A DVS-CAPABLE ULTRA-LOW-POWER SUBTHRESHOLD CMOS TEMPERATURE SENSOR

by

Gregory Toombs

A Thesis

Presented to Lakehead University

in Partial Fulfilment of the Requirement for the Degree of

Master of Science

in

Electrical and Computer Engineering

Thunder Bay, Ontario, Canada

August 2009

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-71765-3
Our file Notre référence
ISBN: 978-0-494-71765-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Abstract

GREGORY TOOMBS. A DVS-Capable Ultra-Low-Power Subthreshold CMOS Temperature Sensor (Under the Supervision of Dr. Ali Manzak).

There are many contemporary contexts in which a small, low-power-consumption temperature sensor is very valuable. Power, area, speed and temperature range factors are important constraints in modern VLSI design. As transistor dimensions decrease, it is possible to lower the operating voltage of circuits, and dynamic voltage scaling (DVS) has been successfully implemented in several commercial applications to reduce power consumption. Power density is increasing, and the resultant temperature issues are being addressed by DVS, considered an efficient dynamic thermal management (DTM) technique. DVS/DTM automation techniques require thermal sensors that operate over a range of supply voltages. Therefore, temperature sensor designs such as this one are needed to address these engineering challenges.

In this thesis, a DVS-capable ultra-low-power subthreshold temperature sensor in 180 nm CMOS technology is proposed. The design is composed of a proportional-to-absolute-temperature (PTAT) current generator modified for insensitivity to power supply variation. The design is monolithic; the included reference current generator is a peaking source whose input is fed back from the output of the sensor. The design procedure includes empirical parameter extraction from BSIM simulations to yield a transistor model viable for design calculations, and adjustments to biasing and transistor dimensions to minimize power consumption and maintain adequate voltage supply independence and linearity. The design utilizes the exponential characteristics of subthreshold CMOS transistors to construct an output current that is a first-order Taylor approximation proportional to the thermal voltage. This is the first time such a design scheme is presented.

Biographical Summary

Gregory Toombs was born in Hamilton, Ontario on May 7, 1984. He studied Electronics Engineering Technology at RCC Institute of Technology in Concord, Ontario, receiving his technician diploma in 2004, technologist diploma in 2005, and Bachelor of Technology degree in late 2005. In 2006, he joined Research in Motion in Waterloo, Ontario as an RF Engineering Technician and worked for a year.

In 2007 Gregory moved to Thunder Bay, Ontario to pursue a Master of Science in Electrical and Computer Engineering degree at Lakehead University. His research area concerns the design and simulation of low-power analogue circuits. He is an active member of the Institute of Electrical and Electronics Engineers (IEEE), the Ontario Association of Certified Engineering Technicians and Technologists (OACETT), and the Ontario Society of Professional Engineers (OSPE).

Gregory has been granted conditional acceptance to a Ph.D. in Electrical and Computer Engineering at the University of Windsor to commence in the fall of 2009, where he will study under the supervision of Dr. Bubaker Boufama and the adjunct supervision of Dr. Rachid Benlamri.

Acknowledgements

My professors at Lakehead University have been instrumental in providing me a consistently practical engineering education and have succeeded in interesting me and my cohorts in what could otherwise be material that runs the risk of blandness.

I thank Dr. Ali Manzak, my graduate supervisor, for supporting my thesis and providing a tireless sense of humour; and Dr. Carlos Christoffersen, my co-supervisor, for his seemingly limitless expertise in analogue circuits and equally limitless amount of ideas to try when nothing works. The nature of this thesis requires a certain type of extreme tolerance for unpredictability, and both of my supervisors have demonstrated this tolerance in ample supply.

I also thank my fellow graduate students, for inspiring me with their successes and for being there to tell me when I have been slaving away in my room for one too many days.

I finally thank my mother and father for sending to the remote northern reaches of Ontario a stream of encouragement, pride, and sustenance.

Gregory Toombs

gtoombs@lakeheadu.ca

Contents

List of Figures....vi

List of Tables....viii

List of Symbols....ix

1 Introduction....1

1.1 Motivation and Objectives of This Study....1

1.2 Thesis Overview....2

2 Literature Review....4

2.1 PTAT Sensors....4

2.2 Summary of Reviewed PTAT Designs....14

2.3 Review of Supporting Circuits....15

3 Theoretical Operation....18

3.1 Circuit Premise....18

3.2 Circuit Equations....24

4 Design....28

4.1 Cascode Current Mirror....28

4.2 Peaking Current Reference....32

4.3 Startup Circuit....35

4.4 Complete Topology....39

4.5 Final Design Parameters....39

5 Simulation Results....41

5.1 Thermal Variation....41

5.2 Power Supply Variation....41

5.3 Monte Carlo Analysis....43

5.4 Reference Behaviour....45

5.5 Comparison with Literature Review....46

6 Conclusions and Future Work....48

6.1 Conclusions........48
6.2 Future Work........48
Appendix A. Technology Parameter Extraction........50
Appendix B. Standard deviation evaluation for the histogram of [5]........57
Appendix C. Initial Dimension Calculations........58
References........63

List of Figures

Figure 2.1: Differential BiCMOS temperature sensor [6]....5
Figure 2.2: Basic bandgap reference [7]....5
Figure 2.3: Differential CMOS temperature sensor [5]....6
Figure 2.4: A fully CMOS PTAT reference [9]....8
Figure 2.5: An all-CMOS PTAT reference with large current mirrors [10]....9
Figure 2.6: A differential all-CMOS PTAT circuit [2]....10
Figure 2.7: A PTAT/bandgap reference using mixed CMOS/BiCMOS transistors [11]....11
Figure 2.8: An all-CMOS subthreshold PTAT reference [12]....12
Figure 2.9: A 65nm CMOS temperature sensor design [4]....13
Figure 2.10: Comparison of thermal ranges for the reviewed designs....15
Figure 2.11: A simple low-voltage current mirror [14]....16
Figure 2.12: Simple peaking current source [15]....16
Figure 3.1: System block diagram....19
Figure 3.2: PMCC-delta transform....22
Figure 3.3: The biasing current generator....24
Figure 3.4: The biasing current subtractor, adder and mirrors....25
Figure 3.5: The translinear cell and output stage....26
Figure 4.1: Isolated low-voltage cascode current mirror....29
Figure 4.2: IOUT, ROUT and error versus variations in VLOAD....30
Figure 4.3: PSRR, IOUT and error versus variations in VDD for 1GΩ input sources....30
Figure 4.4: Isolated peaking current reference circuit....31
Figure 4.5: Iref versus Iout for Rp from 10kΩ to 1MΩ....33
Figure 4.6: Power consumption, PSRR and reference current as a function of supply voltage....33
Figure 4.7: Peaking reference current as a function of temperature....34
Figure 4.8: The startup circuit....35
Figure 4.9: Vdd, Iout, Vstart and Vra transients during startup....36
Figure 4.10: The main thermal biasing section....37
Figure 4.11: The ground-referred cascode bias branch and peaking source....37
Figure 4.12: Thermal current subtractor, adder and mirrors....38
Figure 4.13: The translinear cell and output block....38
Figure 4.14: An example parametric sweep of δ and PMCC against M1-M6 width....40
Figure 5.1: η as derived from the ratio of ID2 to ID1....42
Figure 5.2: ID1 and ID2....42
Figure 5.3: The PTAT output current....42

Figure 5.4: PSRR and IOUT as a function of VDD....43
Figure 5.5: PMCC as a function of VDD....44
Figure 5.6: Monte Carlo histogram of average IOUT for process and mismatch variation....44
Figure 5.7: Output current as a function of temperature and reference current....45
Figure A.1: The parameter extraction circuit....51
Figure A.2: Regression on temperature and corresponding error....53
Figure A.3: Regression on VGS and corresponding error....53
Figure A.4: Regression on VDS and corresponding error....53
Figure A.5: Output resistance (NMOS and PMOS) versus VDS for length from 180nm to 2μm. 55

