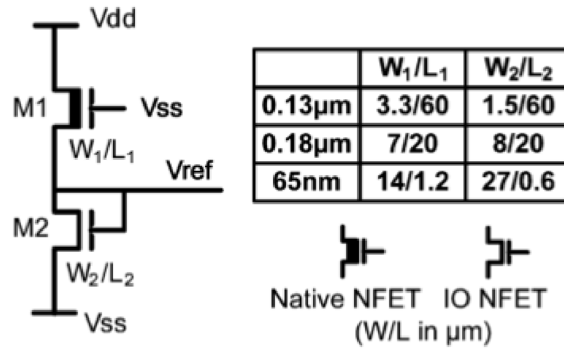


Figure 2.19: Power and area comparisons of recently published voltage references SEOK et al. (2012).

2.3.5.1 *Electrical Configuration*

The idea is to use a native MOSFET in WI (M1) to bias a thick-oxide I/O MOSFET (M2) with different threshold voltage, as in Figure 2.20.



	W1/L1	W2/L2
0.13µm	3.3/60	1.5/60
0.18µm	7/20	8/20
65nm	14/1.2	27/0.6

Figure 2.20: Schematic of proposed 2T voltage reference SEOK et al. (2012).

The analysis is as follows:

$$I_{sub} = \mu C_{OX} S(n-1)\phi_T^2 \exp\left(\frac{V_{GS} - V_{th}}{n\phi_T}\right)\left[1 - \exp\left(-\frac{V_{DS}}{\phi_T}\right)\right] \tag{2.34}$$

Given that both transistors are in saturation and their $V_{DS} > 100$ mV, and that $I_{D1} = I_{D2}$, V_{REF} becomes V_{GS2}.

$$V_{REF} = \frac{n_1 n_2}{n_1 + n_2}(V_{th2} - V_{th1}) + \frac{n_1 n_2}{n_1 + n_2}\phi_T ln\left(\frac{\mu_1 C_{OX1} W_1 L_2}{\mu_2 C_{OX2} W_2 L_1}\right) \tag{2.35}$$

where both the first and second terms can be either PTAT or CTAT. Setting the first derivative equal to zero allows the optimization of $\frac{W_1}{W_2}$, which is:

$$\frac{\partial V_{ref}}{\partial T} = 0 \rightarrow \left(\frac{W_1}{W_2}\right)_{opt} = \frac{\mu_1 C_{OX1} W_1 L_2}{\mu_2 C_{OX2} W_2 L_1} exp\left[\frac{q}{k}(C_{V_{th2}} - C_{V_{th1}})\right] \tag{2.36}$$

where $C_{V_{th1}}$ and $C_{V_{th2}}$ are the threshold voltage first order temperature coefficients. According to the authors, *the longest gate length ($L_1 = L_2 = 60\ \mu m$) allowed by the process design rules can be used for both devices to achieve ultra-low power consumption, while shorter gate lengths can reduce area and drive lower impedance nodes at the expense of power*. Also, longer gate lenghts increase the settling time of the circuit, which increases the cost of the calibration procedure described later on the paper. Process variability of the $C_{V_{th}}$s will strongly affect the circuit's TC, since they are exponentially related.

2.3.5.2 Measured Results

Figure 2.21 summarizes the typical performance of the proposed circuit and its variants, to be described later on. The authors have fabricated the 2T topology on three different technologies, proving the ease of portability of the design, that consists of sizing only two transistors.

For the 0.13μm CMOS process they have measured 49 dies across two different runs, and table 2.6 summarizes the variability results.

The authors mention, for simulation results, that *although transistor parameters vary largely across global variations, the way each parameter is used in V_{ref} and the correlation between two devices reduce their impact on output values. For example, $\sigma/\mu(V_{th2}) = 20\%$, but $\sigma/\mu(V_{th2} - V_{th1}) = 5\%$. Also, $\sigma/\mu\left(\frac{n_1 n_2}{n_1+n_2}\right) = 1.4\%$, and $\sigma/\mu\left(\frac{\mu_1}{\mu_2}\right) = 9\%$.*

TABLE II
MEASUREMENT SUMMARY OF THE PROPOSED 2T VOLTAGE REFERENCE AND ITS VARIANTS

	2T			Trimmable	4T
Process	0.13µm CMOS	0.18µm CMOS	65nm CMOS	0.13µm CMOS	0.13µm CMOS
V_{dd}	0.5-3.0V	0.5-3.6V	0.5-2.5V	0.5-3.0V	0.5-3.0V
V_{out} (min) **V_{out} (max)**	174.9mV 178.7mV	326.8mV 330.0mV	327.2mV 333.0mV	175.2mV 176.5mV	341.5mV 348.1mV
TC (min) **TC (max)**	16.9ppm/°C 231ppm/°C	54.1ppm/°C 176.4ppm/°C	89.1ppm/°C 118.2ppm/°C	5.3ppm/°C 47.4ppm/°C	80.2ppm/°C 142.5ppm/°C
LS	0.033%/V	0.044%/V	0.33%/V	0.036%/V	0.036%/V
PSRR	-53/-62dB (100Hz/10MHz)	-49/-55dB (100Hz/10MHz)	-40/79dB (100Hz/10MHz)	-51/-64dB (100Hz/10MHz)	-58/-59dB (100Hz/100kHz)
Power **(V_{dd}, temp)**	4.4pA (0.5V,25°C) 81pA(3V,80°C)	11pA (0.5V,25°C) 139pA(3V,80°C)	0.48nA (0.5V, 20°C) 8.13nA (2.5V, 80°C)	59pA (0.5V,25°C) 847pA(3V,80°C)	21.7pA (0.5V, 25oC) 400pA (3V, 80oC)
Size	1350µm^2	1425µm^2	900µm^2	9300µm^2	3500µm^2
Comment	2 runs, 49 dies	1 run, 14 dies	1 run, 17 dies	Post-trimming 1 run, 25 dies	1 run, 30 dies

Design size includes output capacitors.

Figure 2.21: Measurement Summary of the Proposed 2T Voltage Reference and its Variants.

Table 2.6: Variability Results (49 dies, two runs)

	σ	μ	σ/μ
TC (ppm/^{o}C)	41	62	66%
V_{ref1} (mV)	1.5	176.1	0.85%
V_{ref2} (mV)	1	176.7	0.57%

2.3.5.3 Calibration Procedure

The circuit can be calibrated across process variations by adjusting the aspect ratio of the native and thick-oxide transistors, according to the scheme of Figure 2.22.

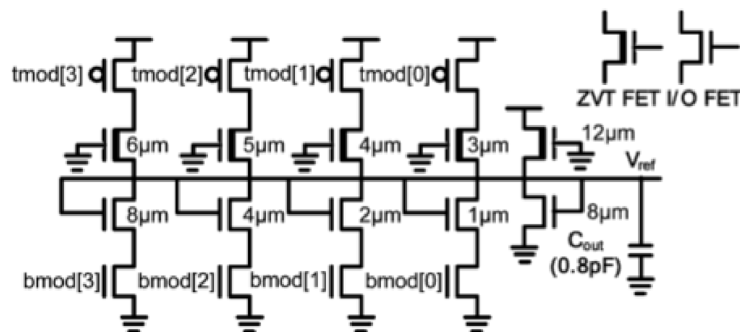


Figure 2.22: Schematic of trimmable 2T voltage reference (L = 60um is used) SEOK et al. (2012).

According to the authors, *By applying control signals bmod and tmod to the switches, the top-to-bottom width ratio varies from 0.52 to 3.75 with 256 different strings. Control signals swing full rail, requiring no extra supply voltage. One-Time Programmable (OTP) memories such as fuses can be used to provide the signals with minimal power overhead. [...] a 0.8 pF output capacitor suppresses the effect of noise on output voltage.*

The goal of the trimming process is to minimize output voltage spread, which can also reduce the TC. There is a correlation between these two parameters, like in the classic Brokaw BGR Optimum Cell Voltage BROKAW (1974) or Bob Pease's V_{magic} PEASE (1990), which is shown in Figure 2.23. According to the authors, *the correlation between the output voltage and the TC is confirmed through SPICE simulations and silicon measurements. We target a TC $<$ 50 ppm/°C and investigate a single temperature point (80°C) trimming for 25 dies (1 run) to minimize time associated with trimming procedures. Since we cannot measure the TC with only one temperature point, the trimming process relies entirely on the output voltage*. The detailed trimming process can be seen in the paper.

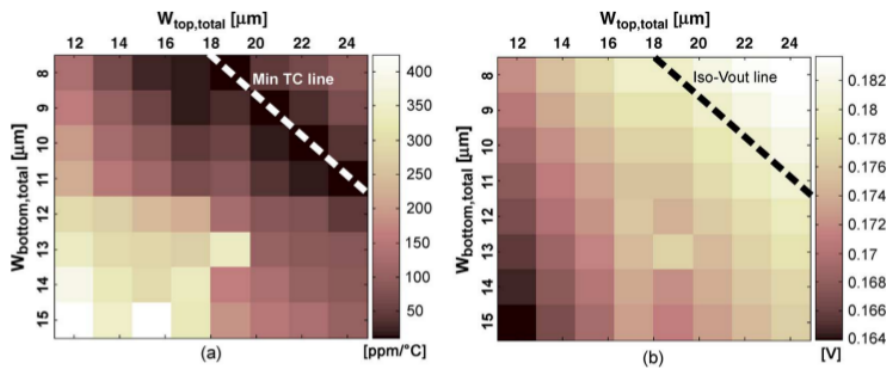


Figure 2.23: Measured (a) TC and (b) output voltage dependency on trim settings SEOK et al. (2012).

Also, Figure 2.24 shows the initial and post-trimming spread, proving the efficacy of the calibration procedure. It reduces the spread of TC and output voltage by 9.6x and 9.8x, respectively, compared to pre-trim results for the 25 dies (1 run). Final TC is below 50 ppm/°C for all 25 chips.

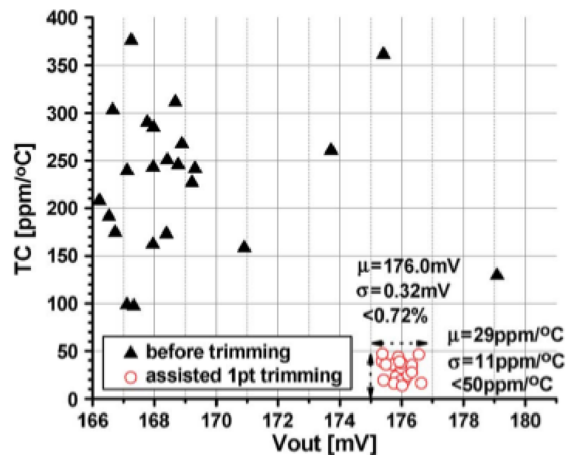


Figure 2.24: Measured reductions of output voltage and temperature coefficient spreads after assisted one temperature point trimming in 0.13um SEOK et al. (2012).

Finally, the authors propose another circuit for generating higher output voltages, that simply stacks the original approach, as is often done with the self-cascode topology. For completeness, the circuit is shown in Figure 2.25 and the results presented in Figure 2.21 as well.

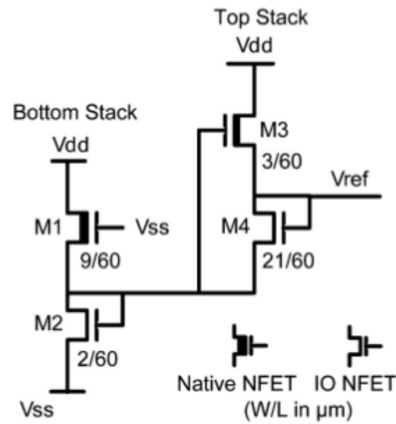


Figure 2.25: Schematic of 4T voltage reference SEOK et al. (2012).

Table 2.7: Performance summary of OSAKI et al. (2013).

	This Work	
CMOS Technology	0.18	μm
Type	BGR	Sub-BGR
Supply Voltage (V)	1.2 - 1.8	0.7 - 1.8
Active Area (mm^2)	0.0294	0.0246
Reference Voltage (V)	1.09	0.548
Temperature Range (°C)	-40 - 125	-40 - 125
TC (ppm/°C)	147	114
Power (μW)	0.1 @ Room Temp.	0.0525 @ Room Temp.
PSRR @ 100 Hz (dB)	-62	-56

2.3.6 Junction Divider and Unbalanced Differential Pair

The main contribution of the OSAKI et al. (2013) paper is the combination of a resistorless nanoampere current source to bias a bipolar junction below 600 mV, summed with cascaded MOSFET unbalanced differential pairs as a PTAT generator. A V_{EB} divider circuit is also proposed to lower the junction voltage and decrease the number of cascaded PTAT stages.

2.3.6.1 Electrical Configuration

Figures 2.26 and 2.27 show the BGR and sub-BGR circuits proposed, with start-up branch, current reference, BJT and PTAT generators.

2.3.6.2 Measured Results

The current reference circuit, the BGR and the sub-BGR circuits occupy an area of 0.0144 mm^2, 0.0150 mm^2 and 0.0102 mm^2, respectively. The authors measured nine sample dies from the same wafer. The average values of the figures of merit are reported in table 2.7.

The authors also present 500 Monte Carlo runs for V_{REF}, being $\sigma/\mu = 1.61$ %.

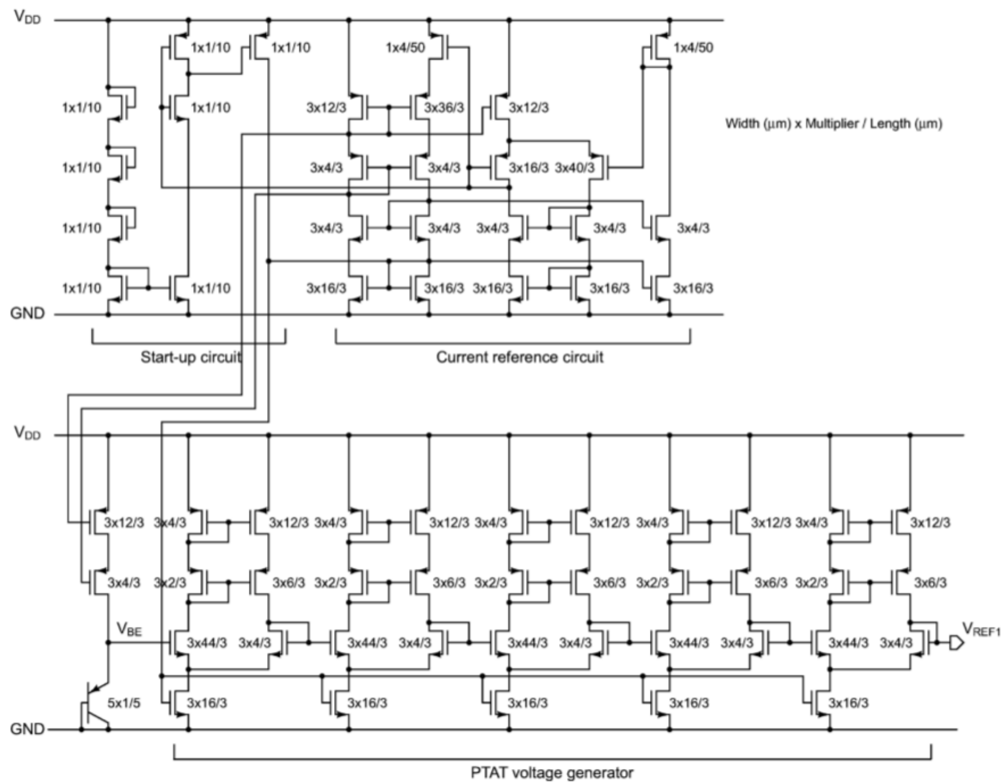


Figure 2.26: BGR Reference Circuit OSAKI et al. (2013).

2.3.7 Summary of Recent Advances

This list is by no means extensive, and several important circuits were left out because of the context of this thesis. In table 2.8 we present a summarized comparison of the main specifications of the reviewed references. All specified metrics are measured according to section 2.2.

From Table 2.8 one can conclude that some issues have already been solved, like low power consumption and power supply, with circuits operating at pico-Watts and 0.45 V supply. Most of the recent references are threshold based, and even though there is not a significant amount of data available, it is fair to say that they suffer from high variability due to average process variations. The simulated σ/μ can reach up to 7 % for the low TC of the Ueno reference, while tighter V_{REF} distributions usually come with expressive temperature variations, as is the case for the other references. The curvature-compensated reference of Ming doesn't operate with low supply voltages and is the most complex circuit, with an area of 0.1 mm^2, consumes five orders of magnitude more power and has to be individually calibrated. Even though the work of Seok presents a cheap one-point voltage calibration with good performance results and portability, we believe that similar results can be achieved without calibration by the use of bipolar transistors and novel topologies.

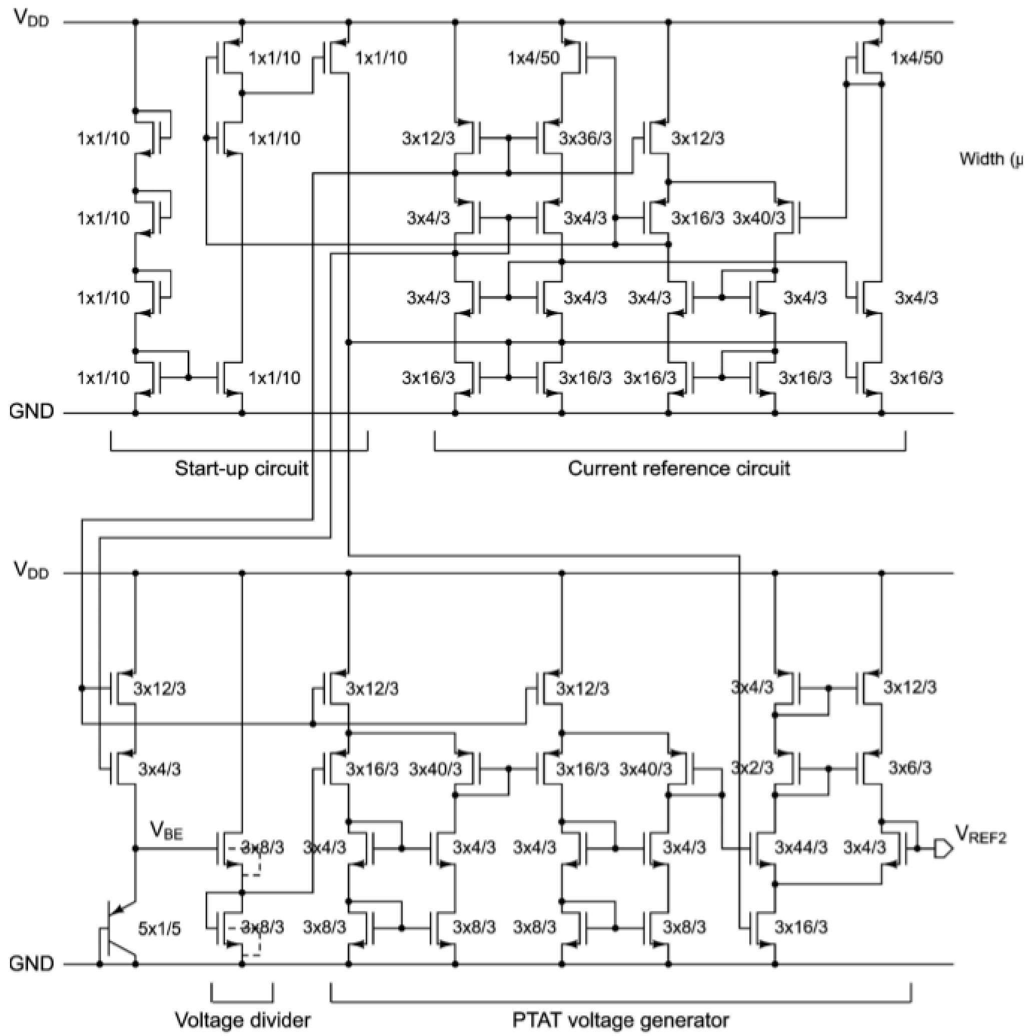


Figure 2.27: Sub-BGR Reference Circuit OSAKI et al. (2013).