List of Tables

Table 2.1: Comparison of metrics from the reviewed PTAT designs....14
Table 4.1: Cascode mirror test parameter values....29
Table 4.2: Peaking source parameters....32
Table 4.3: Final transistor dimensions used....39
Table 5.1: Temperature range and corresponding linearity metrics....43
Table 5.2: Comparison of metrics from the reviewed PTAT designs with the current design....46
Table A.1: Assumed and extracted Matlab regression values....52

List of Symbols

Symbol	Units	Value	Description
$\langle x \rangle$, μ			Average of x
δ			Output nonlinearity
σ			Monte Carlo standard deviation
η		~1.3	Subthreshold model constant of proportionality
λ	V^{-1}	~0.1	Output resistance parameter
R^2			Coefficient of determination linearity metric
PMCC			Pearson product-moment correlation coefficient linearity metric
k	J/°K	1.381×10^{-23}	Boltzmann's constant
I_{ON}	A		Subthreshold model base current parameter
I_{OUT}	A		Main output current
I_{REF}	A		Reference current
PSRR	dB		Power supply rejection ratio
q_e	C	1.602×10^{-19}	Electron charge
T	°K		Absolute temperature
V_{DD}	V		Supply voltage
V_T	V		Thermal voltage
V_{TH}	V		Subthreshold model threshold potential

1 Introduction

1.1 Motivation and Objectives of This Study

Temperature sensors are an important aspect of many different contemporary systems, and are used for diverse reasons. The temperature sensor output data may be used directly; that is, recorded or broadcast in a climatological instrumentation setting or other environmental application. This data is useful not just as a primary system output but as an internal line fed to a control system within a CPU. Such systems commonly dynamically scale voltage or frequency or perform simple CPU usage throttling to regulate system temperature from values high enough to damage the chip, or more straightforwardly issue a warning or shut the system down when internal temperature exceeds such values.

A viable application for this sensor is as the constitutive element in an array of temperature sensors on a large silicon die prone to excessive heat generation. This assumes that the sensors are spaced apart enough to be adequately thermally isolated. In development it may be very beneficial to include such an array to be able to inexpensively monitor a die for the magnitude and position of localized hot spots [1],[2]. The low power consumption of the sensor sections would minimize the thermal contribution from the sensors themselves. This approach would potentially reduce or eliminate the need for expensive and complex infrared thermography imaging techniques. Such a technique could conceivably be extended to module-by-module DVS [3].

Another application is cell phones and other small, wireless transmission devices. Such

devices are usually powered by battery and have very stringent restrictions on area and current capacity of CMOS sections. The power amplifier and, to a lesser extent, the processor of a cell phone both generate an amount of heat significant enough to warrant thermal monitoring in developmental and testing stages, if not also in actual deployment [4].

Environmental and meteorological research equipment is often limited to poor power sources such as batteries and solar power, and low-power temperature sensors are needed in this context as well.

The design goals of this study are to implement a temperature sensor capable of maintaining stable output despite variation in power supply voltage and, to a lesser extent, CMOS process parameters; to quantify linearity and to maintain that linearity at an acceptable level; to minimize power consumption; and to achieve a design that meets or exceeds these metrics as compared to other common sensor designs.

1.2 Thesis Overview

Chapter 2 assesses the state of the art regarding CMOS and BJT temperature sensors, summarizing the papers describing designs relevant to this work. It also assesses several designs for current mirrors and current references, both needed in this circuit.

Chapter 3 contains Section 3.1 that describes the basic characteristics and objectives of the circuit, as well as describing the quantitative metrics used to assess the quality and performance of the circuit. It also contains Section 3.2 that describes the specific equations that are used to derive the PTAT behaviour of the circuit. The main circuit schematic figures are shown and explained in this section.

Chapter 4 details the design of the current mirrors, the current reference, and the startup circuit; the nature of the tuning needed to achieve acceptable circuit performance; and the final transistor dimensions and bias resistor values chosen.

Chapter 5 illustrates the behaviour of the circuit over varying temperature, power supply,

and technology parameters, and lists the performance metrics of the circuit.

Chapter 6 assesses the circuit's strengths and weaknesses both in general and with respect to other sensors in the field, and discusses opportunities for future related work.

Appendix A. describes the work done to formulate a relatively simple mathematical model used to describe the behaviour of CMOS transistors in the subthreshold region for this particular technology (180 nm). It includes details on how Matlab is used to perform regression analysis to extract technology parameters used in the model, as well as the Matlab code used to perform this analysis.

Appendix B. shows the formulae and spreadsheet used to calculate the standard deviation for the output of the circuit in [5].

Appendix C. shows the calculations performed to acquire starting values for transistor dimensions. These values are not used for the final design; they are first optimized with selected simulations.

2 Literature Review

2.1 PTAT Sensors

It is important to review the current state of the art concerning temperature sensors to be able to effectively compare the design of the sensor in this thesis to those that are being developed in industry. The designs are presented in chronological order.

Altet, Rubio, Dilhaire et al. present a BiCMOS thermal sensor circuit for built-in test purposes [6]. It is intended to measure local temperature increases at the surface of the silicon die and to detect abnormal temperature gradients, in the attempt to detect a certain type of circuit defect that creates this gradient. The gradient is detected by comparing differential outputs from two different temperature sensing devices (TSDs) separated by a specific distance; in this case, 500μm.

The circuit is fabricated using a 1.2μm BiCMOS process. The power supply voltage used is 5V. It is specifically optimized for very high sensitivity; under certain conditions the circuit is able to achieve 2.7V/°C. Power consumption, linearity and PSRR are not discussed. The circuit schematic for the differential sensor is shown in Figure 2.1.

Meijer, Wang and Fruett discuss several PTAT concepts and circuits [7]. Among them is a simple bandgap reference with a PTAT ΔV_{BE} node, shown in Figure 2.2. They describe the main drawback of this circuit to be the non-ideality of the applied amplifier and gain inaccuracy. A test circuit using the bipolar bandgap concept was implemented in a 500nm CMOS process. Linearity and PSRR are not discussed. Although power consumption is not discussed directly, the

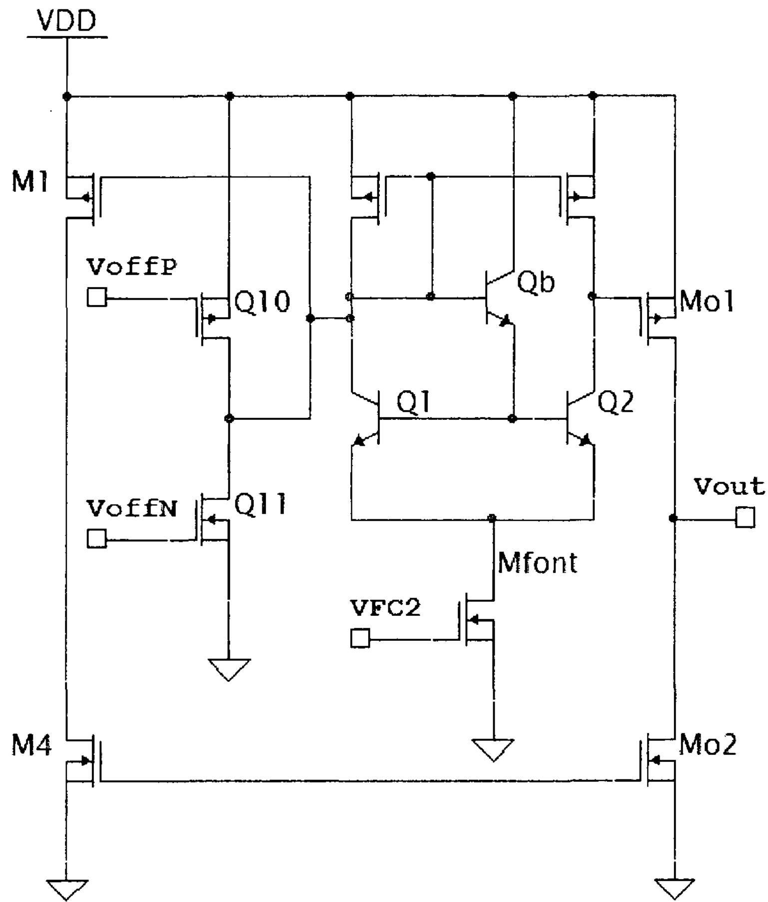


Figure 2.1: Differential BiCMOS temperature sensor [6].

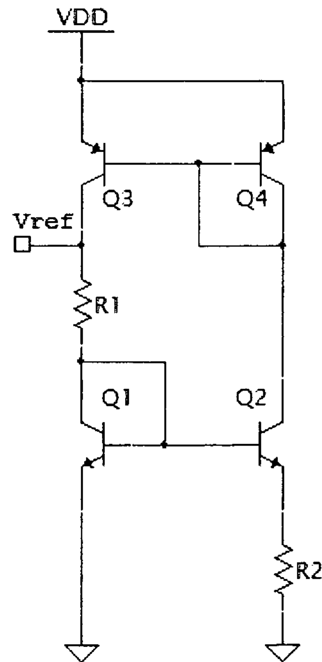


Figure 2.2: Basic bandgap reference [7].

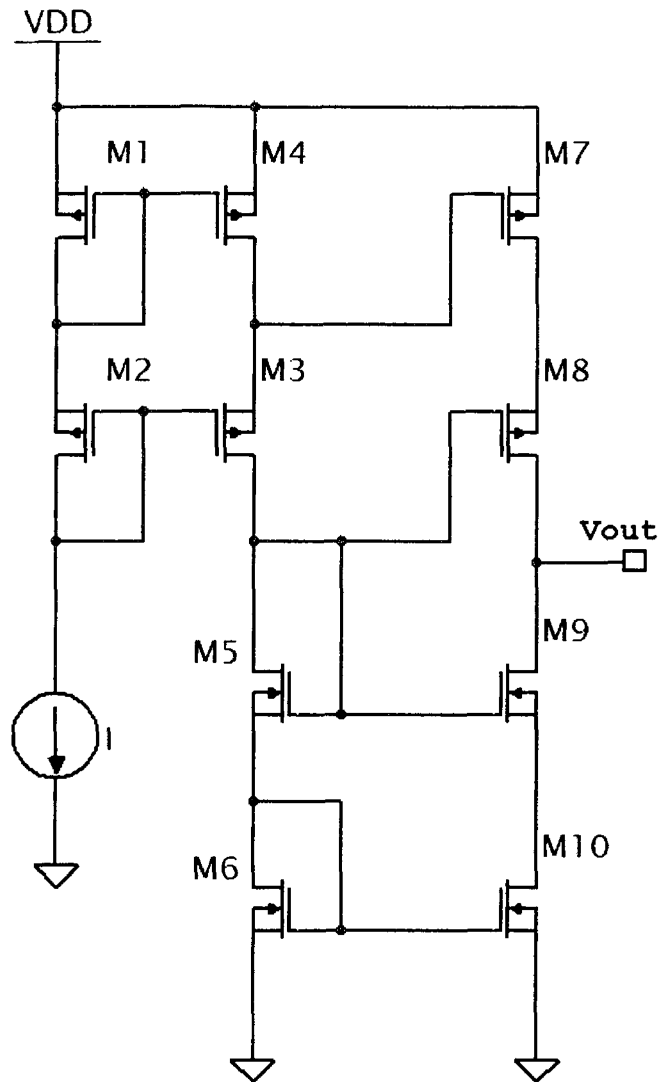


Figure 2.3: Differential CMOS temperature sensor [5].

circuit experiences collector currents greater than 10µA. The tested temperature range is 250°K – 430°K. A more complex reference using dynamic element matching is discussed, but is not fabricated, and no power supply rejection or power consumption performance metrics are included.

Li, Lai, Sin et al. introduce a thermal sensor specifically designed for high-temperature applications of up to 420°C [8]. It is implemented with a custom silicon-on-insulator (SOI) process. The layout has an SOI substrate, and two n+ regions with Ti-W electrodes configured to act as a thermal resistor. The fabricated resistor is 100µm wide by 12 µm long. Experimental results include a bias current of at least 1µA, and a temperature-resistance transfer function that

is roughly linear between 25°C and 350°C. The maximal resistance shown is approximately 16kΩ for the thin-film design; thus, the worst-case power consumption could be assumed to be 16nW before accounting for biasing circuitry costs. Such circuitry would have to be wholly responsible for managing power supply voltage variation insensitivity and signal amplification.

Syal, Lee, Ivanov et al. propose a differential CMOS temperature sensor [5]. The schematic is shown in Figure 2.3. Much like [6], it is intended to detect abnormal temperature gradients induced by circuit defects, but is implemented purely in CMOS instead of the traditional bipolar or BiCMOS technologies that many temperature sensor designs use. Transistor M4 acts as the sensing transistor; all others are for cascoded biasing. The input current is assumed to be ideal. The circuit is not fabricated, but CMOS layout is done in Cadence Virtuoso, and the circuit is simulated in Spice. The circuit is implemented in 180nm CMOS technology and occupies 1600 μm^2. The supply voltage used is 3.3V. The sensitivity is 1.2V/°C. The linearity is 1.3% between 0°C and 75°C, though the metric used to calculate linearity is not described. The maximal power is 50µW for a temperature of 85°C.

The paper also describes a second stage that uses a counter and a serial interface which create an output wave whose frequency is PTAT. The output has a sensitivity of 26.2kHz/°C. For a temperature between 0°C and 80°C, the output frequency varies between approximately 6MHz and 8MHz. The quoted supply sensitivity is 52kHz/V. To calculate a PSRR comparable with that used in this thesis, a form of equation (3.3) must be used:

$$20\log\left(\frac{7\,\text{MHz}}{3.3\,\text{V}\times 52\,\text{kHz/V}}\right)=32.2\,\text{dB} \tag{2.1}$$

The quoted standard deviation from Monte Carlo process variation is 30 kHz, or 0.43% of the nominal output frequency. However, the process variation histogram in the paper suggests a standard deviation closer to 2.45°C, or 6.14% of the nominal 39.87°C (see Appendix B.).

Serra-Graells and Huertas propose an all-CMOS PTAT sensor [9]. The circuit is optimized for low supply voltage and low power consumption. The circuit schematic is illustrated in Figure 2.4. The circuit operates on the premise of log companding. The V_{REF} taken from the source of

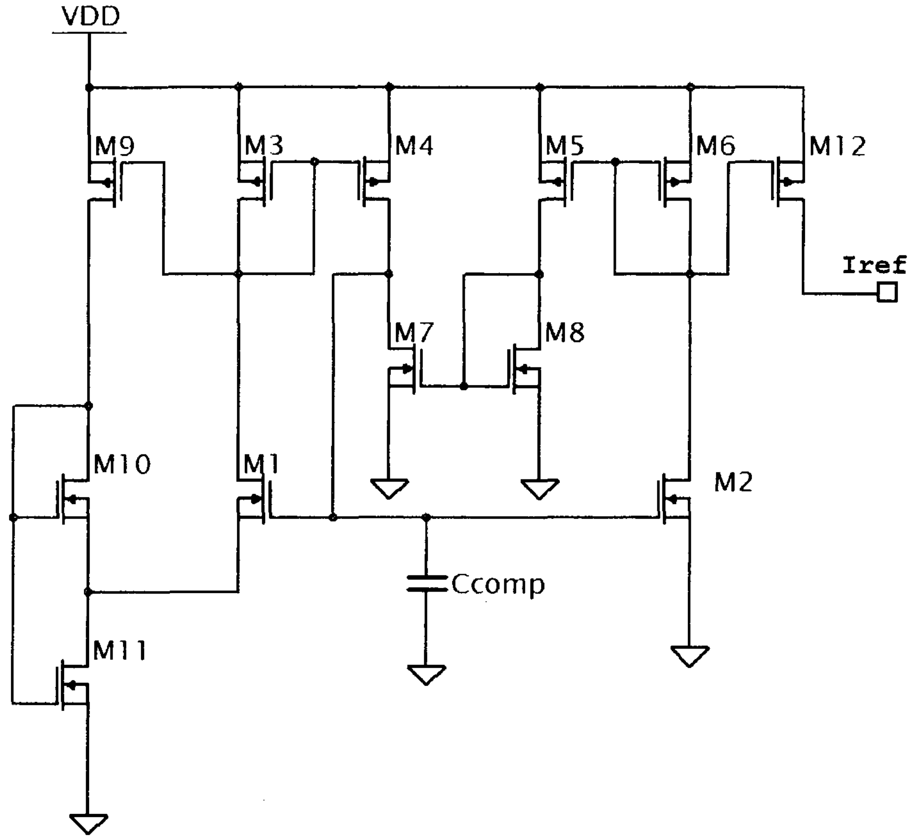


Figure 2.4: A fully CMOS PTAT reference [9].