Table 2.8: Summary of recent resistorless CMOS Voltage References.

[1] individually trimmed dies	De Vita 2007	Ueno 2009	Ming 2010	Magnelli 2011	Seok 2012	Osaki 2013	Unit
Technology	0.35	0.35	0.5	0.18	0.18	0.18	μm
CTAT Voltage	V_{T0}	V_{T0}	V_{EB}	V_{T0}	V_{T0}	V_{EB}	–
Temperature Range	0 – 80	-20 – 80	-40 – 120	0 – 125	-20 – 80	-40 – 120	°C
number of measured samples	20 same batch	17 same batch	5 same batch	40 three batches	49 (22 + 27) two batches	9 same batch	
V_{REF}							
min	607	733	–	240	175	524	mV
avg	670	745	1.23[1]	263.5	176	551	mV
max	732	759	–	280	181	577	mV
σ/μ (Monte Carlo)		7.15	–	–	–	1.61	%
Temperature Coefficient							
min	–	7	–	39	16.9	52	ppm/°C
avg	10	15	11.8[1]	142	62	114	ppm/°C
max	–	45	15.6[1]	400	231	148	ppm/°C
Noise @ 100 Hz	–	–	–	2	16	1.9	μV/$\sqrt{Hz}$
V_{DD}	0.9 – 4	1.4 – 3	3.6	0.45 – 2	0.5 – 3	0.7 – 1.8	V
Power							
min	–	280	–	1	–	–	nW
nom	36	300	$6.48 \cdot 10^5$	3.15	0.0022	52.5	nW
max	–	350	–	15	–	0.081	nW
Line Sensitivity	1.83	0.015	–	1.2	0.077	-	mV/V
PSRR @ 100 Hz	-47	-45	-31.8	-45	-49	-56	dB
Silicon Area	0.045	0.055	0.1	0.043	0.0014	0.0246	mm^2

3 BIPOLAR JUNCTION TRANSISTOR CHARACTERISTICS

We now make a detailed analysis of BJT biasing currents and present a curvature compensation technique based on the collector current temperature dependence, that is going to serve as a theoretical basis for the proposed circuits of chapter 4.

3.1 BJT Temperature Model

The analysis that follows is largely based on the work of TSIVIDIS (1980), where the main non-idealities of the I_C - V_{BE} curve were identified and an analytical model was proposed. We start by recognizing that the silicon bandgap voltage V_G has a temperature dependence

$$V_G(T) = V_G(0) - \frac{\alpha T^2}{T + \beta} \tag{3.1}$$

where $V_G(0) = 1.17$ V is the silicon bandgap voltage at 0 K and $\alpha = 4.73E - 4$ V/K, $\beta = 636$ K are empirical constants. According to TSIVIDIS (1980), this is a best-fit model from experimental data valid from 0 to 400 K, which is in the temperature range of the circuits proposed in this thesis. The main difference from the linear V_G model is that this equation introduces a quadratic behavior to the bandgap voltage, as shown in Figure 3.1.

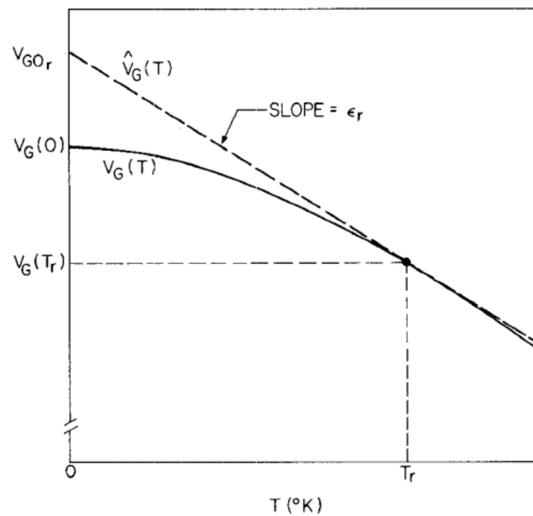


Figure 3.1: Bandgap voltage versus absolute temperature and its first degree approximation (not to scale) TSIVIDIS (1980).

In Figure 3.1 $V_G(T_r)$ is the bandgap voltage at the reference temperature T_r, while V_{G0r} is the linear extrapolation $\hat{V}_G(T)$ of $V_G(T_r)$ to 0 K. The collector current of an NPN transistor in the forward active region, neglecting the Early effect, is given by:

$$I_C(T) = I_S(T) \exp\left(\frac{qV_{BE}}{mkT}\right) \tag{3.2}$$

where T is the absolute temperature, I_C the collector current, V_{BE} the base-emitter voltage, m the non-ideality factor, k is the Boltzmann constant and q is the electron charge. We have extended the result to consider $m > 1$, since it can be relevant for CMOS processes. $I_S(T)$ is a process parameter that depends on the area of the base-emitter junction, the intrinsic carrier concentration and other device physics parameters, as given by:

$$I_S(T) = \frac{qAn_i^2(T)\bar{D}(T)}{N_B} \tag{3.3}$$

with A the base-emitter junction area, $n_i(T)$ the intrinsic carrier concentration, $\bar{D}(T)$ is the "effective" minority carrier diffusion constant in the base, and N_B the Gummel number (total number of impurities per unit area in the base). For a more detailed analysis, the reader is referred to TSIVIDIS (1980). From (3.2) we will derive an accurate expression for the junction voltage.

Consider two temperatures: an arbitrary temperature T and a specified reference temperature T_r. Applying (3.2) for each temperature, and re-writing for V_{BE} leads to:

$$V_{BE}(T) = \frac{T}{T_R}\left[V_{BE}(T_r) + \frac{mkT_r}{q}\ln\left(\frac{I_S(T_r)}{I_S(T)}\frac{I_C(T)}{I_C(T_r)}\right)\right] \tag{3.4}$$

This is the most general equation for the base-emitter voltage, before any approximations are made; it is accurate to the extent that (3.2) is. Defining an "effective" mobility, $\bar{\mu}(T)$ for minority carriers in the base

$$\bar{\mu}(T) = \frac{q\bar{D}(T)}{kT} \tag{3.5}$$

allows us to re-write (3.4) as follows:

$$\begin{aligned} V_{BE}(T) = V_G(T) - \left(\frac{T}{T_r}\right)V_G(T_r) + \left(\frac{T}{T_r}\right)V_{BE}(T_r) \\ + \frac{mkT}{q}\ln\left[\left(\frac{T_r}{T}\right)^4 \frac{\bar{\mu}(T_r)}{\bar{\mu}(T)}\frac{I_C(T)}{I_C(T_r)}\right] \end{aligned} \tag{3.6}$$

This equation is valid for any form of temperature dependence for $\bar{\mu}(T)$. However, if $\bar{\mu}(T)$ can be represent with sufficient accuracy by

$$\bar{\mu}(T) = CT^{-n} \tag{3.7}$$

with C and n appropriate constants, then (3.6) can be written as

$$\begin{aligned} V_{BE}(T) = V_G(T) - \left(\frac{T}{T_r}\right)V_G(T_r) + \left(\frac{T}{T_r}\right)V_{BE}(T_r) \\ -\eta\left(\frac{mkT}{q}\right)\ln\left(\frac{T}{T_r}\right) + \left(\frac{mkT}{q}\right)\ln\left[\frac{I_C(T)}{I_C(T_r)}\right] \end{aligned} \tag{3.8}$$

where $\eta = 4 - n$. We then proceed to a common special case, that assumes a collector current proportional to some power of T:

$$I_C(T) = FT^{\delta} \tag{3.9}$$

Where F and δ are constants. Using (3.9) in (3.8), we obtain a simplified expression for $V_{BE}(T)$:

$$V_{BE}(T) = V_G(T) - \left(\frac{T}{T_r}\right) V_G(T_r) + \left(\frac{T}{T_r}\right) V_{BE}(T_r) - (\eta - \delta)\left(\frac{mkT}{q}\right) \ln\left(\frac{T}{T_r}\right) \tag{3.10}$$

What (3.10) shows is that the base-emitter voltage is defined by process constants $V_G(T)$ and η, as well as design parameters $V_{BE}(T_r)$ and δ. The basic idea is that the base emitter voltage at the reference temperature is chosen according to the desired power consumption, and it roughly defines the average of $\partial V_{BE}/\partial T$, which is a linear term. There are two remaining non-linear terms that come from the silicon bandgap voltage temperature dependence and the $I_S(T)$ and $I_C(T)$ relationship.

3.2 BJT Curvature Correction

One way to implement curvature correction is to compensate for this non-linear terms with another non-linear voltage of opposite sign, or to subtract from V_{BE} another V_{BE} with the same non-linearities. Another approach, that is going to be used in this work, consists of adjusting the coefficient δ to cancel out η from the last non-linear term of (3.10).

In Figure 3.2(a) we present a comparison of $V_{BE}(T)$ using (3.10) under four different I_C conditions: a DC current ($\delta = 0$), a PTAT current ($\delta = 1$), a PTAT^2 current ($\delta = 2$), the ideal compensation current ($\delta = 3.5$) and a fourth exponent ($\delta = 4$). For all curves, the reference voltage is $V_{BE}(T_r) = 550$ mV at $T_r = 300$ K, $\eta = 2$ and $m = 1.075$. The collector current $I_C(T_r) = 3.5$ nA was determined from the X-FAB 180nm process for a 2μm x 2μm emitter area.

By looking at the derivative of $V_{BE}(T)$ in Figure 3.2(b) it becomes clear that the junction voltage can be linearized through an almost cubic collector current, or $\delta = 3.5$. We also note that the average derivative becomes slightly smaller as δ increases. The same result is presented in a similar way at MEIJER; SCHMALE; VAN ZALINGE (1982).

It is useful to define a quantitative metric for the linearity of $V_{BE}(T)$. We propose that this junction voltage be added to a linear PTAT term, as in the classical BGR, where the linearity will be measured indirectly through the remaining temperature coefficient of the voltage reference generated from it. The PTAT thermal coefficient γ can be calculated through

$$\gamma = -avg\left(\frac{\partial V_{BE}}{\partial T}\right) \tag{3.11}$$

and the resulting reference voltage is given by

$$V_{REF}(T) = V_{BE}(T) + \gamma T \tag{3.12}$$

The resulting reference voltage can be seen in Figure 3.3. We can then compare the linearity of the proposed currents in a quantitative way, using TC_{EFF} from (2.4). This is shown in Table 3.1.

Clearly, the non-linearity of the junction can be decreased by increasing the temperature dependency of the collector current. Moreover, there is a crossing point where the junction will be 'over-compensated', and that is around $\delta = 3.5$.

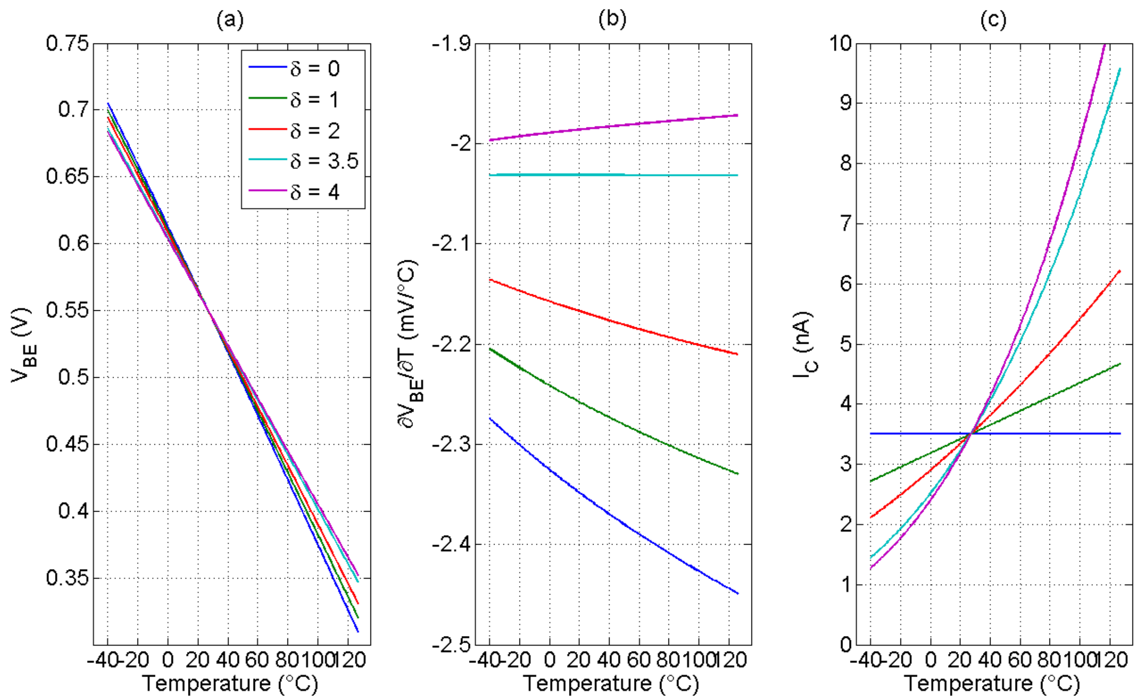


Figure 3.2: Different δ values comparison. (a) V_{BE}; (b) first derivative; and (c) I_C versus temperature.

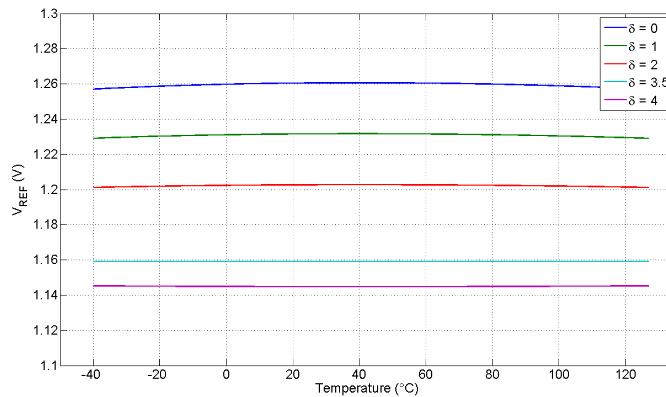


Figure 3.3: Resulting V_{REF} for different δ values comparison.

3.3 Fabrication Variability

Up to now we have assumed nominal process conditions for the variables n, δ and $V_{BE}(T_r)$. They are of course process dependent, so we now proceed to verify the impact of each one of them in the total effective temperature coefficient of the resulting voltage reference.

Each variable $V_{BE}(T_r)$, δ and n was considered independent and uncorrelated, meaning they have been varied individually, where we assumed a gaussian distribution with standard deviation over the mean $\sigma/\mu = 1\%$. The mean values μ are the ones used for the nominal case of the previous section. Figure 3.4 shows the spread of the output voltage V_{REF} for 1000 Monte Carlo runs, where the PTAT coefficient γ is a constant calculated from the nominal case, as would be done for a real BGR design.

Table 3.1: Linearity comparison for different δs.

δ	0	1	2	3.5	4	Unit
TC_{EFF}	17.4	12.8	7.6	0.04	2.7	ppm/°C

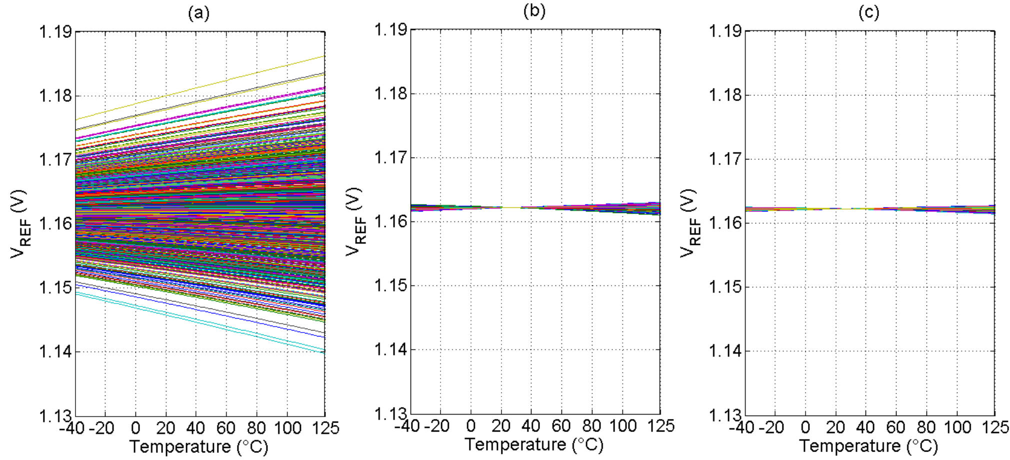


Figure 3.4: Impact of the design parameters variability on V_{REF}. (a) $V_{BE}(T_r)$; (b) δ and (c) n, for $\sigma/\mu = 1\%$.

Clearly, each design variable has a very different impact on the reference voltage thermal stability. It also becomes evident that the most relevant variable is the junction voltage at the reference temperature $V_{BE}(T_r)$, since it not only defines the absolute value of V_{REF}, but also the average derivative of the CTAT term V_{BE}.

To summarize, Tables 3.2, 3.3 and 3.4 present the yield of the circuit for different variability conditions, assuming that a maximum TC of 25 ppm/°C is desired.

From this analysis it is possible to conclude that curvature compensation is a useful technique, but hard to implement when accounting for process variations. To illustrate this, an ideal constant current source biasing the junction at 550 mV at room temperature presents σ/μ = 0.4 %, due only to variability in the device. As shown in table 3.4, the yield falls sharply for a $\sigma/\mu > 0.5$ %. There are also many other factors that will influence the value of $V_{BE}(T_r)$ such as packaging stresses and long term drifts (aging) FRUETT; MEIJER (2001) and FRUETT; MEIJER; BAKKER (2003), normally outside the control

Table 3.2: Yield (%) for δ = 1.

σ/μ (%)	$V_{BE}(T_R) = 550$ mV	$\delta = 1$	$n = 2$
0.1	100	100	100
0.5	99.5	100	100
1	84.3	100	100
2.5	41.8	100	100
5	20.2	100	100
10	12.1	98.8	83

Table 3.3: Yield (%) for δ = 2.

σ/μ (%)	$V_{BE}(T_R) = 550$ mV	$\delta = 2$	$n = 2$
0.1	100	100	100
0.5	99.9	100	100
1	88.5	100	100
2.5	43.8	100	100
5	24.6	100	99.9
10	13.9	87.7	89.3

Table 3.4: Yield (%) for δ = 3.5.

σ/μ (%)	$V_{BE}(T_R) = 550$ mV	$\delta = 3.5$	$n = 2$
0.1	100	100	100
0.5	99.9	100	100
1	88.4	100	100
2.5	51	100	100
5	23	96.6	100
10	13.3	71.6	89.9

of the circuit designer. Bottom line is: it does not make much sense to compensate for μV deviations due to non-linearities, when the nominal value at room temperature varies in the order of tens of mV.

4 PROPOSED CIRCUITS

4.1 BJT Bias and Curvature Compensation

In the previous section the linearization of $V_{BE}(T)$ based on the collector current thermal dependence was discussed. The discussion was based on a theoretical model, that needs to be adapted when designing BGRs for standard digital CMOS processes, where the available bias terminal is the emitter. One way of doing that would be to insert β into the equations, including its thermal dependence. That would require the extraction of the process parameters for the whole desired temperature range, which can be complicated.