M1 is ideally proportional to the product of the thermal voltage and the natural logarithm of the reciprocal of the gain through a specially designed attenuation loop, thus:

$$V_{REF} = U_T \ln P \quad (2.2)$$

The design is fabricated in both 1.2μm and 350nm CMOS technologies. For the 1.2μm design, the minimum supply voltage is 900mV, the maximum power consumption is 5μW, the sensitivity is 220μV/°C for V_{REF} and 1.13 nA/°C for I_{REF}, and the PSRR is greater than 60dB in the operational range of V_{DD} > 950mV. Linearity is not quantified. Although no temperature range is explicitly stated, if one is inferred from the transfer function graph in the paper, the output is valid for the range of 0°C to 50°C at a minimum. The standard deviation of V_{REF} over Monte Carlo variations is less than ~3mV. With a nominal V_{REF} of 65mV, this implies a standard deviation of less than ~5%.

Aside from the excellent supply rejection ratio, power consumption and minimal supply

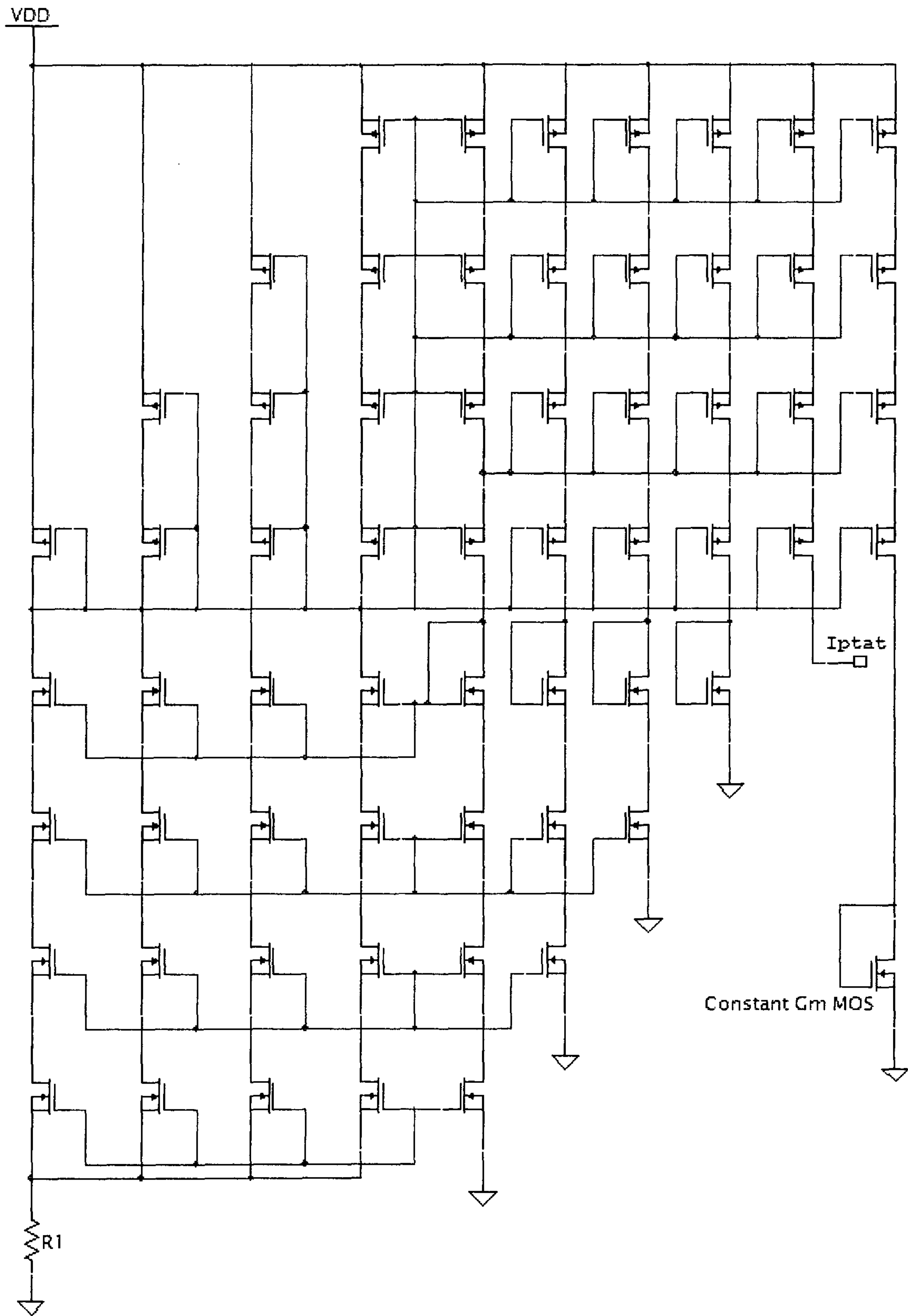


Figure 2.5: An all-CMOS PTAT reference with large current mirrors [10].

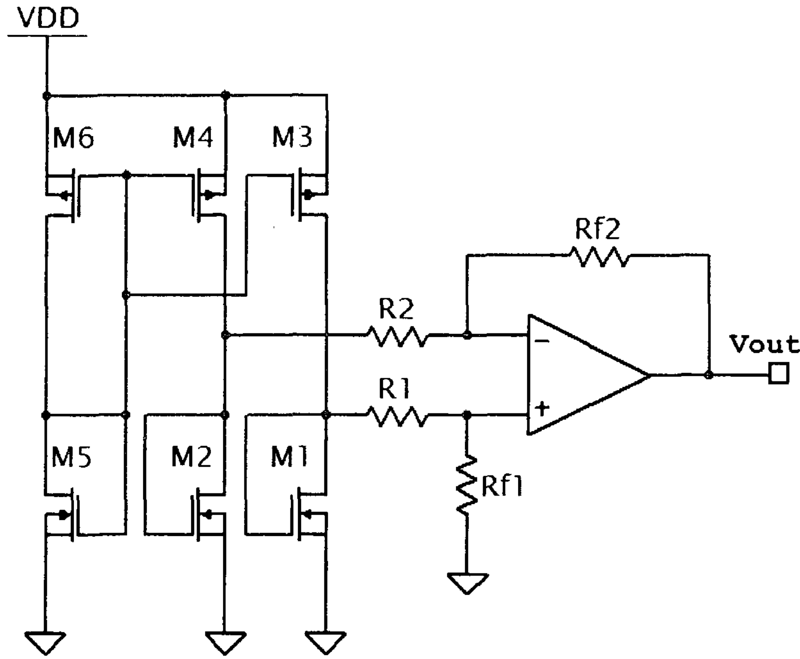


Figure 2.6: A differential all-CMOS PTAT circuit [2].

voltage, the design has the advantage of being monolithic; that is, it requires no external reference to operate.

Danaie and Lotfi present an all-CMOS PTAT current reference [10]. The circuit schematic is shown in Figure 2.5. It uses a very large high-swing quadruple current mirror and a large biasing section. The main objective of this circuit is to greatly increase PSRR. The circuit is not fabricated, but is simulated with the use of a BSIM3 model. The experimental results quote a temperature range of -50°C to 130°C, a supply of 1.3V, a power consumption of 80μW, and a PSRR consistently over 100dB. No sensitivity or linearity metrics are provided, but it is stated that the PSRR remains above 100dB in all process corners.

Chen, Meterelliyoz and Roy propose an all-CMOS PTAT sensor circuit for use in a temperature-adaptive voltage regulation scheme [2]. The circuit schematic is shown in Figure 2.6. The design generates two reference voltages that differ due to specifically chosen transistor dimensions, and subtracts them with a simple op-amp configured to act as a differential amplifier. From the temperature-output transfer function shown in the paper it can be seen that the circuit produces a very linear output for a wide range of temperatures, with an acceptable sensitivity. However, no power consumption or PSRR data are provided. The paper describes a

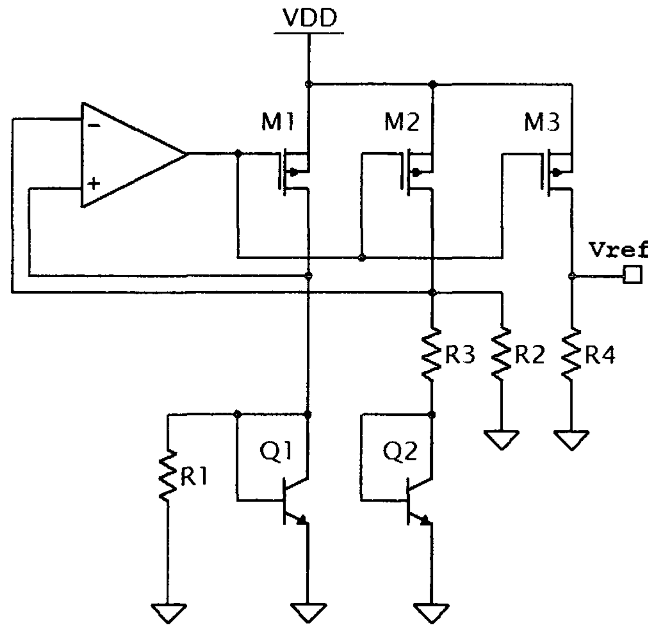


Figure 2.7: A PTAT/bandgap reference using mixed CMOS/BiCMOS transistors [11].

tendency for rather large inter-die process variation (±50mV at a nominal voltage of ~500mV, so ~±10%).

Lee, Hsu and Luo propose a simple PTAT/bandgap reference circuit using two BiCMOS transistors Q1 and Q2 for their thermal properties [11]. The schematic is shown in Figure 2.7. It relies on an operational amplifier and operates in a similar fashion to [2], but there are biasing resistors added to control circuit biasing currents. The design includes a calibration scheme to minimize error. The design is fabricated in a 250nm TSMC CMOS process. The temperature range considered is 20°C to 60°C. The nominal supply used is 2.5V, and can range from 1.8V-3.3V. The power consumption ranges between 300μW and 2.2mW. The design occupies less than $1mm^2$. No PSRR or linearity metric data are provided.

Ueno, Hirose, Asai, et al. propose an ultra-low-power all-CMOS PTAT sensor that does not require the use of any bipolar or BiCMOS transistors, and does not require a custom process [12]. The circuit schematic is shown in Figure 2.8. The circuit relies on the exponential characteristics of CMOS transistors in the subthreshold region. It generates two currents I_{D1} and I_{D2} that differ due to a biasing resistor. These two currents are then passed through an adder, subtractor and a translinear cell to manually construct an output current with a linear dependence on temperature

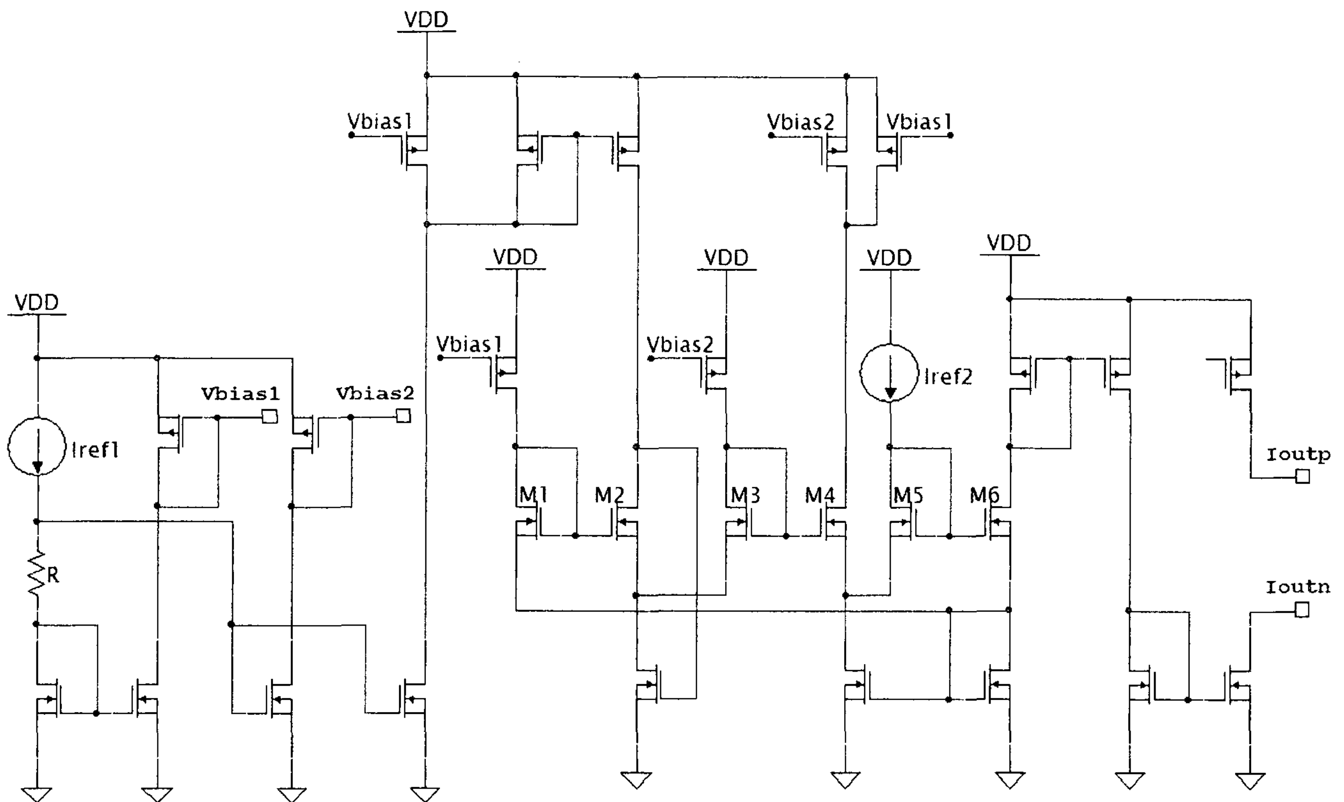


Figure 2.8: An all-CMOS subthreshold PTAT reference [12].

and (ideally) minimized dependence on technology parameters. The current thesis employs many of the concepts of this circuit; for a detailed discussion refer to Section 3.2.

This circuit also includes a charge-balancing ADC. The sensor design is implemented in a 350nm CMOS technology, and simulated with SPICE but not fabricated. The nominal power supply used is 1.5V. The sensitivity is 113 pA/°C from the bare PTAT section, or 21.4 Hz/°C from the ADC. The operational temperature range is -20°C - 100°C. The worst-case power consumption is 5.8μW at 100°C. No quantified linearity or PSRR data are provided.

Duarte, Geannopoulos, Mughal et al. describe a temperature sensor implemented in a relatively modern 65nm CMOS process [4]. The design is instantiated three times in the Intel Pentium 4 Processor. The circuit is intended to be part of a complete thermal management control system. The circuit is specifically optimized for high linearity over all process corners. The design includes a thermal sensor block, a DAC, a comparator and a thermal management control unit. Figure 2.9 shows a much-simplified schematic.