Instead, we now introduce a novel circuit that can be sized for the chosen $V_{EB}(T_r)$ and δ independently, extracting the process parameters for T_r only, and relying on the complete SPICE model to simulate the behavior of β.

4.1.1 Circuit Concept

The basic concept of our topology MATTIA; KLIMACH; BAMPI (2014d) is shown in Figure 4.1. In this circuit, the BJT junction voltage is counterbalanced by the gate-source voltage of N stacked nMOS transistors. The resulting V_{GS} defines the BJT bias current, through a feedback path that uses a current mirror with gain K. By defining N and K, a non-zero equilibrium DC operating point can be reached, that reflects the current-voltage behavior of both the BJT and the MOSFETs.

The emitter current I_E of a BJT is given by:

$$I_E = I_{SE} exp\left(\frac{V_E}{m\phi_t}\right) \tag{4.1}$$

where I_{SE} is the reverse saturation current, V_E is the emitter-base voltage and m represents the slope factor. The temperature dependence of each parameter is implicit, unless otherwise specified. Using the ACM MOSFET model CUNHA; SCHNEIDER; GALUP-MONTORO (1998), the drain current of a transistor operating saturated and in subthreshold regime can be simplified to:

$$I_D = 2eI_{SQ}S \exp\left(\frac{V_G - V_{T0}}{n\phi_t} - \frac{V_S}{\phi_t}\right) \tag{4.2}$$

where the variables have their usual meaning. Appendix B shows the detailed model. Substituting (4.1) and (4.2) into the equality $I_E = KI_{D1}$, and solving for the junction voltage V_E leads to:

$$V_E = \frac{\phi_t}{\left(\frac{1}{m} - \frac{1}{nN}\right)}\left[ln\left(2eK\frac{W}{L}\frac{I_{SQ}}{I_{SE}}\right) - \frac{V_{T0}}{n\phi_t}\right] \tag{4.3}$$

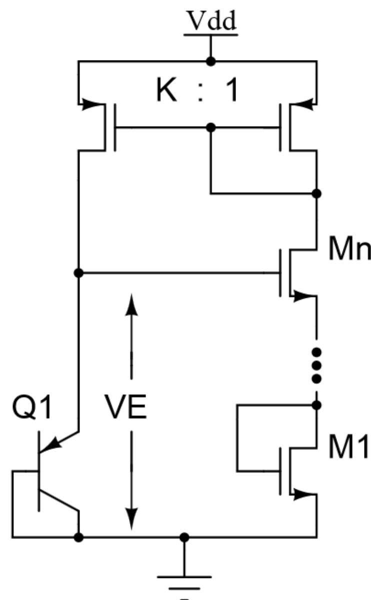


Figure 4.1: Schematic of the bias circuit concept.

By changing the number of stacked transistors N and the current gain K, different $\partial V_E/\partial T$ derivatives can be generated. The MOSFETs drain current will be heavily dependent on the number N. As an example, consider the circuits of Figure 4.2, where $N = 2$ and $N = 3$.

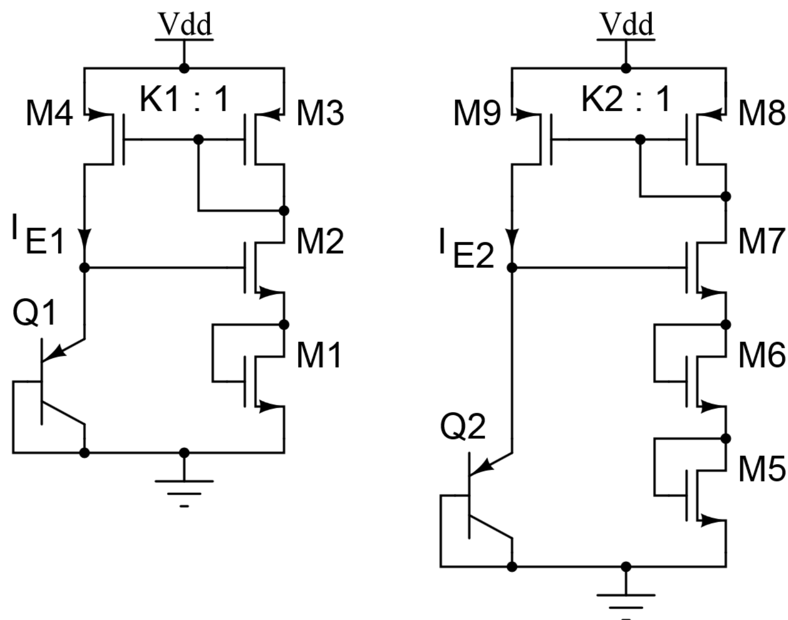


Figure 4.2: Schematic of different bias circuit examples. (a) $N = 2$; (b) $N = 3$

Suppose we want to design this circuit for a junction voltage of 550 mV at room temperature. In X-FAB 0.18μm technology, the resulting emitter currents versus temperature are shown in Figure 4.3.

Since M5-7 work under much lower inversion levels than M1,2, the current mirror gain $K_2 = 19.6$ while $K_1 = 1$. Current I_{E2} has a much higher temperature dependence

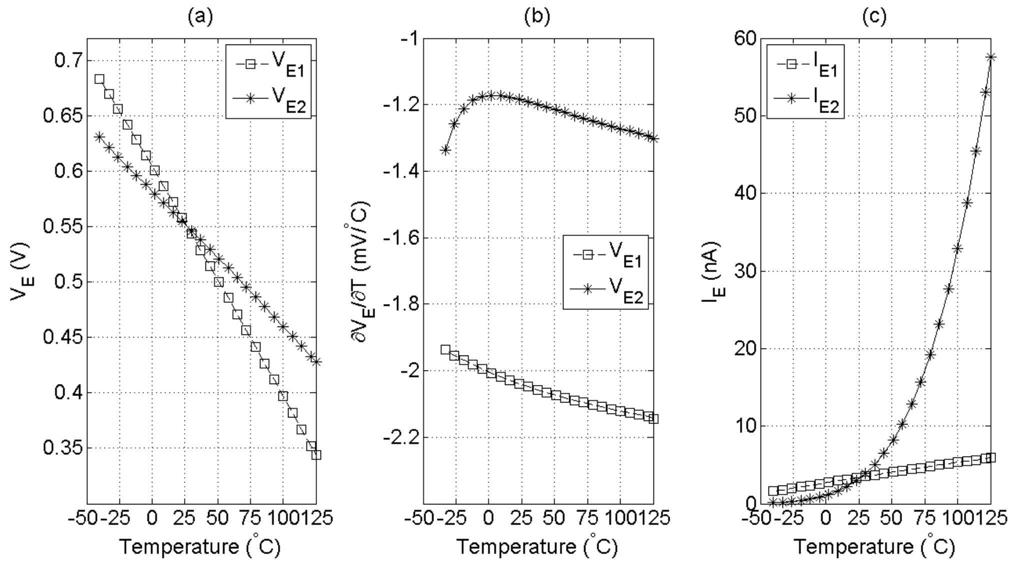


Figure 4.3: Behavior over temperature. (a) V_{E1} and V_{E2}; (b) $\partial V_{E1}/\partial T$ and $\partial V_{E2}/\partial T$; (c) I_{E1} and I_{E2}.

than I_{E1}, because of the three stacked nMOS, but still both are quite inadequate if compared to the ideal linearization current of Figure 3.2. However, with the circuit proposed in the next session these two currents can be combined with appropriate gains, making it possible to adjust δ and, therefore, to compensate for the junction curvature.

4.1.2 Proposed Circuit

The proposed circuit is shown in Figure 4.4. In this topology the BJT emitter current I_E is defined by the sum of two currents, a bias current I_{bias} and a compensation current I_{comp}. These currents are generated by a counterbalance between the gate-source voltage of stacked nMOS transistors M1-M5, where the resulting drain currents I_{D2} and I_{D5} are fed back to the emitter through current mirrors M6-M9.

All transistors are operating in weak inversion, saturated and have the same aspect ratios. As long as the current mirror gains K_1 and K_2 have roughly the same order of magnitude, the drain current contribution of M1-M2 will dominate over the drain current contribution of M3-M5 at room temperature, since the former work under much higher inversion levels than the latter. If that is the case, $V_{BE}(T_r)$ can be defined by sizing M1-M2 and K_1, while δ is determined through the sizing of M3-M5.

Substituting (4.1) and (4.2) into the equality $I_E(T_r) = K_1 I_{D1}(T_r)$ and solving for the junction voltage at room temperature leads to:

$$V_E(T_r) \approx \frac{\phi_t}{\left(\frac{1}{m} - \frac{1}{2n}\right)} \left[\ln \left(2eK_1 \frac{W_1}{L_1} \frac{I_{SQ}(T_r)}{I_{SE}(T_r)} \right) - \frac{V_{T0}(T_r)}{n\phi_t} \right] \tag{4.4}$$

From the approximate equation (4.4) it can be seen that the emitter voltage at T_r depends on process parameters that define the circuit's equilibrium point, namely: the MOSFET threshold voltage V_{T0} and specific current I_{SQ}, and the junction reverse saturation current I_{SE}. The extraction of these parameters at room temperature provides a reasonable accuracy to determine $V_E(T_r)$.

A complete set of equations for the analytical design of variable δ is a complex derivation that includes two feedback paths and the temperature dependent equations of several

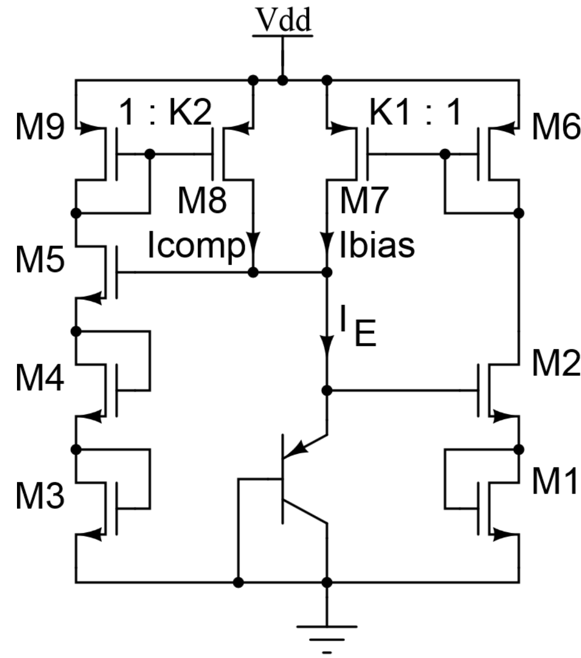


Figure 4.4: Self-biased curvature compensation circuit.

process parameters I_{SQ}, I_{SE}, V_{T0} and β, to name a few. Instead, we propose the following methodology:

1. extract process parameters V_{T0}, I_{SQ}, I_{SE}, n and m through simulation or experimental data at room temperature T_r only;
2. size $S_{1,2}$ and K_1 to achieve the desired $V_E(T_r)$, through the approximate expression (4.4);
3. use SPICE simulation to fine tune S_{3-5} and K_2 for the maximum linearity ($\delta = 3.5$).

This methodology has been successfully implemented in two different processes with three different MOSFET types. As an example, in Figure 4.5(a) and (b) we present a comparison of the junction voltage, and in Figure 4.5(c) the collector current, for both the analytical model of chapter 3 and the SPICE simulation data using the X-FAB 0.18μm process.

Clearly, the model matches the simulation results very well. A small difference in the junction voltage at 27 °C has caused a higher derivative for the SPICE simulation, and it is also slightly less linear, probably due to the approximated analytical silicon bandgap expression. To conclude the demonstration of the circuit's concept and potential, we now compare it with the traditional approaches of biasing the junction with a DC or a PTAT current. These currents, when added to an ideal compensation PTAT voltage, would result in the references shown in Figure 4.6. Here, V_{DC} and V_{PTAT} correspond to a junction biased with a DC and PTAT currents, respectively. The V_{IDEAL} corresponds to the circuit of Figure 4.4 implemented with ideal current mirrors, while V_{REAL} is the implementation with real pMOS current mirrors.

The resulting temperature coefficients are 20.15, 14.93, 1.79 and 4.6 ppm/°C for the DC, PTAT, ideal current mirrors and real implementation circuits, respectively.

Table 4.1 presents a comparison of the analytical model of chapter 3 versus the SPICE simulation results of XFAB 180nm and IBM 130nm circuits.

We have thus presented a resistorless BJT bias and curvature compensation circuit

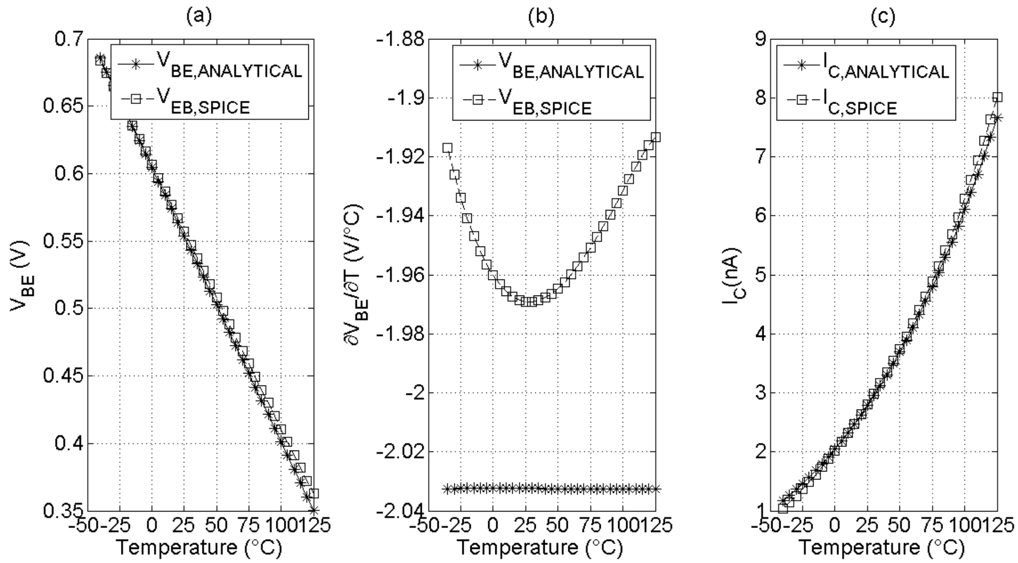


Figure 4.5: Analytical and SPICE results comparison versus temperature: (a) junction voltage; (b) first derivative; (c) collector current

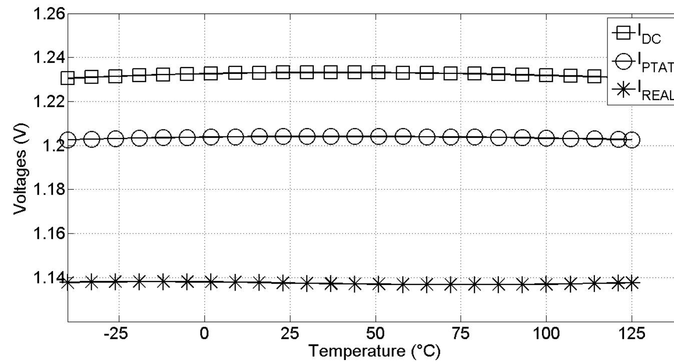


Figure 4.6: SPICE results comparing V_{REFs} with different junction bias currents.

that works in the nanoampere range. If this circuit is to implement a bandgap reference, it obviously cannot work with a supply lower than 1 V. In the next session we detail the emitter voltage division that enables a lower reference voltage, and introduce the PTAT cells used to compose the complete sub-bandgap reference circuit.

4.2 Sub-Bandgap Voltage Reference

Referring to Figure 4.7 we identify the same bias circuit of the previous section, but without the curvature compensation branch. We have shown that curvature compensation is a very difficult thing to implement successfully due to process variations. We have opted for a sub-BGR design without compensation mainly because we want to investigate the performance of the circuit without any calibration scheme.

On Figure 4.7, M1 and M2 have the same drain current, so voltage V_E appears divided at the gate of M1. This voltage is then added to a PTAT voltage generated by three self-cascode (SC) PTAT structures M2-M7 to provide a temperature independent output V_{REF}. To better understand the analysis, we start by explaining the SC PTAT generator, followed

Table 4.1: Linearity performance (TC_{EFF}) of analytical model and SPICE results.

TC_{EFF}	I_{DC}	I_{PTAT}	I_{REAL}	
Analytical	17.40	12.80	0.04	ppm/°C
XFAB 180nm	28.06	16.94	12.83	ppm/°C
IBM 130nm I/O FET	20.15	14.93	5.44	ppm/°C
IBM 130nm LP FET	20.15	14.93	4.61	ppm/°C

by the BJT bias and V_E divider. The equation for V_{REF} is derived last.

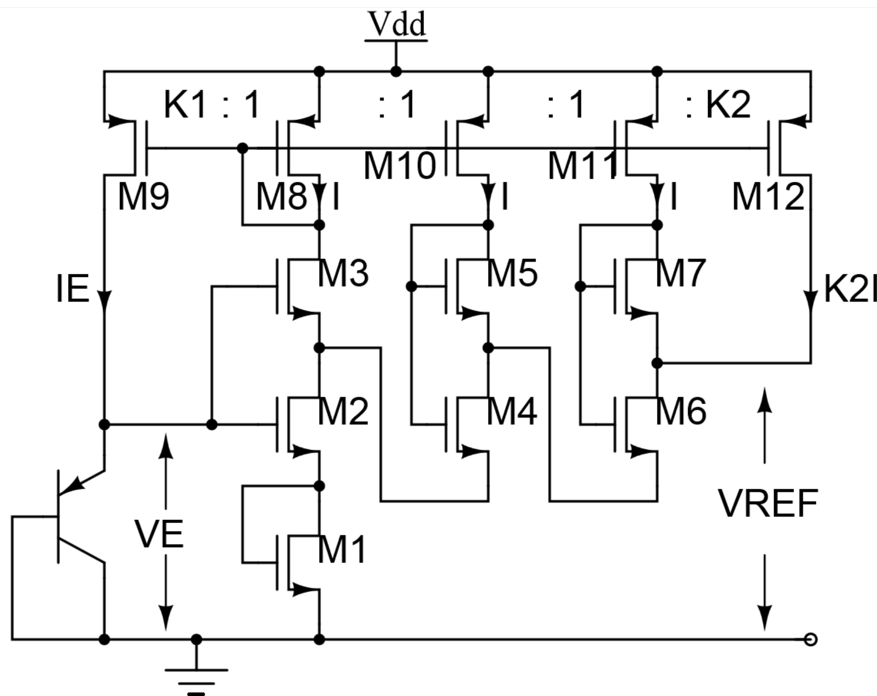


Figure 4.7: Schematic of the proposed sub-BGR circuit.

4.2.1 Self-Cascode PTAT Generator

According to the ACM MOSFET model CUNHA; SCHNEIDER; GALUP-MONTORO (1998), the drain current of a transistor operating in subthreshold regime is given by

$$I_D = 2eI_{SQ}S \exp\left(\frac{V_G - V_{T0}}{n\phi_t} - \frac{V_S}{\phi_t}\right)\left[1 - \exp\left(-\frac{V_{DS}}{\phi_t}\right)\right] \tag{4.5}$$

where the terms have their usual meaning, and have been defined before. In this work, the MOSFET is considered saturated under weak inversion operation when $V_{DS} > 4\phi_t$.

A PTAT voltage can be generated by the traditional self-cascode structure VITTOZ; NEYROUD (1979). In the circuit of Figure 4.7, three SC pairs are used, composed of transistors M2-M7. All MOSFETs are operating in the subthreshold region, where the here called 'lower' transistor (M2, M4 and M6) can be in triode or in saturation, while the 'upper' one (M3, M5 and M7) must be saturated. The difference of their gate-source voltages appear across the drain-source terminals of the lower transistor,

and this voltage is proportional to the thermal potential. Using (4.5) and knowing that $V_{DS(lower)} = V_{GS(lower)} - V_{GS(upper)}$ leads to (4.6).

$$V_{DS(lower)} = \phi_t \ln\left(\frac{I_{lower}}{I_{upper}}\frac{S_{upper}}{S_{lower}} + 1\right) \tag{4.6}$$

The logarithmic dependence limits the PTAT voltage that can be generated by a single SC cell. If a high temperature derivative is needed, several pairs must be cascaded, as done in the circuit of Figure 4.7. The total PTAT voltage generated is given by the sum of the drain-source voltages of M2, M4 and M6. The proportionality constant inside the logarithm can be scaled by adjusting the current across each device and their aspect ratios, as shown in (4.7).

$$V_{PTAT} = \phi_t \ln\left[\left((K_2+3)\frac{S_3}{S_2}+1\right)\left((K_2+2)\frac{S_5}{S_4}+1\right)\left((K_2+1)\frac{S_7}{S_6}+1\right)\right] \tag{4.7}$$

Where $K_2 = \frac{I_{D12}}{I_{D11}}$ is the current gain defined by current mirrors M8-M12.

4.2.2 V_E Divider and BJT Bias Circuit

Referring again to Figure 4.7, we can see that $V_{G2} = V_E$ and $I_{D1} = I_{D2}$. The use of (4.5) then leads to the complete expression for the gate voltage of M1, given by (4.8).

$$V_{G1} = \frac{V_E - n\phi_t \ln\left[\frac{S_1}{S_2} + \frac{S_1}{(K_2+3)S_3}\right]}{n+1} \tag{4.8}$$

What equation (4.8) shows is that the emitter voltage will be divided by approximately 2.3 ($n \approx 1.3$ for WI), then subtracted by a term proportional to the thermal voltage defined by the aspect ratio of transistors M1, M2, M3 and current gain K_2. The V_E division is thus fairly insensitive to process variations, since it depends mainly on geometrical factors, and on the subthreshold factor. In a practical design, $(K_2 + 3)S_3 \gg S_2$ is used, since the first SC stage must produce a reasonable PTAT voltage. If $S_1 = S_2$ is applied for simplicity, one can see that (4.8) reduces to (4.9).

$$V_{G1} \approx \frac{V_E}{n+1} \tag{4.9}$$

Substituting (4.1), (4.5) and (4.9) into the equality $I_E = K_1 I_{D8}$, also noting that $I_{D8} = I_{D1}/(K_2 + 3)$ and solving for the junction voltage V_E leads to (4.10).

$$V_E \approx \frac{m(n+1)}{n(n+1)-m}\left[n\phi_t \ln\left(2e\frac{I_{SQ}}{I_{SE}}\frac{K_1}{K_2+3}S_1\right) - V_{T0}\right] \tag{4.10}$$

From the approximate equation (4.10) the main design variable $V_E(T_r)$ can be calculated, once process parameters are determined at T_r.

4.2.3 Reference Voltage

The reference voltage output is thus the sum of the divided junction voltage at the gate of M1 (4.9), plus the PTAT voltage (4.7), and it is given by (4.11).

$$V_{REF} \approx \frac{V_E}{n+1} + V_{PTAT} \tag{4.11}$$

The first term of (4.11) represents the CTAT voltage, while the second term is a constant that multiplies ϕ_t. We now discuss some trade-offs regarding the power consumption and area of the circuit.

4.2.4 Design Methodology

This topology has a wide design space, and some limitations have to be imposed on the available variables for a real implementation. The choice of aspect ratios and current mirror gains is restricted to integer values that provide good layout regularity and facilitate common-centroid structures. There are trade-offs between area and power consumption, so we also define $S_3 = 10S_2$ to avoid a too large area for M3, and $K_2 = 2$ to provide sufficient gain for the PTAT cells while keeping the power consumption low. The smallest characterized BJT provided in the design kit is used, with a 2 x 2 μm^2 emitter area, which has the lowest I_E for a given V_E. The XFAB 180nm technology estimated parameters at 27 °C are $V_{T0}(T_r) = 435$ mV, $I_{SQ}(T_r) = 85$ nA, $n = 1.3$, $I_{SE}(T_r) = 8.9$ aA and $m = 1.05$. An exploration of the design space, done in Matlab, is shown in Figure 4.8. Variables K_1 and $S_1 = S_2$ are varied, and the resulting emitter voltage V_E is analyzed, together with currents I_E and I_{D1}.

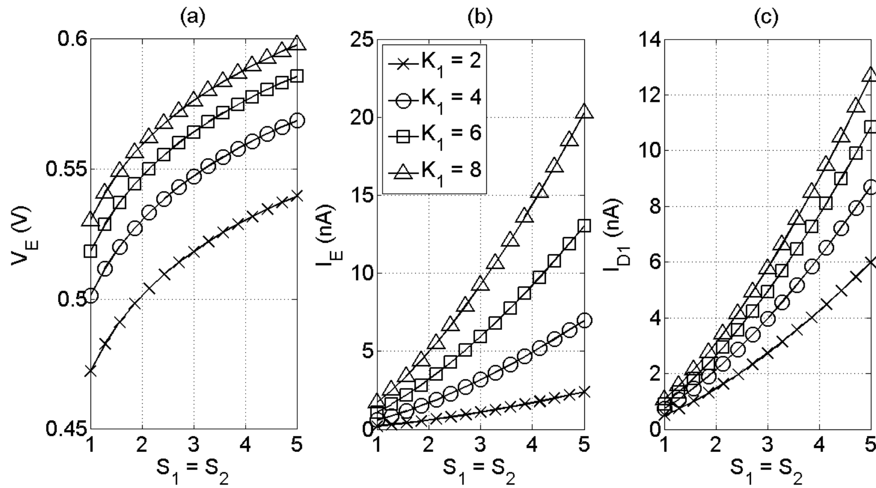


Figure 4.8: Design space exploration (a) V_E; (b) I_E; and (c) I_{D1}.

Low power consumption is an important metric used in this design, and a significant amount of current is due to the BJT bias. For this technology, the simulated emitter voltage thermal derivative diverges from an almost linear behavior for junction voltages lower than 520 mV. The designer then must pay attention to the minimal bias condition defined for V_E, since a low temperature coefficient of the reference voltage is also an important performance metric. Another restriction is that the minimum drain current for the MOSFETs must be much higher than the leakage current, which for this technology is in the order of hundreds of fA at 125 °C.

Considering these restrictions, we define $K_1 = 4$ and $S_1 = S_2 = 4$, resulting in an emitter voltage of approximately 550 mV, where $I_E \approx 3.5$ nA and $I_{D8} \approx 600$ pA at 27 °C. Simulations of the circuit over temperature provide the CTAT derivative, and the SC cells can then be sized to compensate it. The first SC cell has already been defined by S_3/S_2 and K_2, while the remaining PTAT voltage is provided by sizing the remaining two cells through S_4-S_7.

5 SIMULATION AND MEASUREMENT RESULTS

The proposed circuits have been implemented in two different CMOS processes, XFAB 180nm and IBM 130nm. We have experimented with standard, low-power and I/O transistors to demonstrate the portability of the topologies, and will present some of these results in the next section.

All of the laid out circuits presented here follow good practices such as common-centroid and regular structures, guard-rings, dummies, no metal connections over gates and so on. Also, all of the available measurement pins have ESD diode protections that were simulated accordingly and should not affect the working principle of the circuits, despite the very low power consumption achieved.

We have received 40 samples from an IBM 130 nm fabrication run, where the dies are divided into 10 unpackaged and 30 packaged on a QFN80. Only the unpackaged samples have been measured since we are waiting for the printed circuit board to measure the packaged dies. The fabricated chip contains the works of 7 M.Sc. and Ph.D. students and can be seen on Figure 5.1.

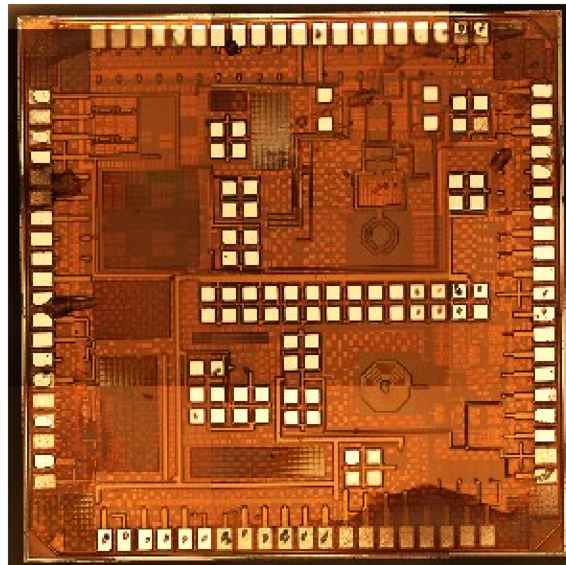


Figure 5.1: Micrograph of the overall chip fabricated in IBM 130 nm with an area of 2.5 x 2.5 mm^2.

5.1 BJT Bias and Curvature Compensation

We repeat the schematic of the circuit on Figure 5.2 for the convenience of the reader.

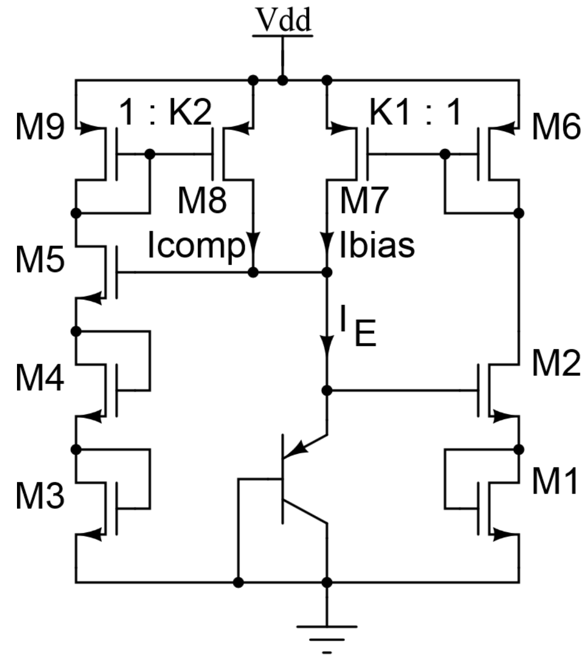


Figure 5.2: Self-biased curvature compensation circuit.

5.1.1 Simulation Results

The following results are for post-layout extraction simulations on SPICE using low-power transistors from the IBM PDK. These devices have V_{T0} = 577 mV and I_{SQ} = 140 nA at room temperature. The implemented layout has an area of 0.00067 mm^2, as shown in Figure 5.3.

The circuit was designed for a junction voltage of 550 mV at room temperature, as shown in Figure 5.4. The power consumption is 2.5 nW at room temperature and reaches 8.8 nW at 125 °C for V_{DD} = 0.8 V.

The sensitivity of the circuit against power supply variations is presented in Figure 5.5. The chosen low-power transistors have a poor output conductance characteristic, and the PSRR measured at 100 Hz is -33 dB. The line sensitivity is 19.78 mV/V for V_E and 1.1 nA/V for I_E, both for V_{DD} varying from 0.7 to 1.2 V.

The noise measured at the junction emitter at 100 Hz and room temperature is 1.6 μV/$\sqrt{Hz}$, as shown in Figure 5.6.

To analyze the fabrication variability of the circuit, Monte Carlo (MC) simulation was done separately for local mismatch effects and average process variations, with 100 runs each. For average process MC, all the transistors have their parameters changed equally in each run. For local mismatch MC, the parameters of each transistor are varied individually in each run. Both effects are taken into account in a combined variability analysis. Figure 5.7 (a) shows the spread of the junction voltage, with a $\sigma/\mu = 5.1\%$ for mean process variation, while local mismatch, shown in Figure 5.7 (c), yields $\sigma/\mu = 0.31\%$. A combined variability analysis yields $\sigma/\mu = 4.12\%$ for V_E. Despite the very small area, it is clear that the major problem in the proposed bias circuit is the average process variation, which has much larger impact in MOSFETs operating in the subthreshold condition.

These detailed results are then summarized in Table 5.1 for another MOSFET of the IBM 130 nm process, and the standard XFAB 180nm device as well. All circuits employ the same junction emitter area of 2 x 2 μm^2 and have similar performance, demonstrating

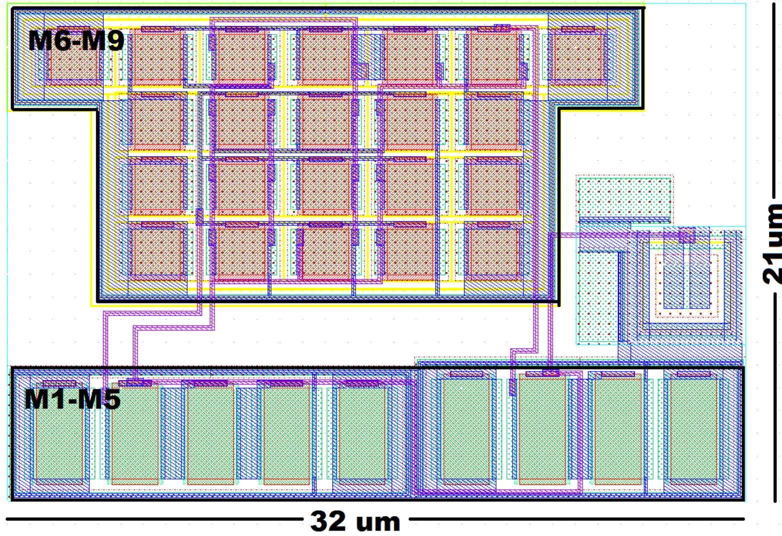


Figure 5.3: Layout of the proposed bias and curvature compensation circuit in IBM 130nm.

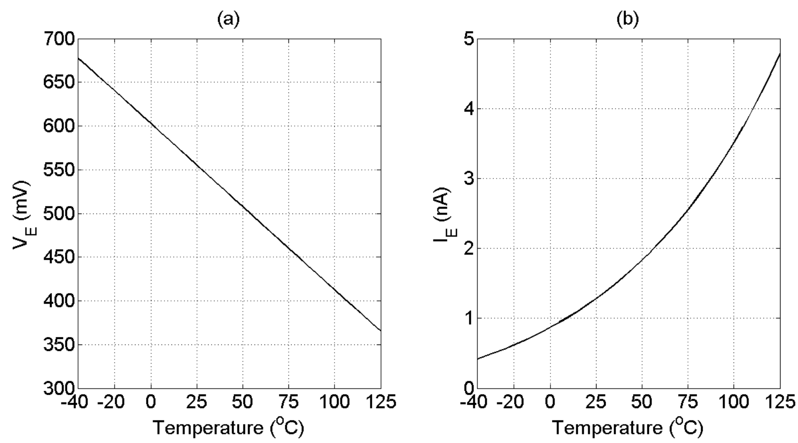


Figure 5.4: (a) V_E; and (b) I_E over temperature, V_{DD} = 0.8 V.

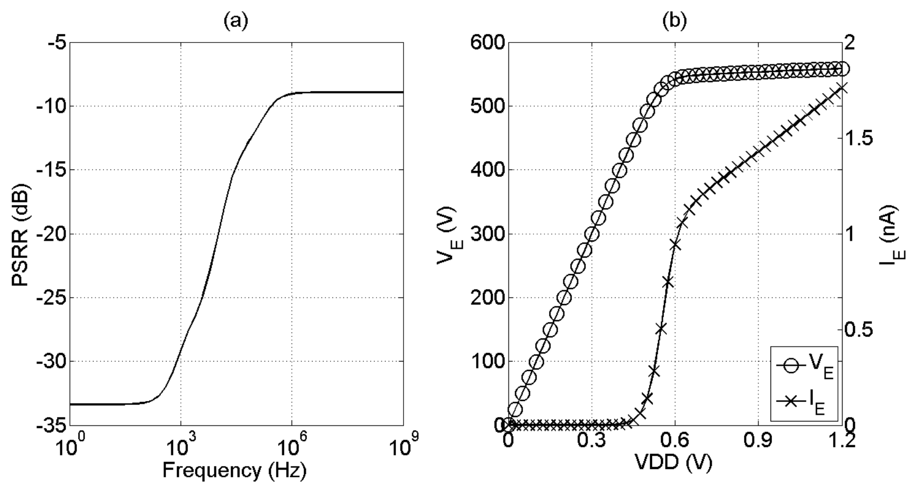


Figure 5.5: (a) PSRR versus frequency; (b) V_E and I_E versus power supply, 27°C.

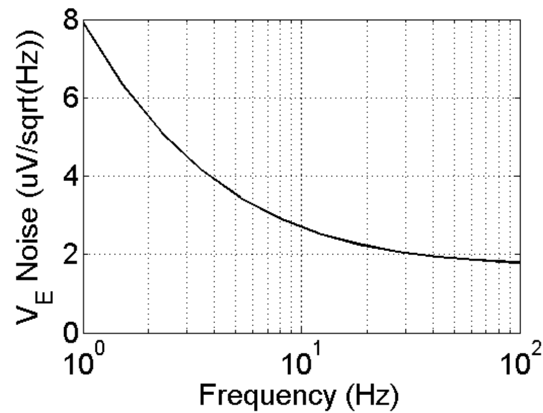


Figure 5.6: V_E noise versus frequency at 27 °C and V_{DD} = 0.8 V.

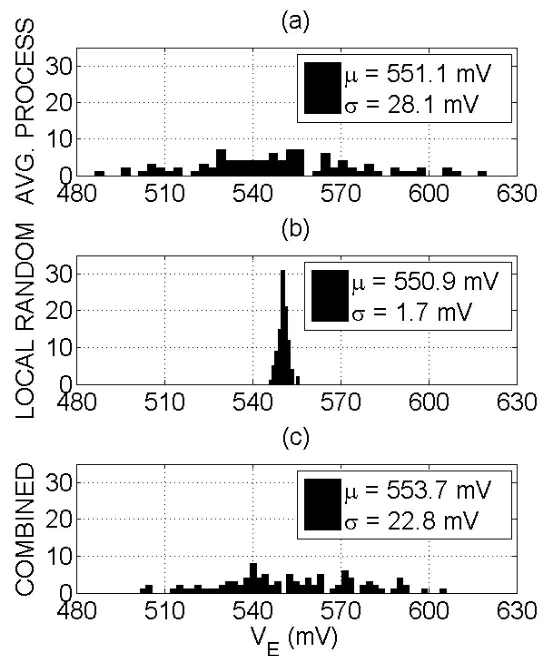


Figure 5.7: V_E Monte Carlo results from 100 runs. Average process variation on (a); Local random mismatch on (b); and Combined analysis on (c). V_{DD} = 0.8 V.

the portability of the circuit.

5.1.2 Experimental Results

5.1.2.1 LP Devices at Room Temperature

The following results are for 10 unpackaged bare dies measured at 21 °C. The fabricated circuit uses the LP device from IBM 130nm. A section of the fabricated chip can be seen in Figure 5.8, with the highlighted circuit area, which is much smaller than an I/O pad, for example.

Due to package limitations an independent V_{DD} pin was not available to measure the power consumption of the circuit, so it was estimated from the junction voltage and current gain K_1. It has a mean of 10 nA at 0.8 V, leading to a power consumption of 8 nW.

Table 5.1: Performance comparison of the BJT Bias simulated circuits.