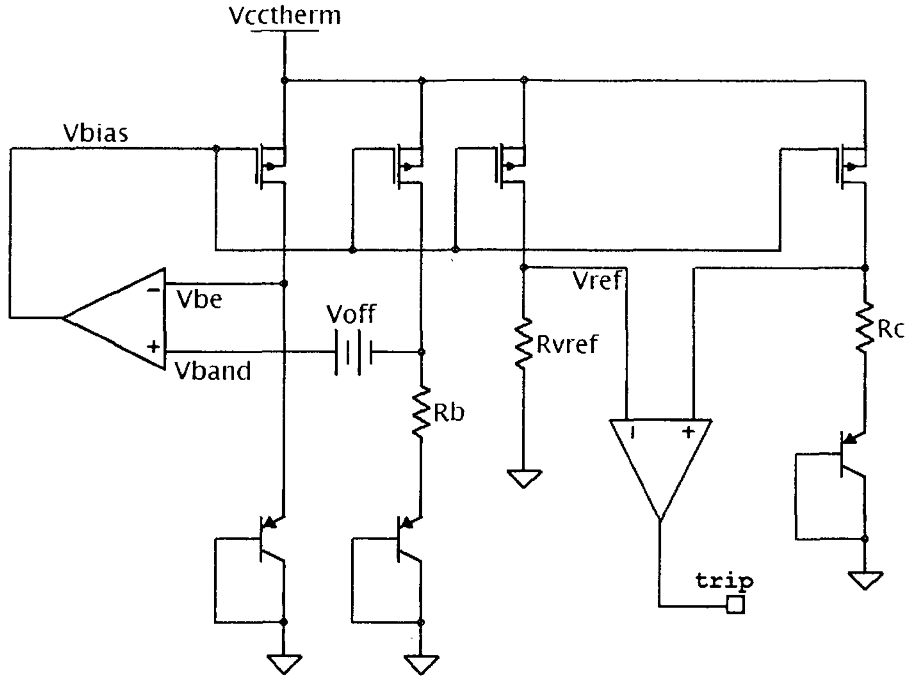


Figure 2.9: A 65nm CMOS temperature sensor design [4].

No power consumption or PSRR data are provided. However, the linearity is well-quantified. The paper employs the Pearson product-moment correlation coefficient (PMCC) metric, found by

$$r=\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{s_X}\right)\left(\frac{Y_i-\overline{Y}}{s_Y}\right) \tag{2.3}$$

where Y is a function of X, $\overline{X}$ and $\overline{Y}$ are the sample means, and s_X and s_Y are the sample standard deviations. A PMCC value of 1 indicates a perfectly linear, positively correlated relationship between X and Y. Thus, the PMCC values of 0.9982, 0.9980 and 0.9996 for the three respective instances of the sensor in the Pentium 4 processor indicate excellent linearity.

Tsai and Chiueh have manufactured a sensor of the same topology as that of [9], but have refined the performance and have included more metric information [13]. Their design is fabricated in TSMC 130nm and 180nm CMOS processes. For the 180nm design, the range for which there is a linear output is -55°C to 170°C. The power consumption is 28.53 µW. The sensitivity is 264 µV/°C. The design occupies 1260 µm^2.

This paper has well-quantified linearity metrics. It employs the "R^2" index, more commonly known as the coefficient of determination. It can be found by

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \langle y \rangle)^2} \tag{2.4}$$

where y represents the actual data, f represents a best-fit linear function, and $\langle y \rangle$ is the data mean. As with the PMCC, a value of 1 indicates a perfectly linear function, and thus the listed R^2 value of 0.99788 indicates excellent linearity.

2.2 Summary of Reviewed PTAT Designs

Src.	V_{DD}	Power	PSRR	T range	Sensitivity	Linearity	Area	Technology	Monte Carlo deviations
[2]				30°C - 150°C	~1.8mV/°C	Good		50nm CMOS predictive model	~10%
[4]				10°C - 90°C	0.95 – 1.15 °C/bit	Excellent		65nm CMOS	
[6]	5V				2.7V/°C			1.2μm BiCMOS	
[7]				-23°C - 157°C				500nm CMOS	
[8]		> 16nW		25°C - 350°C	~ 37Ω/°C	Acceptable	>1200μm²	Custom SOI	
[5]	3.3V	< 50μW	32.2dB	0°C - 75°C	1.2V/°C	Good	1600μm²	180nm CMOS	0.43% or 6.14%
[9]	>950 mV	< 5μW	>60dB	0°C - 50°C	220μV/°C, 1.13 nA/°C	Good	0.05mm²	1.2μm and 350nm CMOS	<~5%
[10]	1.3V	80μW	>100dB	-50°C - 130°C				180nm CMOS	0.5%
[11]	2.5V (1.8V-3.3V)	300μW-2.2mW		20°C - 60°C			< 1mm²	250nm CMOS	Calibrated
[12]	1.5V	<5.8μW		-20°C - 100°C	113pA/°C, 21.4Hz/°C	Good		350nm CMOS	
[13]		28.53μW		-55°C - 170°C	264μV/°C	Excellent	1260μm²	130nm and 180nm CMOS	

Table 2.1: Comparison of metrics from the reviewed PTAT designs.

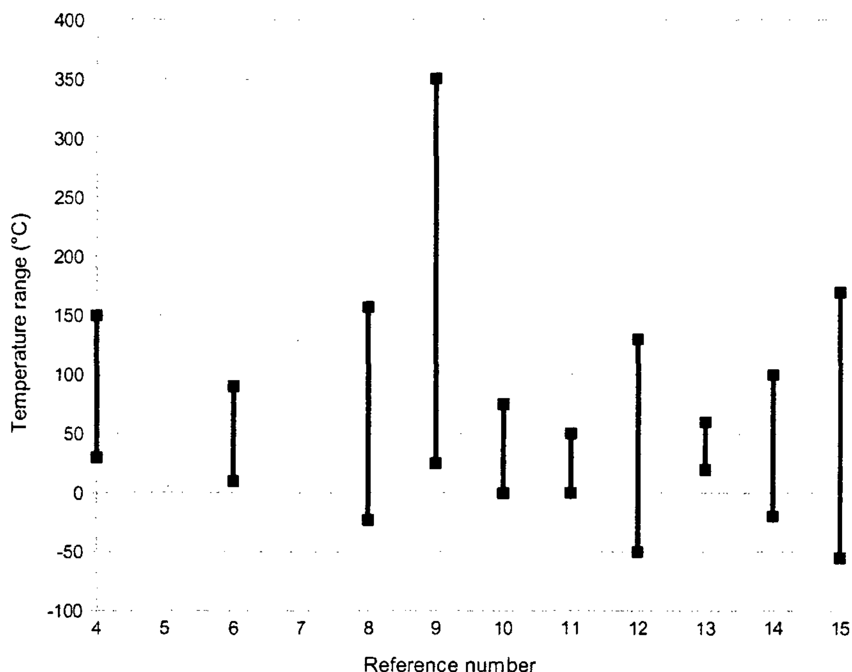


Figure 2.10: Comparison of thermal ranges for the reviewed designs.

2.3 Review of Supporting Circuits

In addition to a review of PTAT sensor circuit literature, it is important to review supporting circuits needed for the construction of the main circuit: current mirrors, and current references.

Yan and Sanchez-Sinencio have published an excellent general-purpose evaluation of several low-voltage current mirror designs [14]. They present the topology of Figure 2.11 as a conventional high-swing, low-voltage current mirror. They describe it as a regulated cascode structure capable of rendering an accurate transfer ratio depending on the matching between the left and right halves of the circuit. This circuit is capable of a much better output resistance than that of a simple current mirror with a single transistor for input and output, respectively. Thus, it is well-suited to applications such as DVS that require a higher PSRR than circuits that can assume a constant supply voltage. Of the fourteen mirrors described in the Low Voltage Current Mirror section of [14], it is the simplest high-swing design that does not require the use of operational amplifiers or complex biasing schemes.