Technology	XFAB 180nm	IBM 130nm	IBM 130nm	Unit
MOSFET	Regular	I/O	Low-Power	–
$V_{T0}(T_r)$	435	481	577	mV
$I_{SQ}(T_r)$	85	111	140	nA
Emitter Area	2 x 2	2 x 2	2 x 2	μm x μm
$V_E(T_r)$	552	547	551	mV
$I_E(T_r)$	3.84	1.13	1.32	nA
V_{DD}	0.7 – 1.8	0.7 – 1.2	0.7 – 1.2	V
Power ($V_{DD} = 0.8V$, 27°C)	4.32	2.55	3.12	nW
Linearity	12.0	5.4	4.6	ppm/°C
Noise @ 100 Hz	1.3	1.5	1.6	μV/$\sqrt{Hz}$
LS (V_E)	2.0	9.1	19.78	mV/V
PSRR @ 100 Hz	-48.5	-38.8	-33.0	dB
Area	0.00129	0.00057	0.00067	mm^2
$V_E(\sigma/\mu)$ @ 27 °C	2.76	4.03	4.12	%

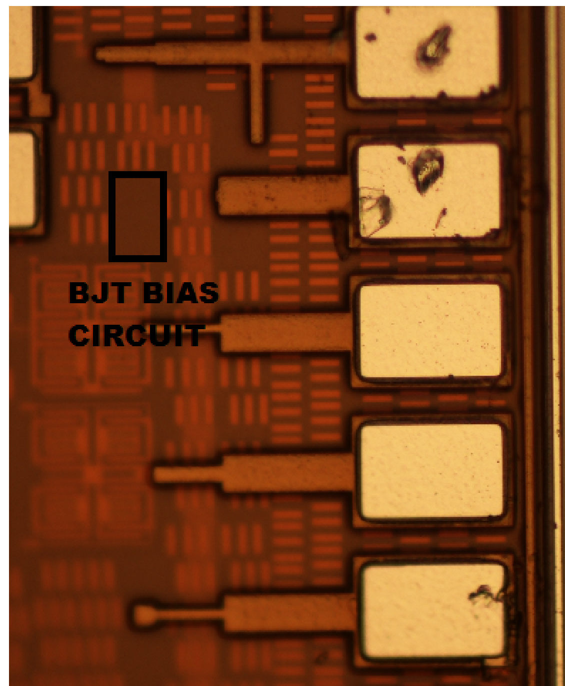


Figure 5.8: Detail of the fabricated chip and the proposed BJT bias circuit area.

In Figure 5.9 we show that the minimum V_{DD} is around 0.7 V, while on Figure 5.10 one can see in detail the fabrication spread and the sensitivity of the junction voltage against V_{DD}.

On average, shown in Figure 5.11 V_E = 610 mV, with a σ = 8.7 mV. The average LS of V_E is 17 mV/V from 0.7 to 1.2 V.

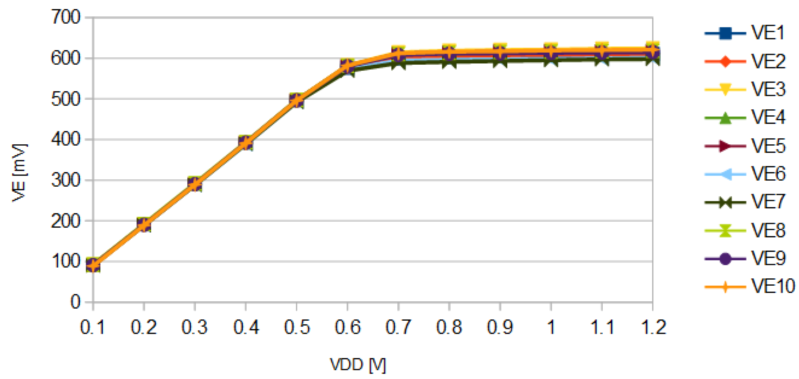


Figure 5.9: V_E vs. V_{DD} @ 21°C.

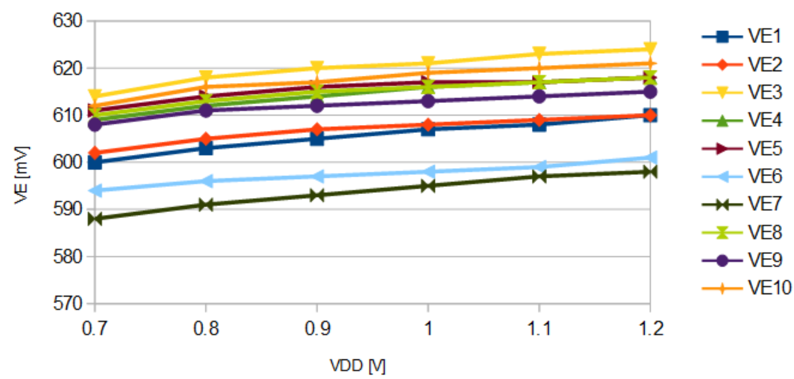


Figure 5.10: V_E vs. V_{DD} @ 21°C, from 0.7 V to 1.2 V.

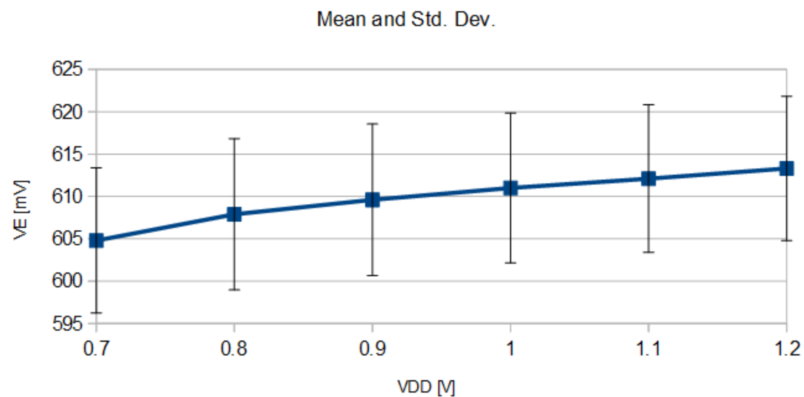


Figure 5.11: V_E vs. V_{DD}, Mean and Std. Dev. @ 21°C.

5.1.2.2 I/O Devices versus Temperature

The behavior of the circuit was then characterized versus temperature, using a probe station and a thermal chuck. The following results are for one unpackaged bare die using the I/O MOSFETs from the 130nm PDK. Figure 5.12 shows the emitter voltage for a temperature range of 0 to 125 °C and a supply variation of 0.6 V to 1.2 V with 25 mV steps. The junction voltage is 585 mV at room temperature, while the average temperature derivative is -1.56 mV/°C.

The total power consumption of the circuit is presented in Figure 5.13, being 5 nA at room temperature and reaching 27 nA at 125 °C under a supply of 0.7 V.

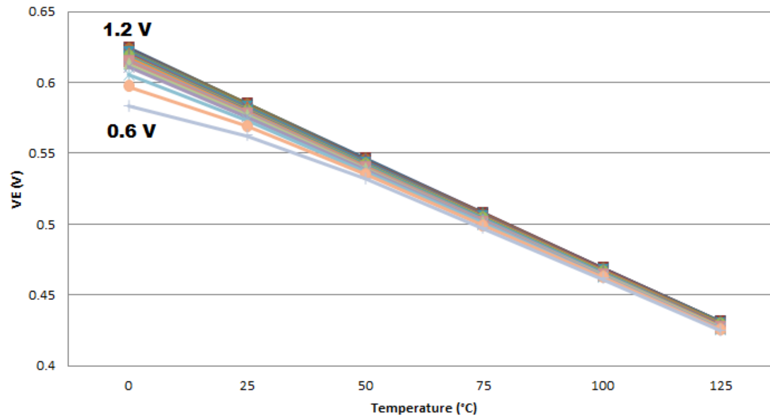


Figure 5.12: V_E vs. Temperature, with varying V_{DD}.

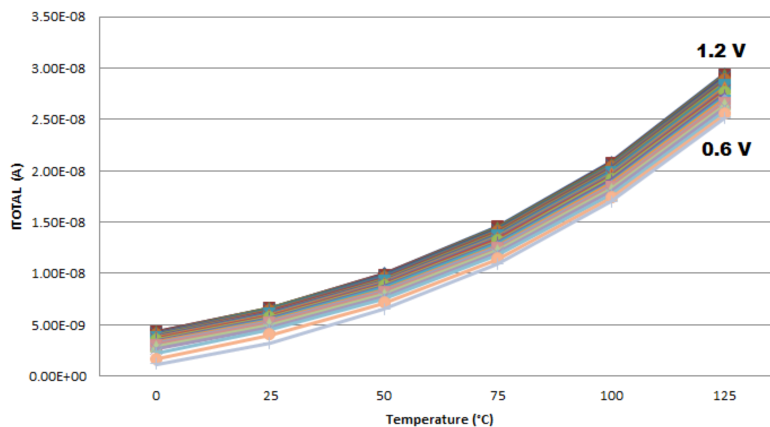


Figure 5.13: I_{TOTAL} vs. Temperature, with varying V_{DD}.

5.2 Sub-Bandgap Voltage Reference

Again we repeat the schematic of the circuit on Figure 5.14 for the convenience of the reader.

5.2.1 Simulation Results

The results presented here are for post-layout simulations using standard nMOSFETs in XFAB 180nm. These transistors have V_{T0} = 435 mV and I_{SQ} = 85 nA at room temperature. The silicon area, already considering common-centroid structures and dummies, is shown in Figure 5.15. It is a very small topology, occupying only 0.0012 mm^2.

The voltage reference obtained is around 479 mV, with a slight curvature due to the non-linearity of the BJT's emitter voltage, as shown in Figure 5.16 (a). The effective temperature coefficient is 8.79 ppm/^{o}C for the 0 to 125 ^{o}C temperature range, with V_{DD} = 0.9 V.

In Figure 5.16 (b) we show the V_E voltage, that is divided by approximately two, then added to V_{PTAT} to form the temperature independent V_{REF}. Figure 5.16 (c) presents the currents in each branch. It is 5.4 nA for the whole circuit at 27 ^{o}C, reaching a maximum of 20.5 nA at 125 ^{o}C, which leads to a power consumption of 4.9 nW and 18.5 nW, respectively, under V_{DD} = 0.9 V. Startup behavior of the circuit was simulated, having a settling time of less than 200 μs, which is acceptable for our proof of concept. We expect

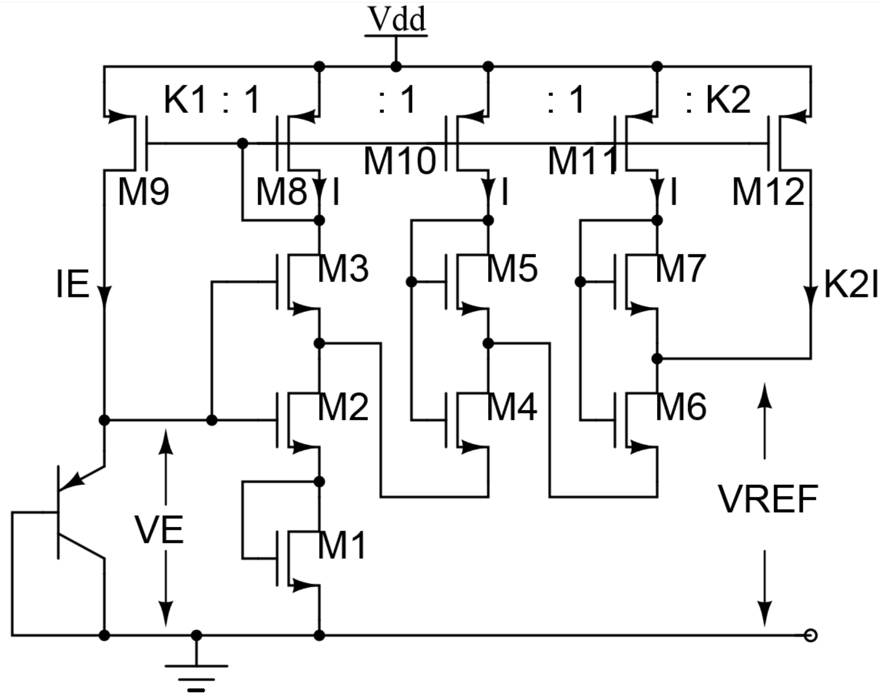


Figure 5.14: Schematic of the proposed sub-BGR circuit.

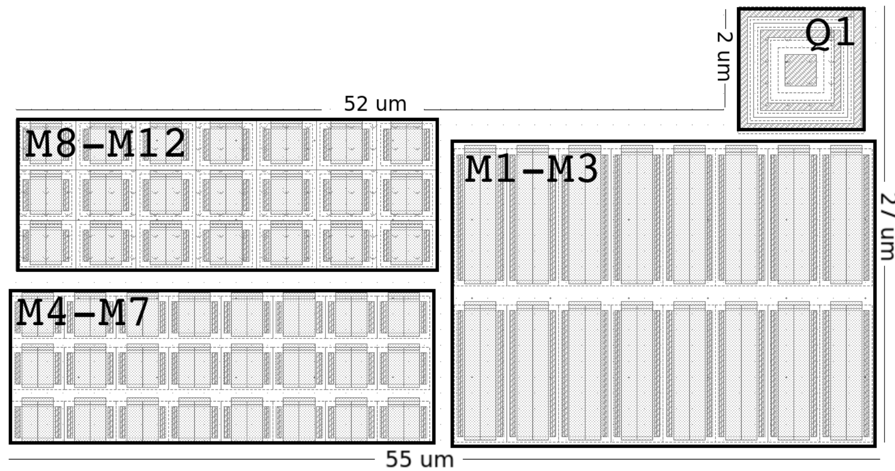


Figure 5.15: Layout floorplanning of the proposed sub-bandgap reference in XFAB 180nm.

that leakage currents are enough to start the circuit, but in real applications a startup circuit could be necessary for faster settling. The noise, shown in Figure 5.17 equals 1.57 μV/$\sqrt{Hz}$ simulated at 100 Hz.

PSRR measured at 100 Hz and V_{DD} = 0.9 V, is -48 dB - Figure 5.18 (a). The effective temperature coefficient is minimum for the optimal V_{DD} of 0.9 V, and reaches a maximum of 16 ppm/°C at 1.8 V. Even though the nominal supply voltage of the process used is 1.8 V, this implementation starts operating around 0.85 V, as shown in Figure 5.18 (c). The line sensitivity of V_{REF} is 2.112 mV/V from 0.85 V to 1.8 V, while the current consumption sensitivity is 69 pA/V - Figure 5.18 (c). Below 0.85 V, the PTAT cells are not working yet, and have zero drain current, which makes the feedback loop increase the current in the M1-M3 branch to achieve equilibrium with the junction voltage V_E.

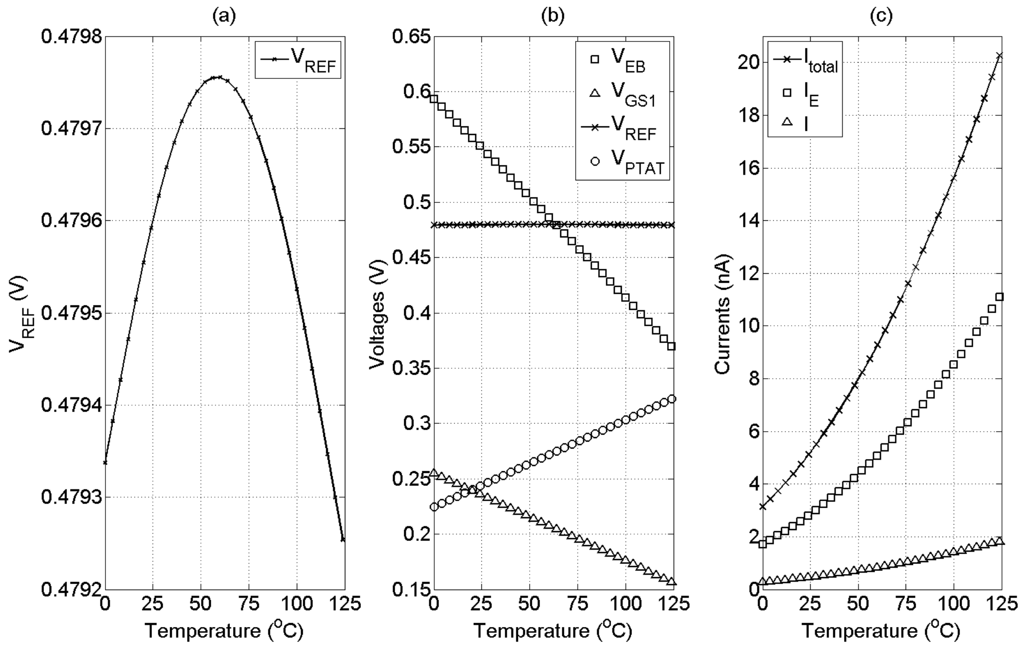


Figure 5.16: (a) V_{REF}; (b) V_E, V_{GS1} and V_{PTAT} voltages; (c) I_{TOTAL}, I_E and I currents over temperature. V_{DD} = 0.9 V.

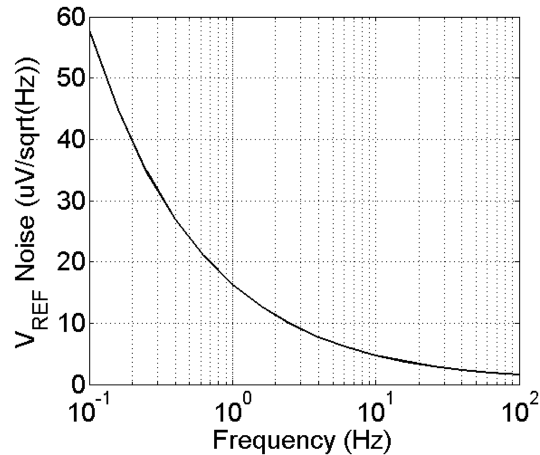


Figure 5.17: V_{REF} noise vs. frequency.

Both line sensitivity and PSRR could be increased by adding cascode current sources, for example, at the penalty of increasing the minimum supply voltage.

To analyze the fabrication variability of the circuit, Monte Carlo (MC) simulation was done separately for local mismatch effects and average process variations, with 100 runs each. For average process MC, all the transistors have their parameters changed equally in each run. For local mismatch MC, the parameters of each transistor are varied individually in each run. Figure 5.19 (a) shows the spread of the reference voltage, with a $\sigma/\mu = 2\%$ for mean process variation, while local mismatch, shown in Figure 5.19 (c), yields $\sigma/\mu = 0.8\%$. Figure 5.19 (b) presents the spread of the temperature coefficient, where 96 % of the parts yield a TC below 50 ppm/°C for mean process variations, while for local mismatch only the spread is much smaller and all parts have TC < 12.5 ppm/°C. Despite the very small area, it is clear that the major problem in the proposed reference

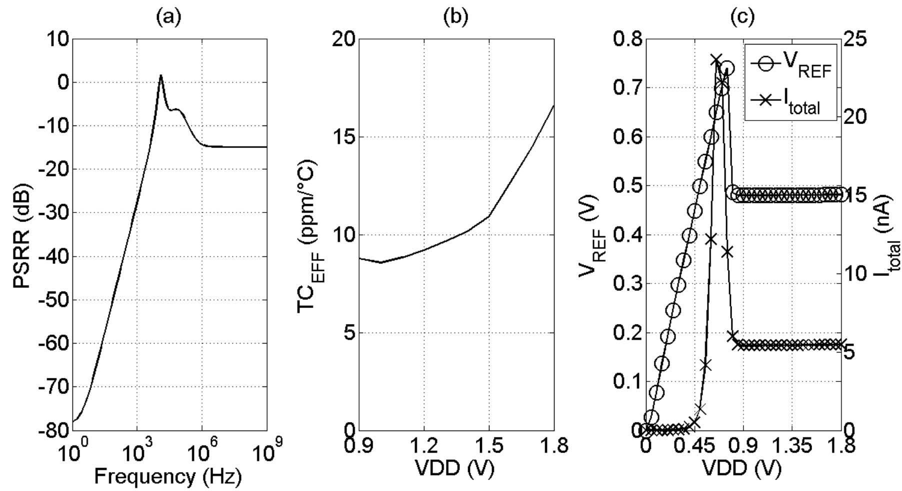


Figure 5.18: (a) PSRR versus frequency; (b) TC_{EFF}; (c) V_{REF} and I_{TOTAL} vs. V_{DD}.

is the average process variation, which has much larger impact in MOSFETs operating in the subthreshold condition.

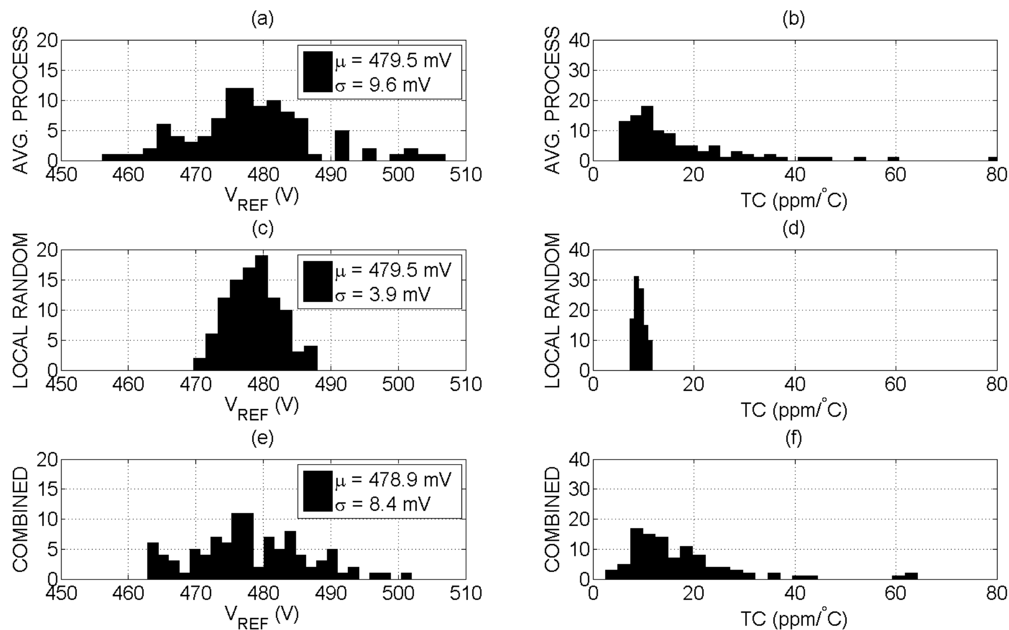


Figure 5.19: V_{REF} and TC_{EFF} Monte Carlo results. Average process variation on (a) and (b); Local mismatch on (c) and (d);

The detailed results are again summarized in Table 5.2 for a MOSFET of the IBM 130 nm process and the standard XFAB 180nm device as well. All circuits employ the same junction emitter area of 2 x 2 μm^2 and have similar performance, demonstrating the portability of the circuit.

5.2.2 Experimental Results

We wanted to present simulation results from another process to show the portability of the circuit with similar results. The experimental results we present next are for the IBM 130nm I/O device. The layout of the resulting reference in 130nm can be seen in

Table 5.2: Performance comparison of the Sub-BGR simulated circuits.

Technology	XFAB 180nm	IBM 130nm	–
MOSFET	Regular	I/O	–
$V_{T0}(T_r)$	435	481	mV
$I_{SQ}(T_r)$	85	111	nA
V_{REF}	479	550	mV
TC (0 - 125°C)	8.79	14	ppm/°C
Noise @ 100 Hz	1.57	2.85	μV/$\sqrt{Hz}$
V_{DD}	0.9 – 1.8	0.9 – 1.2	V
Power @ $V_{DD,min}$, 27°C	4.9	4.5	nW
LS	2.11	1.65	mV/V
PSRR @ 100 Hz	-48	-41	dB
Area	0.0012	0.0022	mm^2
$V_{REF}(\sigma/\mu)$	2	2.59	%
Yield (TC < 50 ppm/°C)	96	71	%

Figure 5.20, occupying an area of 0.0021 mm^2.

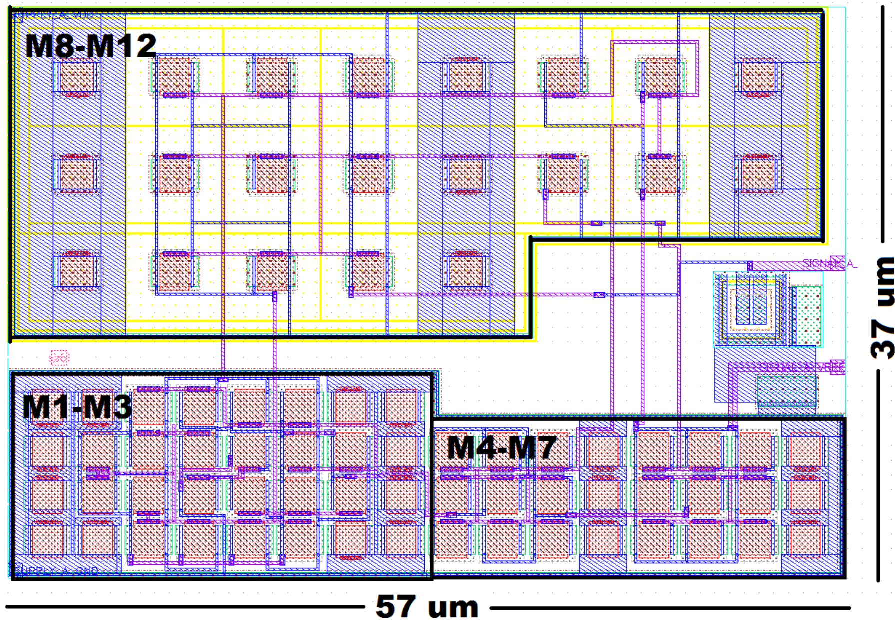


Figure 5.20: Fabricated layout of the proposed sub-bandgap reference in IBM 130nm.

The 10 samples have been measured as unpackaged bare dies with a probe station and the semiconductor parameter analyzer 4145B from HP. A highlight of the fabricated circuit can be seen in Figure 5.21, which is again much smaller than an I/O pad and the associated ESD protection.

Finally, Figures 5.22, 5.23, 5.24, 5.25, 5.26, 5.27, 5.28, 5.29 and 5.30 present the

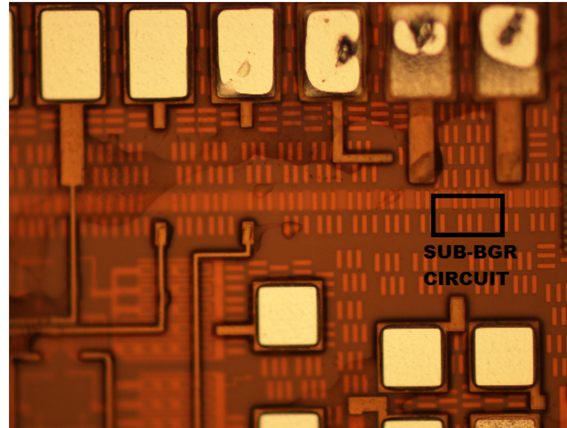


Figure 5.21: Detail of the fabricated chip and the proposed sub-BGR circuit area.

behavior versus supply voltage at 21 °C for the available variables V_E, V_{REF} and I_{TOTAL}.

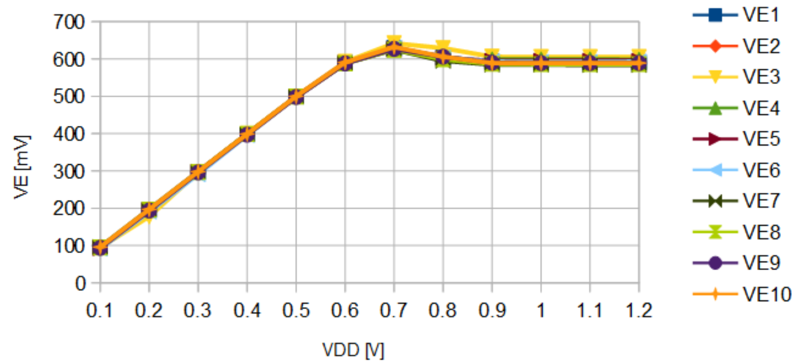


Figure 5.22: V_E vs. V_{DD} @ 21°C.

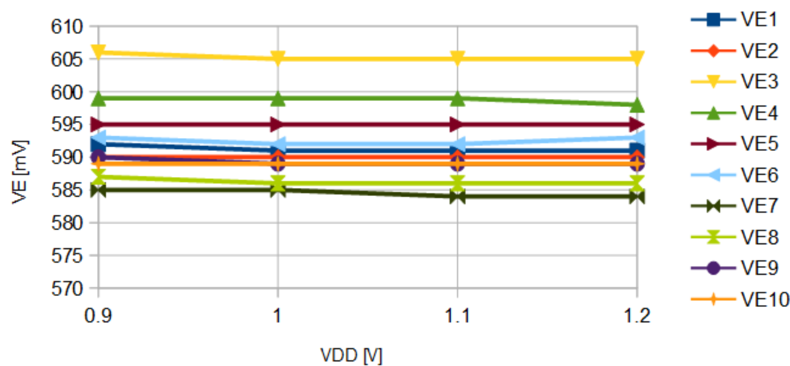


Figure 5.23: V_E vs. V_{DD} @ 21°C, from 0.9 V to 1.2 V.

All circuits work with the minimum supply of 0.9 V, and have a mean current consumption of 9 nA. The average V_{REF} measured was 550 mV, with $\sigma = 6.7$ mV and average line sensitivity of 3.3 mV/V.

5.2.2.1 Temperature Measurements

The voltage reference was then measured against variations in temperature and in power supply. Three unpackaged bare dies were measured, where Figure 5.31 shows the

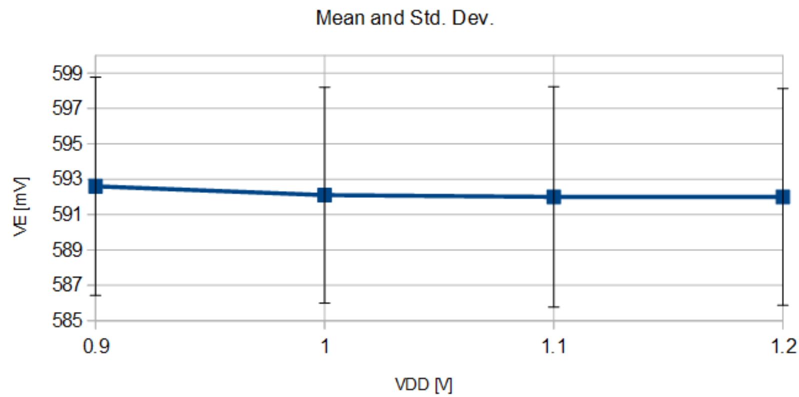


Figure 5.24: V_E Mean and Std. Dev. vs. V_{DD} @ 21°C.

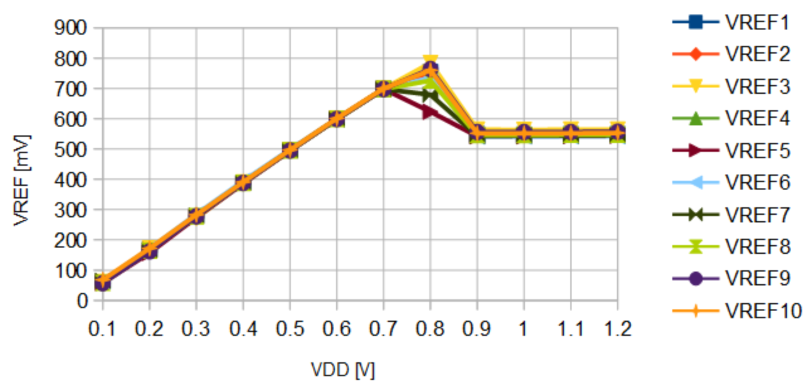


Figure 5.25: V_{REF} vs. V_{DD} @ 21°C.

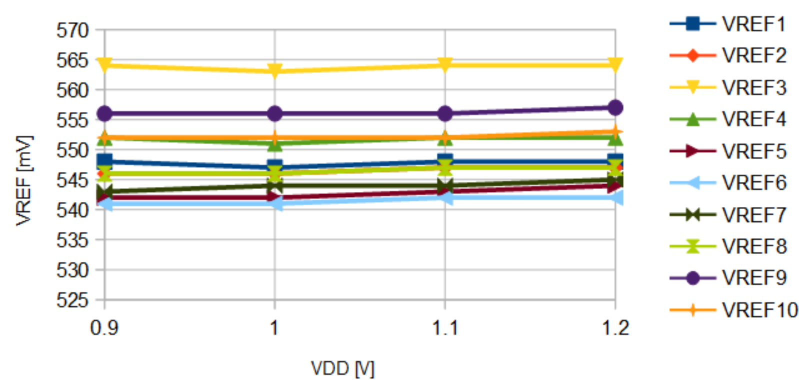


Figure 5.26: V_{REF} vs. V_{DD} @ 21°C, from 0.9 V to 1.2 V.

behavior of the reference voltage under such conditions for the first die.

As can be seen, the circuit suffers a maximum variation of under 4 mV, if a minimum supply of 925 mV is considered. The variation increases for lower temperature because the minimum supply is higher, due to a higher junction voltage. The same graph is presented for the other two dies in Figures 5.32 and 5.33.

The temperature coefficient of these references can then be calculated, and it is shown as a function of the supply voltage on Figure 5.34. It is higher for supply voltages that approach the 0.9 V, as expected. Still, the minimum TC measured was 11 ppm/°C while

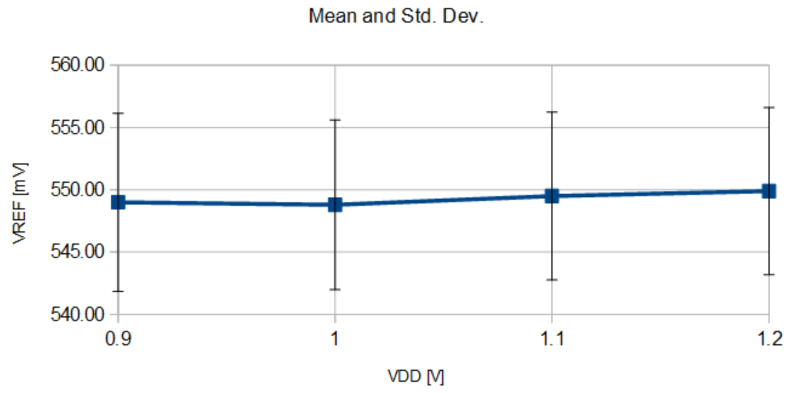


Figure 5.27: V_{REF} Mean and Std. Dev. vs. V_{DD} @ 21°C.

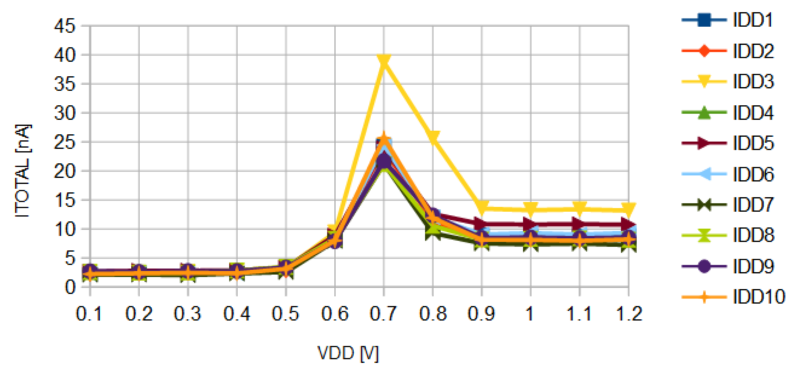


Figure 5.28: I_{TOTAL} vs. V_{DD} @ 21°C.

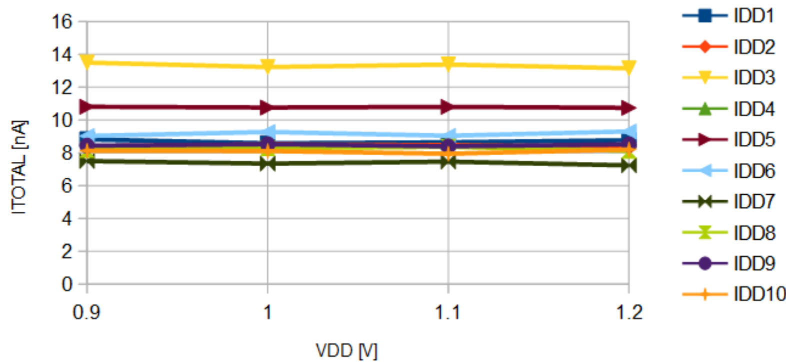


Figure 5.29: I_{TOTAL} vs. V_{DD} @ 21°C, from 0.9 V to 1.2 V.

the maximum was 86 ppm/°C.

The measured coefficient is in agreement with the values predicted by simulation, and below 80 ppm/°C for two samples and below 25 ppm/°C for the best die. Finally, the total power consumption against temperature and supply variations are presented in Figure 5.35, again for the three measured dies.

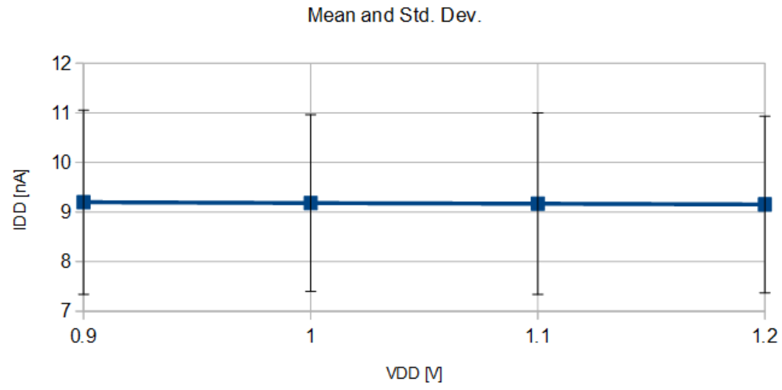


Figure 5.30: I_{TOTAL} Mean and Std. Dev. vs. V_{DD} @ 21°C.

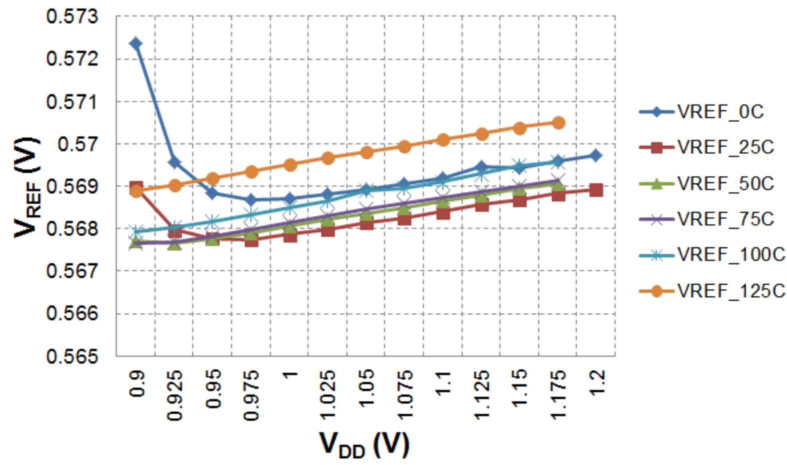


Figure 5.31: V_{REF} vs. Temperature, with varying V_{DD} - Die # 1.