Most analogue circuits, including this thesis, rely on the presence of a stable current reference source with acceptable performance in the areas of supply independence, voltage

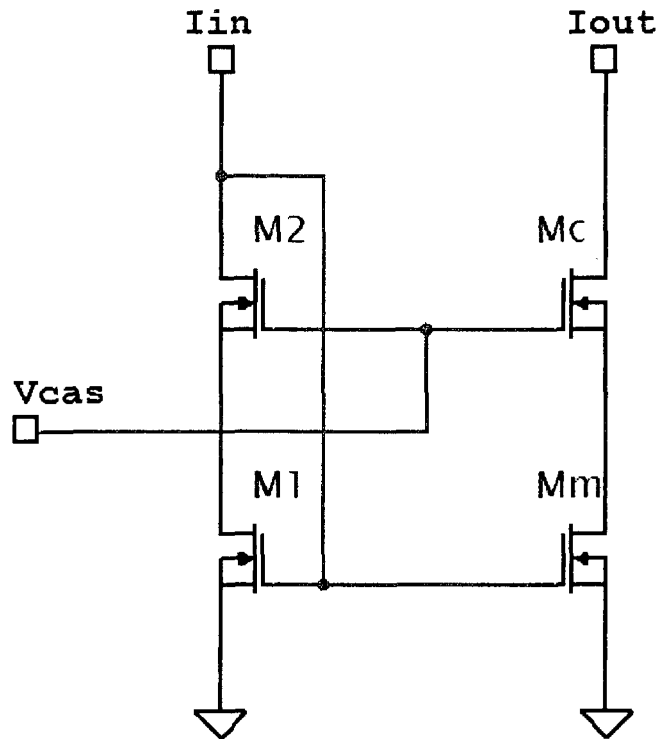


Figure 2.11: A simple low-voltage current mirror [14].

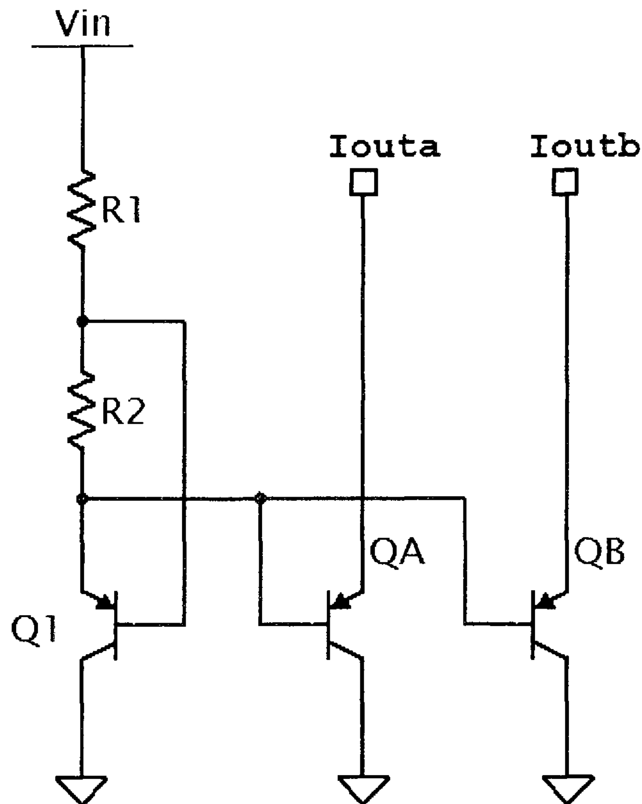


Figure 2.12: Simple peaking current source [15].

requirement, and output resistance. Peaking current sources are current sources with both an input current and output current, where the transfer function between the two exhibits a significant local maximum at a certain value. The design is popular and is well-described in the literature [15],[16],[17],[18]. Figure 2.12 shows the basic circuit, whose topology need not change between CMOS and BJT implementations. R1 may be omitted. The peaking behaviour is due to the effect of the voltage drop over R_2 on Q_1. Before the peak, as the input increases, Q_1 gradually turns on and the current mirrored to Q_A/Q_B accordingly increases. After the peak, as the effect of increasing voltage over R_2 on the gate of Q_A/Q_B becomes non-negligible, the current mirrored in Q_A/Q_B begins to decrease.

Presupposing the presence of a linearly varying output current at the last stage of the main circuit of this thesis, it can be justified that there exists a strong case for the use of a peaking current source whose input is mirrored from the output of the main circuit. The system will then constitute a feedback loop where there exists a point of local stability at the peak. The application of this circuit and its detailed operation is further discussed in Chapter 4.

3 Theoretical Operation

3.1 Circuit Premise

This thesis is comprised of a circuit design that is a hybrid of several existing circuits. The circuit operates in the subthreshold region and utilizes concepts from [12] to derive an output current that is a linear function of temperature. This derivation relies on the exponential characteristics of CMOS transistors in the subthreshold region, a CMOS-fabricated (on-die) resistor to generate two different biasing currents that are each a function of temperature, and a translinear cell to implement a mathematical function that generates an output current that is a linear function of absolute temperature to a first order of approximation in a Taylor series. The block diagram for the entire system is shown in Figure 3.1. Note that the feedback reference scheme shown in the block diagram is a modification not present in [12]; this feedback will be explained in Section 4.2.

One of the many challenges of CMOS circuit implementation in the subthreshold region is output insensitivity to changes in the voltage supply, characterized by the metric of power supply rejection ratio. The subthreshold biasing current and translinear scheme implemented with simple current mirrors produces a power supply rejection ratio insufficient for practical use in a dynamic voltage scaling context, and thus several significant circuit modifications must be performed. The most important of these is near-ubiquitous use of a standard dual-input cascoded current mirror design. Such current mirrors greatly increase power supply insensitivity at the cost of increased circuit area and slightly increased power consumption.

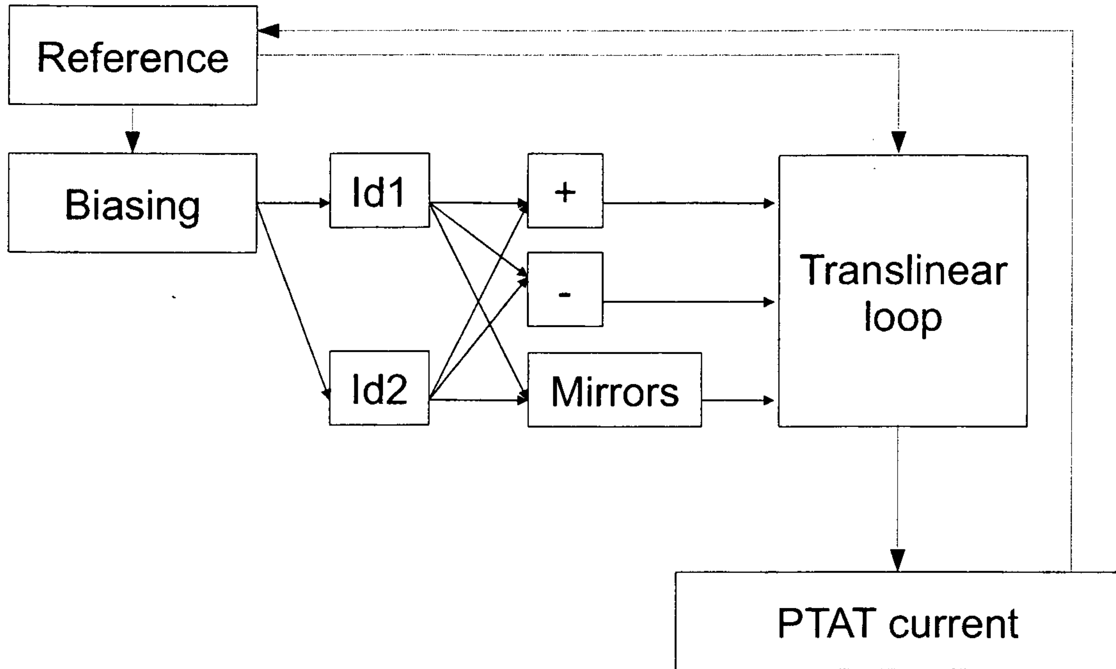


Figure 3.1: System block diagram.