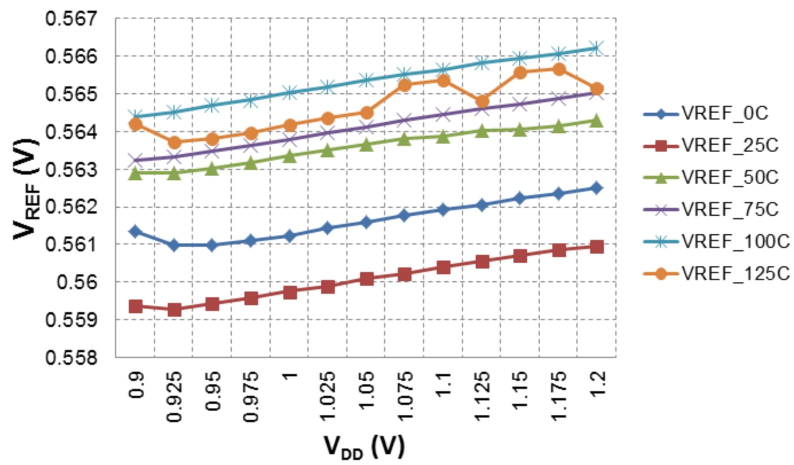


Figure 5.32: V_{REF} vs. Temperature, with varying V_{DD} - Die # 2.

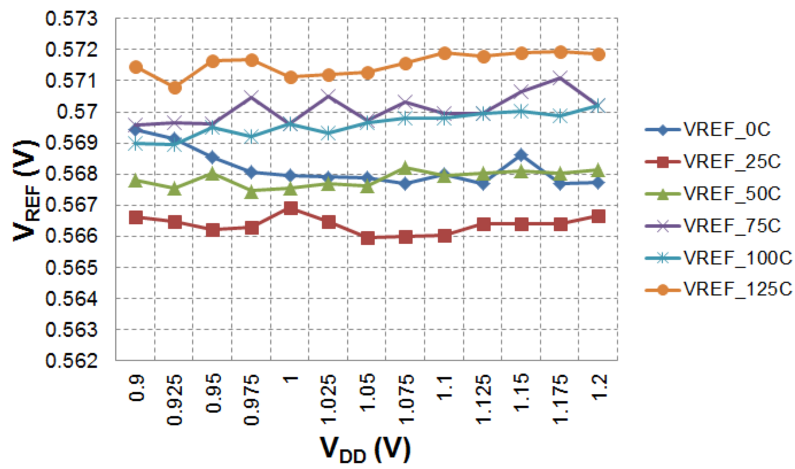


Figure 5.33: V_{REF} vs. Temperature, with varying V_{DD} - Die # 3.

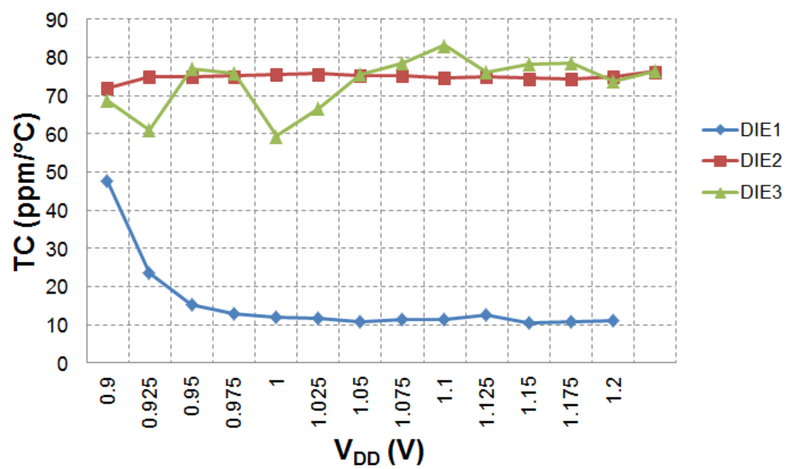


Figure 5.34: TC_{EFF} vs. V_{DD} - Dies # 1-3.

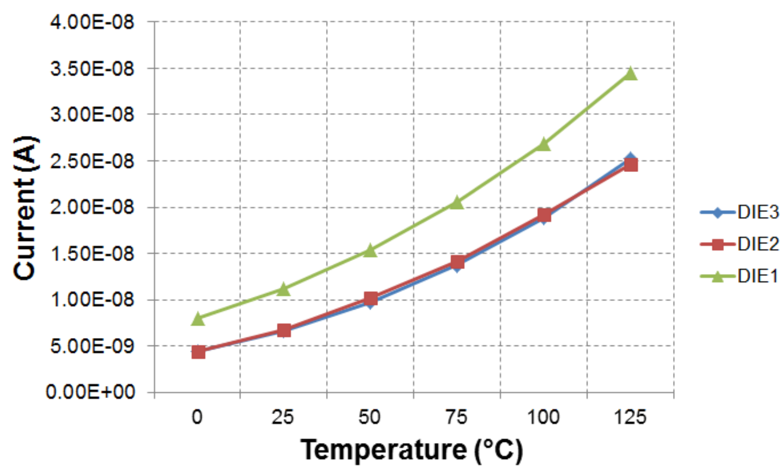


Figure 5.35: I_{TOT} vs. Temperature, with varying V_{DD} - Dies # 1-3.

6 CONCLUSION

This thesis has covered the most recent applications of voltage references in CMOS technologies, specifically those that operate with sub-1 V supply, nanowatt power consumption and are implemented in resistorless processes. The main applications for these types of circuits are found in portable battery-operated and energy harvesting devices, smart sensors and identification tags like RFID.

We have then detailed the non-linearity present in the traditional bandgap voltage reference, and explained how it can be corrected through the use of heavily temperature dependent collector currents. A novel circuit topology was introduced to generate such a current, and a design methodology was presented that implements curvature compensation independently of the desired bias voltage. Simulation and experimental results using Low Power FETs from IBM 130nm show that it operates from 0.7 V supply, consuming 3.1 nW (for an emitter voltage of 550 mV) at room temperature and occupying a silicon area of 0.00067 mm^2. The emitter voltage varies up to σ/μ = 4.1 % due to the MOSFETs threshold voltage dependency. Similar simulation results were obtained using I/O FETs from the same process, and regular FETs from the XFAB 180nm process. Experimental results at room temperature from 10 unpackaged dies of the same run show a mean emitter voltage of 610 mV, which is within the expected average process variation, and with σ/μ = 1.4 %. The same circuit implemented with I/O FETs was characterized against temperature for one unpackaged bare die, providing a mean junction voltage derivative over temperature of 1.56 mV/°C.

Derived from this bias circuit was a junction voltage divider and consequent sub-bandgap reference using the well known self-cascode cell. We demonstrated a voltage reference of 479 mV that consumes 4.9 nW under a 0.9 V supply, simulated in XFAB 180nm. The main advantages are the low temperature coefficient of 9 ppm/°C for a 0 to 125 °C temperature range, with a silicon area of 0.0012 mm^2. Monte Carlo simulations show that the expected spread of the reference voltage is σ/μ = 2 % for a combined variability analysis. The same circuit was designed and fabricated on IBM 130nm using I/O FETs, and experimental results at room temperature from 10 unpackaged dies of the same run show a mean reference of 550 mV with σ/μ = 1.3 %. Temperature measurements of three samples have shown a minimum temperature coefficient of 11 ppm/°C and maximum of 86 ppm/°C, consuming less than 10 nW at room temperature and less than 40 nW at maximum temperature.

Comparisons with the state of the art are some times difficult because of the scarcity of data, especially regarding variability results. We did our best to provide a fair comparison in Table 6.1, while more detailed results were presented in the beginning of the thesis.

It is fair to say that our topology has the potential to improve significantly the state of the art in important metrics such as absolute voltage accuracy and temperature coefficient

Table 6.1: Comparison of recent resistorless CMOS Voltage References.

(*) simulation results	De Vita 2007	Ueno 2009	Ming 2010	Magnelli 2011	Seok 2012	Osaki 2013	**This Work**	Unit
Technology	0.35	0.35	0.5	0.18	0.18	0.18	**0.13**	μm
CTAT Voltage	V_{T0}	V_{T0}	V_{EB}	V_{T0}	V_{T0}	V_{EB}	V_{EB}	–
Temperature Range	0 – 80	-20 – 80	-40 – 120	0 – 125	-20 – 80	-40 – 120	**0 – 125**	°C
Number of samples	20 same batch	17 same batch	5 same batch	40 three batches	49 (22 + 27) two batches	9 same batch	10 (3) same batch	
V_{REF}								
min	607	733	–	240	175	524	**540**	mV
avg	670	745	1.23[1]	263.5	176	551	**550**	mV
max	732	759	–	280	181	577	**565**	mV
σ/μ (Monte Carlo)	–	7.15*	–	–	–	1.61*	**2.59***	%
Temperature Coefficient								
min	–	7	–	39	16.9	52	**11**	ppm/°C
avg	10	15	11.8[1]	142	62	114	**–**	ppm/°C
max	–	45	15.6[1]	400	231	148	**83**	ppm/°C
Noise @ 100 Hz	–	–	–	2	16	1.9	**2.9***	μV/$\sqrt{Hz}$
V_{DD}	0.9 – 4	1.4 – 3	3.6	0.45 – 2	0.5 – 3	0.7 – 1.8	**0.9 – 1.2**	V
Power								
min	–	280	–	1	–	–	**6.3**	nW
nom	36	300	$6.48 \cdot 10^5$	3.15	0.0022	52.5	**8.1**	nW
max	–	350	–	15	0.081	–	**12.6**	nW
Line Sensitivity	1.83	0.015	–	1.2	0.077	–	**0.9**	mV/V
PSRR @ 100 Hz	-47	-45	-31.8	-45	-49	-56	**-41***	dB
Silicon Area	0.045	0.055	0.1	0.043	0.0014	0.0246	**0.0022**	mm^2

[1] individually trimmed dies

over a wide temperature range, while consuming one of the lowest currents and occupying the smallest area for a bandgap based reference by a factor of 10. These claims are backed by thorough simulation results and preliminar experimental measurements.

6.1 Future Work

There are, of course, several issues to be addressed in the developed topologies, especially if they are to improve significantly on the state of the art in terms of variability. Several other topologies were designed during the M.Sc. but were left out of this thesis, and they could use some improvement as well. We list some possible topics for a future work as:

1. Experimental temperature characterization of all the 40 dies received from IBM;
2. Run at least one more fabrication to address batch-to-batch variations;
3. Two V_{T0} extractor topologies were also developed but not included in this thesis. The extracted value could be used to compensate for variability in the BJT bias. One was published in MATTIA; KLIMACH; BAMPI (2014e) and a voltage reference was derived from it MATTIA; KLIMACH; BAMPI (2014c). The other topology remains unpublished so far;
4. A detailed study of the factors that impact the variability of BJTs and MOSFETs, aiming to reduce the average process dependence, that remains the most restrictive performance aspect;
5. Study and develop cost-efficient calibration/trimming schemes;
6. Provide some sort of "load" regulation, where the reference can source up to a determined amount of current;

7. Improve the line sensitivity through the use of cascode current mirrors or symmetrically matched current-voltage mirrors LAM; KI (2010).

REFERENCES

AITA, A. et al. Low-Power CMOS Smart Temperature Sensor With a Batch-Calibrated Inaccuracy of $\pm 0.25°C$ ($\pm 3\sigma$) From -70°C to 130°C. **Sensors Journal, IEEE**, [S.l.], v.13, n.5, p.1840–1848, May 2013.

ANNEMA, A.-J. Low-power bandgap references featuring DTMOSTs. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.34, n.7, p.949–955, Jul 1999.

BANBA, H. et al. A CMOS bandgap reference circuit with sub-1-V operation. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.34, n.5, p.670–674, May 1999.

BROKAW, A. A simple three-terminal IC bandgap reference. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.9, n.6, p.388–393, Dec 1974.

BUCK, A. et al. A CMOS bandgap reference without resistors. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.37, n.1, p.81–83, Jan 2002.

CUNHA, A.; SCHNEIDER, M.; GALUP-MONTORO, C. An MOS transistor model for analog circuit design. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.33, n.10, p.1510–1519, 1998.

DE VITA, G.; IANNACCONE, G. An ultra-low-power, temperature compensated voltage reference generator. In: CUSTOM INTEGRATED CIRCUITS CONFERENCE, 2005. PROCEEDINGS OF THE IEEE 2005, 2005. **Anais...** [S.l.: s.n.], 2005. p.751–754.

DE VITA, G.; IANNACCONE, G. A Sub-1-V, 10 ppm/C, Nanopower Voltage Reference Generator. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.42, n.7, p.1536–1542, 2007.

DE VITA, G.; IANNACCONE, G.; ANDREANI, P. A 300 nW, 12 ppm//spl deg/C Voltage Reference in a Digital 0.35 /spl mu/m CMOS Process. In: VLSI CIRCUITS, 2006. DIGEST OF TECHNICAL PAPERS. 2006 SYMPOSIUM ON, 2006. **Anais...** [S.l.: s.n.], 2006. p.81–82.

DEVICES, A. ADuCRF101: precision analog microcontroller arm cortex m3 with rf transceiver. **Data Sheet**, [S.l.], 2013.

FRUETT, F.; MEIJER, G. Experimental investigation of piezojunction effect in silicon and its temperature dependence. **Electronics Letters**, [S.l.], v.37, n.22, p.1366–1367, Oct 2001.

FRUETT, F.; MEIJER, G.; BAKKER, A. Minimization of the mechanical-stress-induced inaccuracy in bandgap voltage references. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.38, n.7, p.1288–1291, July 2003.

GUPTA, V.; RINCóN-MORA, G. A. A Low-impedance, Sub-bandgap 0.6 μm CMOS Reference with 0.84 % Trimless 3-σ; Accuracy and -30 dB Worst-case PSRR Up to 50 MHz. **Analog Integr. Circuits Signal Process.**, Hingham, MA, USA, v.62, n.3, p.345–359, Mar. 2010.

KLIMACH, H. et al. Resistorless switched-capacitor bandgap voltage reference with low sensitivity to process variations. **Electronics Letters**, [S.l.], v.49, n.23, p.1448–1449, 2013.

KUIJK, K. A precision reference voltage source. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.8, n.3, p.222–226, June 1973.

LAM, Y.-H.; KI, W.-H. CMOS Bandgap References with Self-biased Symmetrically Matched Current-voltage Mirror and Extension of Sub-1-V Design. **IEEE Trans. Very Large Scale Integr. Syst.**, Piscataway, NJ, USA, v.18, n.6, p.857–865, June 2010.

LEUNG, K. N.; MOK, P. A sub-1-V 15-ppm/ deg;C CMOS bandgap voltage reference without requiring low threshold voltage device. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.37, n.4, p.526–530, Apr 2002.

LEUNG, K. N.; MOK, P. A CMOS voltage reference based on weighted Delta;VGS for CMOS low-dropout linear regulators. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.38, n.1, p.146–150, Jan 2003.

MAGNELLI, L. et al. A 2.6 nW, 0.45 V Temperature-Compensated Subthreshold CMOS Voltage Reference. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.46, n.2, p.465–474, 2011.

MATTIA, O. E.; KLIMACH, H.; BAMPI, S. 0.9 V, 5 nW, 9 ppm/oC resistorless sub-bandgap voltage reference in 0.18um CMOS. In: CIRCUITS AND SYSTEMS (LASCAS), 2014 IEEE 5TH LATIN AMERICAN SYMPOSIUM ON, 2014. **Anais...** [S.l.: s.n.], 2014. p.1–4.

MATTIA, O. E.; KLIMACH, H.; BAMPI, S. 0.7 V Supply, 8 nW, 8 ppm/C Resistorless Sub-Bandgap Voltage Reference. In: CIRCUITS AND SYSTEMS (MWSCAS), 2014 IEEE 57TH MIDWEST SYMPOSIUM ON, 2014. **Anais...** [S.l.: s.n.], 2014. p.1–4.

MATTIA, O. E.; KLIMACH, H.; BAMPI, S. 2.3 ppm/C 40 nW MOSFET-Only Voltage Reference. In: INTERNATIONAL SYMPOSIUM ON LOW POWER ELECTRONICS DESIGN (ISLPED), 2014., 2014. **Anais...** [S.l.: s.n.], 2014.

MATTIA, O. E.; KLIMACH, H.; BAMPI, S. Sub-1 V Supply Nano-Watt MOSFET-Only Threshold Voltage Extractor Circuit. In: INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS (SBCCI 2014), 2014. **Anais...** [S.l.: s.n.], 2014. p.1–4.

MATTIA, O.; KLIMACH, H.; BAMPI, S. Resistorless BJT bias and curvature compensation circuit at 3.4 nW for CMOS bandgap voltage references. **Electronics Letters**, [S.l.], v.50, p.863–864(1), June 2014.

MEIJER, G.; SCHMALE, P. C.; VAN ZALINGE, K. A new curvature-corrected bandgap reference. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.17, n.6, p.1139–1143, 1982.

MING, X. et al. A High-Precision Compensated CMOS Bandgap Voltage Reference Without Resistors. **Circuits and Systems II: Express Briefs, IEEE Transactions on**, [S.l.], v.57, n.10, p.767–771, 2010.

OSAKI, Y. et al. 1.2-V Supply, 100-nW, 1.09-V Bandgap and 0.7-V Supply, 52.5-nW, 0.55-V Subbandgap Reference Circuits for Nanowatt CMOS LSIs. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.48, n.6, p.1530–1538, 2013.

PEASE, R. The design of band-gap reference circuits: trials and tribulations. In: BIPOLAR CIRCUITS AND TECHNOLOGY MEETING, 1990., PROCEEDINGS OF THE 1990, 1990. **Anais...** [S.l.: s.n.], 1990. p.214–218.

SEOK, M. et al. A Portable 2-Transistor Picowatt Temperature-Compensated Voltage Reference Operating at 0.5 V. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.47, n.10, p.2534–2545, 2012.

SONG, B.-S.; GRAY, P. A precision curvature-compensated CMOS bandgap reference. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.18, n.6, p.634–643, 1983.

TORRES, E. O.; MILNER, L. A.; RINCON-MORA, G. A. Hybrid supplies for wireless micro-systems. **The Electrochemical Society's Interface**, [S.l.], v.17, n.3, p.57–60, 2008.

TSIVIDIS, Y. Accurate analysis of temperature effects in I/SUB c/V/SUB BE/ characteristics with application to bandgap reference sources. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.15, n.6, p.1076–1084, 1980.

TSIVIDIS, Y.; ULMER, R. A CMOS voltage reference. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.13, n.6, p.774–778, 1978.

UENO, K. et al. A 300 nW, 15 ppm/C, 20 ppm/V CMOS Voltage Reference Circuit Consisting of Subthreshold MOSFETs. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.44, n.7, p.2047–2054, 2009.

VITTOZ, E.; FELLRATH, J. CMOS analog integrated circuits based on weak inversion operations. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.12, n.3, p.224–231, 1977.

VITTOZ, E.; NEYROUD, O. A low-voltage CMOS bandgap reference. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.14, n.3, p.573–579, 1979.

WIDLAR, R. New developments in IC voltage regulators. **Solid-State Circuits, IEEE Journal of**, [S.l.], v.6, n.1, p.2–7, 1971.

APPENDIX A LIST OF PUBLICATIONS

1. MATTIA, O. E., KLIMACH, H. and BAMPI, S. *"0.9 V, 5 nW, 9 ppm/°C Resistorless Sub-Bandgap Voltage Reference in 0.18μm CMOS"*. Proceedings of the 5th IEEE Latin American Symposium on Circuits and Systems **(LASCAS'14)**. Santiago, Chile. February 2014.
2. MATTIA, O. E., KLIMACH, H. and BAMPI, S. *"Resistorless BJT bias and curvature compensation circuit at 3.4 nW for CMOS bandgap voltage references"*. **Electronics Letters**. Vol. 50, Issue 12, p. 863-864. June 2014.
3. MATTIA, O. E., KLIMACH, H. and BAMPI, S. *"0.7 V Supply, 8 nW, 8 ppm/°C Resistorless Sub-Bandgap Voltage Reference"*. Proceedings of the 57th Midwest Symposium on Circuits and Systems **(MWSCAS'14)**. College Station, Texas, USA. August 2014.
4. MATTIA, O. E., KLIMACH, H. and BAMPI, S. *"2.3 ppm/°C 40 nW MOSFET-Only Voltage Reference"*. Proceedings of the International Symposium on Low Power Electronics and Design **(ISLPED'14)**. La Jolla, California, USA. August 2014.
5. MATTIA, O. E., KLIMACH, H. and BAMPI, S. *"Sub-1 V Supply Nano-Watt MOSFET-Only Threshold Voltage Extractor Circuit"*. Proceedings of the 27th Symposium on Integrated Circuits and Systems Design **(SBCCI'14)**. Aracaju, Brazil. September 2014.

   Still under review are the following publications.
6. MATTIA, O. E., SCHNEIDER, M., KLIMACH, H. and BAMPI, S. *"0.7 V Supply Self-Biased NanoWatt MOS-Only Threshold Voltage Monitor"*. Submitted to **ISCAS'14**.
7. MATTIA, O. E., KLIMACH, H. and BAMPI, S. *"Sub-1 V Supply Resistorless Sub-Bandgap Voltage Reference in 0.13μm CMOS"*. Submitted to the **Journal of Analog Integrated Circuits and Signal Processing** as invited paper due to the presentation of the LASCAS'14 paper.

APPENDIX B ACM MOSFET MODEL

In the ACM model, the drain current I_D of a long-channel MOSFET is expressed as

$$I_D = I_F - I_R = SI_{SQ}(i_f - i_r) \tag{B.1}$$

where I_F and I_R are the forward and reverse currents, $S = W/L$ is the aspect ratio, W being the width and L the length of the transistor. i_f and i_r are the forward and reverse inversion coefficients, related to the source and drain inversion charge densities, while I_{SQ} is the sheet normalization transistor current

$$I_{SQ} = \frac{1}{2} n\mu C'_{ox}\phi_t^2 \tag{B.2}$$

where n is the subthreshold slope factor, μ is the channel effective mobility (both slightly dependent on the gate voltage V_G), C'_{ox} is the gate capacitance per unit area, and ϕ_t is the thermal voltage. The relationship between inversion level i_f and i_r and terminal voltages is given by

$$\frac{V_P - V_{S(D)}}{\phi_t} = F(i_{f(r)}) = \sqrt{1 + i_{f(r)}} - 2 + \ln(\sqrt{1 + i_{f(r)}} - 1) \tag{B.3}$$

where V_S and V_D are the source and drain voltages (all terminal voltages are referenced to the transistor bulk), and V_P is the pinch-off voltage, approximated by

$$V_P \simeq \frac{V_G - V_{T0}}{n} \tag{B.4}$$

where V_G is the gate voltage, and V_{T0} is the threshold voltage for zero bulk bias. Over time, many operational and device physics dependent (inversion charge, for instance) definitions for V_{T0} were used. In the ACM MOSFET model, the threshold voltage has a universal physical meaning, defined as the condition where the drift and diffusion components of the drain current have equal magnitude.

The first term (the square root one) in the right side of (B.3) is related to the drift component of the drain current, being predominant under strong inversion. The last term (the logarithmic one) is related to the diffusion component, being predominant under weak inversion operation. In the forward saturation condition, $I_F \gg I_R$, and consequently, $I_D \simeq I_F = SI_{SQ}i_f$. In this thesis the V_{T0} value is then rigorously defined based on (B.3).

APPENDIX C SUMMARY IN PORTUGUESE

C.1 Introdução

Este capítulo contextualiza o uso de referências de tensão em sistemas eletrônicos, e apresenta as principais necessidades das aplicações contemporâneas, assim como as restrições que as mesmas impõem no projeto de circuitos. Define-se o que seria uma tensão de referência ideal e apresenta-se o conceito básico de compensação em temperatura, que é a métrica de desempenho mais importante, enquanto aponta-se para soluções utilizadas para melhorar outras métricas. Os principais objetivos da dissertação são então descritos, assim como a estrutura do trabalho.

C.1.1 Motivação

Referências de tensão são uma parte fundamental da maioria, se não de todos circuitos e sistemas eletrônicos. Elas são extensivamente utilizadas nas áreas de analógico, sinal misto, radio frequência e até em circuitos digitais como memórias, sendo que a exatidão necessária varia significativamente através dos domínios de aplicação. Para ilustrar a sua aplicação, a Figura 1.1 mostra um *System on Chip* (SoC) moderno da Analog Devices, que integra um microprocessador, um sistema de aquisição de dados de precisão, onde uma referência de tensão e mostrada explicitamente, e um transceptor de RF. Existem outras referências de tensão e/ou corrente - não mostradas na figura - distribuídas pelo sistema, que fornecem referências locais para comparadores ou polarização, por exemplo.

C.1.2 Aspectos de Baixa Tensão e Baixo Consumo

A evolução contínua das tecnologias CMOS é o principal fator por trás da operação a baixa tensão, rapidamente atingindo alimentações abaixo de 1 V para nós de processo abaixo de 130 nm. Os sistemas atuais, operador por bateria, ou os futuros auto-alimentados e auto-sustentáveis requerem muito baixo consumo de potência, em torno de alguns nA até alguns μA dependendo da função a ser executada, conforme ilustrado pelo perfil de carga de micro-sensores na Figura 1.2. Até mesmo SoCs grandes como o da Figura 1.1 possuem um consumo de potência em *stand-by* de 280 nA em modo *power-down* e de 1.9 μA com retenção de memória do processador e do transceptor. O bloco de referência de tensão normalmente é mantido ligado uma vez que ele fornece o limiar de chaveamento para os circuitos de *power on reset*, e uma referência exata para regulação da alimentação.

C.1.3 Aspectos do Processo de Fabricação

Estas limitações de alimentaçao impõem um limite na faixa dinâmica disponível para o processamento de sinais, e o constante aumento da complexidade dos sistemas requer que especificações desafiadoras sejam atingidas. De acordo com o relatório da *International Technology Roadmap for Semiconductors* (ITRS) de 2013, um dos desafios chave para integração é a manufatura de dispositivos passivos de qualidade dentro dos chips, que introduzem complexidades no processo e podem levar a preocupações de controle de manufatura e custo. Isso faz com que abordagens que não utilizem resistores sejam muito desejadas, mas diversas questões, principalmente de descasamento entre os dispositivos passivos e variações médias de processo devem ser abordadas.

A questão do descasamento local, devido à variações aleatórias na concentração de dopantes, por exemplo, pode ser resolvida através de boas práticas de projeto e layout. Já as variações médias de processo, que afetam *wafers* inteiros em diferentes rodadas, não podem ser resolvidas diretamente através de projeto, e afetam significativamente o desempenho de todos os circuitos, especialmente de referências de tensão. É por isso que as referências tradicionais são constituídas de constantes físicas, como a tensão térmica e a tensão de *bandgap* do silício. Ainda assim, não idealidades degradam o desempenho da referência, sendo que no caso de variações médias de processo as mesmas só podem ser compensadas através de duas alternativas custosas: esquemas complexos de compensação de dependências opostas com relação ao processo de fabricação, ou através de calibração.

Como já foi dito, o uso de referências de tensão é universal em sistemas eletrônicos. Exemplos de atividades recentes em áreas onde elas ainda necessitam melhorias são em sensores inteligentes, *energy harvesters*, dispositivos biomédicos implantáveis, sistemas analógicos e de RF assistidos digitalmente, em identificação por etiquetas inteligentes e dispositivos da Internet das Coisas.

C.1.4 Tensão de Referência Ideal

Idealmente uma tensão de referência é um bloco que fornece uma tensão constante, insensível à variações em:

- temperatura;
- tensão de alimentação;
- corrente de carga;
- deformação do encapsulamento;
- variabilidade no processo de fabricação.

Todos estes items são importantes na especificação de uma tensão de referência. A corrente de carga pode ser desconsiderada quando a referência e utilizada para polarizar a porta de um MOSFET, por exemplo, mas os outros parâmetros costumam ser preocupações universais. Nem todas as topologias de circuito tratam das questões mencionadas acima, mas todas utilizam uma estratégia comum para implementar uma referência independente da temperatura. O conceito básico é claramente ilustrado pela referência *bandgap* clássica (BGR) WIDLAR (1971), que é o principal foco desta dissertação, e mostrada na Figura 1.3. Outras referências baseadas em diodos Zener ou no ponto de *Zero Temperature Coefficient* (ZTC) dos MOSFETs são parecidas com a BGR, no sentido de explorarem a compensação de dois fenômenos com dependências opostas em temperatura, mas não são o foco deste trabalho.

Na BGR duas variáveis com sensibilidades opostas em função da temperatura são combinadas em uma soma ponderada que neutraliza o coeficient de temperatura da referência em uma dada faixa de temperaturas. Um termo linear e proporcional à temperatura absoluta (PTAT), baseado na tensão termica é amplificado por M e contra-balanceado pela tensão de uma junção de um BJT, que é complementar à temperatura absoluta (CTAT) e levemente não-linear. A referência e chamada de *bandgap* porque o valor da tensão de referência V_{OUT} é aproximadamente igual à tensão de *bandgap* do silício à 0 K. O circuito utilizado para gerar a corrente de polarização e o potencial térmico será detalhado mais a frente.

Este comportamento ideal é obviamente degradado por não-idealidades nos termos compositores. Por exemplo, a Figura 1.4 mostra o comportamento não-linear da tensão da junção, que leva a uma 'curvatura' bem conhecida na tensão de referência final. Existem esquemas que compensam tal curvatura, e eles serão detalhados no capítulo 3.

A curvatura da junção não é a única fonte de erro no circuito ideal da Figura 1.3, sendo outras a tensão de *offset* dos amplificadores, o descasamento entre os dispositivos, a resistência de base do transistor bipolar e o ganho de corrente finito. Estas não-idealidades foram tratadas extensivamente na literatura de referências *bandgap* PEASE (1990), assim como em sensores de temperatura AITA et al. (2013). Existem várias maneiras de melhorar a exatidão DC e AC desta topologia, onde algumas das soluções (e os respectivos compromissos) podem ser listados como GUPTA; RINCóN-MORA (2010):

- *trimming* (custo, tempo de teste);
- *Dynamic Element Matching* (complexidade, ruído de chaveamento);
- *buffer* em série (*offset* sensível à temperatura);
- fontes de corrente cascode (menor alimentação).

C.1.5 Objetivos

Reconhece-se a importância das técnicas apresentadas na seção anterior, e entende-se que circuitos de alto desempenho, sem restrições de custo ou consumo de potência, utilizem das mesmas para obter melhores resultados. No entanto, o principal objetivo desta dissertação é o de implementar novas topologias que atinjam muito baixo consumo de potência, com tensão de alimentação reduzida e que possam ser fabricadas em processos digitais CMOS sem resistores. Explora-se provas de conceito que possuem características diferentes das alternativas correntes, avaliando o seu potencial para atingir melhor desempenho através do uso destas técnicas complementares.

C.1.6 Organização

Esta dissertação está organizada como segue: no capítulo 2 uma revisão bibliográfica resumida revisa as principais referências de tensão em ordem cronológica, e uma síntese voltada à resultados e construída com base nos avanços mais recentes. No capítulo 3 as características do transistor bipolar de junção são exploradas, e seus principais aspectos e limitações sao apresentados. Também é introduzido um conceito de compensação de curvatura. O capítulo 4 introduz uma nova topologia de circuito que implementa tal conceito, assim como é derivada uma referência de tensão *sub-bandgap* a partir desta topologia. Resultados de simulação detalhados seguem no capítulo 5, acompanhados de resultados experimentais de uma rodada de fabricação. A dissertação conclui com um comparativo entre os resultados obtidos e o estado da arte, e aponta para direções futuras no desenvolvimento destas topologias.

C.2 Resumo do Texto

O segundo capítulo apresenta um modelo não-linear em função da temperatura para e tensão de bandgap do silício, e deriva-se uma expressão para tensão base-emissor. Essa tensão baseia-se em algumas aproximações com relação à mobilidade efetiva dos portadores na base, e assume-se uma dependência com temperatura para corrente de coletor do transistor. A expressão final resulta em termos não lineares relacionados à esta dependência e à tensão de bandgap, que podem ser então compensados através de uma corrente de coletor com $\delta = 3.5$. Faz-se uma análise de sensibilidade com relação às variáveis que compoem essa tensão base-emissor, e motra-se que o parâmetro de projeto mais importante é a tensão base-emissor na temperature de referência $V_{BE}(T_r)$.

O terceiro capítulo introduz um circuito de polarização de bipolares para tecnologia CMOS, e que implementa a técnica de compensação de curvatura descrita no capítulo 2. Uma metodologia de projeto é descrita, e um exemplo de implementação é mostrado em tecnologia XFAB 180 nm, onde compara-se os resultados preliminares com o modelo analítico, mostrando boa coerência entre ambos.

O quarto capítulo utiliza o circuito de polarização do capítulo 3 para implementar um circuito de referência sub-bandgap. A divisão da tensão base-emissor é apresentada, assim como as células *self-cascode* que geram o termo PTAT. Uma expressão para tensão de referência e derivada, e uma metodologia de projeto é apresentada com base nesta equação.

No capítulo 5 são apresentados resultados de simulação para o circuito de polarização e compensação de curvatura do BJT, utilizando dois processos diferentes, XFAB 180 nm e IBM 130 nm. Além disso, resultados experimentais de uma rodada de fabricação em IBM 130 nm comprovam o funcionamento da topologia para faixa de temperatura proposta. Na sequência, os resultados do circuito de referência sub-bandgap são apresentados, novamente para duas tecnologias XFAB 180 nm e IBM 130 nm. Da mesma forma, resultados experimentais em IBM 130 nm comprovam o funcionamento do circuito de referência.

C.3 Conclusão

Esta dissertação cobriu os avanços mais recentes em referências de tensão em tecnologia CMOS, especificamente aquelas que operam com alimentação abaixo de 1 V, consomem nanoWatts e são implementadas em processos sem resistores. As principais aplicações para estes tipos de circuitos se encontram em dispositivos operados a bateria e auto-sustentáveis, sensores inteligentes e etiquetas de identificação como RFID.

Detalhou-se a principal não-linearidade presente na referência *bandgap* tradicional, e explicou-se como a curvatura pode ser corrigida através de uma corrente de coletor fortemente dependente com a temperatura. Uma nova topologia de circuito foi introduzida para gerar tal corrente, e uma metodologia de projeto foi apresentada que implementa a compensação de curvatura independentemente da tensão de polarização escolhida. Resultados de simulação e experimentais utilizando *Low Power* FETs do processo IBM 130 nm mostram que o circuito opera com 0.7 V de alimentação, consumindo 3.1 nW (para uma tensão de emissor de 550 mV) à temperatura ambiente e ocupando uma área de silício de 0.00067 mm^2. A tensão de emissor varia até σ/μ = 4.1 % devido à variação na tensão de limiar dos MOSFETs. Resultados similares foram obtidos utilizando FETs de I/O no mesmo processo, e FETs regulares do processo XFAB 180 nm. Resultados experimen-

tais à temperatura ambiente para 10 amostras não encapsuladas de uma mesma rodada mostram uma tensão de emissor média de 610 mV, o que está de acordo com as variações de processo esperadas, e com σ/μ = 1.4 %. O mesmo circuito implementado com FETs de I/O foi caracterizado em função da temperatura para uma amostra não encapsulada, fornecendo uma derivada média da tensão da junção contra temperatura de - 1.56 mV/°C.

Derivado deste circuito de polarização foi um divisor da tensão de junção e, consequentemente, uma referência *sub-bandgap* utilizando a bem conhecida célula *self-cascode*. Demonstramos uma tensão de referência de 479 mV que consome 4.9 nW sob uma alimentação de 0.9 V, simulada em XFAB 180 nm. As principais vantagems são o baixo coeficiente de temperatura de 9 ppm/°C para uma faixa de 0 a 125 °C, com uma área de silício de 0.0012 mm^2. Simulações de Monte Carlo mostram que a variação esperada para a referência de tensão e de σ/μ = 2 % para uma análise combinada de variabilidade. O mesmo circuito foi projetado e fabricado em IBM 130 nm utilizando FETs de I/O, e resultados experimentais à temperatura ambiente para 10 amostras não encapsuladas apresentam uma referência média de 550 mV com σ/μ = 1.3 %. Medidas em temperatura para três amostras demonstram um coeficiente de temperatura mínimo de 11 ppm/°C e máximo de 86 ppm/°C, consumindo menos de 10 nW à temperatura ambiente e menos de 40 nW na temperatura máxima.

Comparações com o estado da arte sao difícieis devido à escassez de dados, especialmente relacionados à variabilidade de processo. Fizemos o melhor para apresentar uma comparação justa na Tabela 6.1, enquanto que resultados mais detalhados foram apresentados no começo da dissertação.

É justo dizer que nossa topologia tem o potencial de melhorar significativamente o estado da arte em métricas importantes como a exatidão absoluta da tensão de referência e o seu coeficiente de temperatura sob uma faixa larga de temperatura, consumindo uma das menores correntes e ocupando a menor area para referências *bandgap* por um fator de 10. Estas afirmações sao baseadas em detalhados resultados de simulação e resultados experimentais preliminares.

C.3.1 Trabalhos Futuros

Existem é claro várias questões a serem abordadas no desenvolvimento das topologias apresentadas, especialmente se elas objetivam melhorar significativamente o estado da arte em termos de variabilidade de processo. Várias outras topologias foram projetadas durante o período deste mestrado, mas ficaram de fora desta dissertaçao. As mesmas podem ser melhoradas também. Listamos abaixo alguns possíveis tópicos de trabalho futuro como:

1. Caracterização em temperatura de mais amostras;
2. Efetuar pelo menos mais uma rodada de fabricação para verificar o impacto de variações entre lotes diferentes;
3. Duas topologias de extração de V_{T0} foram desenvolvidas mas não incluídas nesta dissertação. O valor extraído pode ser utilizado para compensar a variabilidade no circuito de polarização do BJT apresentado. Uma destas topologias foi publicada em MATTIA; KLIMACH; BAMPI (2014e) e uma tensão de referência foi realizada a partir dela em MATTIA; KLIMACH; BAMPI (2014c). A outra topologia permanece não publicada até então;
4. Um estudo detalhado dos fatores que impactam a variabilidade de BJTs e MOSFETs, visando a redução das variações médias de processo, que permanecem como

o fator mais restritivo em termos de desempenho;

5. Estudar e desenvolver um sistema de calibração eficiente em termos de custo;
6. Prover alguma forma de regulação de carga, onde as referências possam alimentar até uma determinada quantidade de corrente;
7. Melhorar a sensibilidade com relação à alimetação através do uso de fontes de corrente *cascode* ou de um espelho de tensão e corrente simétrico LAM; KI (2010).