Another significant challenge in the design and implementation of this circuit is a stable reference current. It is sometimes assumed by papers such as [12] that an ideal current reference is provided beforehand. In this case an ideal current source implies complete insensitivity to changes in temperature, as in (3.1).

$$\frac{\partial I_{ref}}{\partial T} \rightarrow 0 \tag{3.1}$$

In this context an ideal current source also implies complete insensitivity to changes in voltage supply. Traditionally, PSRR is calculated with (3.2).

$$PSRR = 20\log\left(\frac{\partial V_{DD}}{\partial V_{OUT}}\right) \tag{3.2}$$

Of course, current sources imply I_{OUT} instead of V_{OUT}, and so this thesis uses (3.3),

$$PSRR = 20\log\left(\frac{\langle I_{OUT}\rangle}{\langle V_{OUT}\rangle}\frac{\partial V_{DD}}{\partial I_{OUT}}\right) \tag{3.3}$$

where $\langle x \rangle$ signifies the average value of x. In this case, the averages would be taken over the full domain of the independent variable V_{OUT}. Thus, an ideal insensitivity to power supply variation implies

$$20 \log \left(\frac{\langle I_{REF} \rangle}{\langle V_{OUT} \rangle} \frac{\partial V_{DD}}{\partial I_{REF}} \right) \rightarrow \infty \tag{3.4}$$

In practise, the only time this can occur is at values for which functions experience local maxima or minima, but such an asymptote in the PSRR is less meaningful than its steady-state value. High PSRR is difficult to achieve for low-voltage CMOS circuits in the subthreshold region. Although circuit designs do exist for high-PSRR current references with acceptable temperature independence, it is shown in this thesis that the complexity of such a design is unnecessary and that a standard peaking source may function instead. It is shown that the peaking source is not temperature-independent, but that the thermal variation effects from the reference do not negatively affect circuit performance, and that the reference is acceptably stable during power supply variation.

The circuit is implemented in a standard 180 nm CMOS technology process, and simulated with the Cadence® Spectre®/Virtuoso® suite. CMC has kindly provided Lakehead University the tools and technologies required to support this software. Cadence uses a version of the BSIM (Berkeley Short-channel IGFET Model) simulation model. The complexity of this model makes it such that direct use of its equations would render basic design calculations hopelessly and impractically difficult. Such complexity calls for simplification for the purposes of estimate calculations; to this end, a simplified subthreshold behaviour model has been derived and is described in detail in Appendix A. The model is derived by constructing simple circuits in Cadence, selecting biasing currents in the general range suitable for the design of this sensor circuit, transferring a table of simulated values into Matlab, and running a multidimensional regression to fit the chosen model equations. It is shown that the correspondence between the behaviour simulated by BSIM and the simplified model functions is quite close.

After the circuit topology is laid out, parametric simulations must be performed in Cadence

to fine-tune the operation of the circuit. During this tuning process, the modified (independent) variables are the dimensions of the transistors in the circuit, the value of the biasing resistors, the temperature, and the supply voltage. There are dozens of dependent variables monitored, including the circuit currents and node voltages, and functions of these currents and voltages that produce meaningful system metrics. The main metrics pertinent to this circuit are power consumption, power supply rejection ratio, minimal operative supply voltage, output linearity, and output current standard deviation over process and mismatch variation during Monte Carlo analysis. Most of these metrics are referenced during the design process and are thus pertinent to a general understanding of the circuit and its performance.

Output linearity is a quantitative metric undiscussed in all but two of the reviewed papers, but is very important in assessing the merit of the sensor. A nonlinear output usually requires the use of some expensive calibration scheme or other post-processing feature to generate a linear function useful in practical circuit and computation environments. The PMCC metric of (2.3) and especially the R^2 metric of (2.4) are difficult to realize in the Cadence Wavecalc simulation suite. In this thesis nonlinearity of a function is primarily calculated with the simpler devised δ of (3.5):

$$\delta\left(y(x)\right)=\left\langle\left|\frac{\partial y}{\partial x}\Big/\left\langle\frac{\partial x}{\partial y}\right\rangle-1\right|\right\rangle \tag{3.5}$$

or equivalently

$$\delta\left(I_{OUT}(T)\right)=\frac{1}{T_{MAX}-T_{MIN}}\int_{T_{MIN}}^{T_{MAX}}\left|\frac{\partial I_{OUT}}{\partial T}\div\left(\frac{1}{T_{MAX}-T_{MIN}}\int_{T_{MIN}}^{T_{MAX}}\frac{\partial I_{OUT}}{\partial T}dT\right)-1\right|dT \tag{3.6}$$

This formula in effect takes the instantaneous partial derivative of the dependent with respect to the independent and compares it to the average value of the same derivative over the whole domain of the independent, producing an error function. Thus, in an ideal temperature sensor,

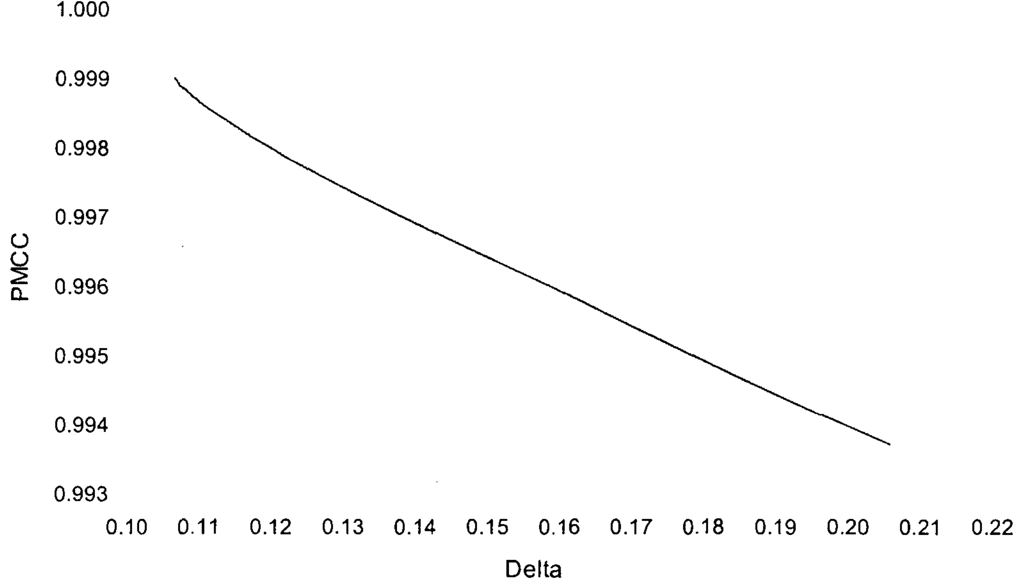


Figure 3.2: PMCC-delta transform.

$$\left\langle \left| \frac{\partial I_{OUT}}{\partial T} \left\langle \frac{\partial T}{\partial I_{OUT}} \right\rangle - 1 \right| \right\rangle \rightarrow 0 \tag{3.7}$$

In practise, an average δ nonlinearity of approximately 10% produces acceptably linear output. The PMCC requires a value of at least 0.999 to indicate a linearity comparable to δ = 10%. There exists a certain transform between the δ-value of a function and its PMCC, shown in Figure 3.2. This is plotted by taking numerical δ and PMCC data from Cadence simulations.

A method to calculate the PMCC for comparison with [4] has been devised with (3.8), and is computationally equivalent to (2.3):

$$PMCC = \frac{1}{\sigma_T \sigma_{I_{OUT}}} \left\langle \left(T - \langle T \rangle\right)\left(I_{OUT} - \langle I_{OUT} \rangle\right) \right\rangle \tag{3.8}$$

σ_T and $\sigma_{I_{OUT}}$ are the standard deviations of T and I_{OUT}, respectively. Internally, the Cadence software finds the outermost average by invoking a numerical integral that, accounting for the standard deviation factorization, is mathematically equivalent to the sigma term in (2.3). Implementing this function in Cadence requires an artificial voltage source equal to the current temperature, only possible by reverse-engineering Spectre's model for temperature coefficients

and obtaining (3.9):

$$V_{\text{effective}} = V_{DC}\left(\mu_1\left(T - T_{NOM}\right) + 1\right) \tag{3.9}$$

Letting V_{DC} = 1, μ_1 = 1 and T_{NOM} = 1 makes the the effective voltage equal to the temperature which can in turn be used in (3.8).

The integral form of the PMCC expression is lengthy, but is important to understand how Cadence may evaluate the PMCC in a continuous dimension instead of a set of discrete data. Let $\phi_{I_{OUT}}$ and ϕ_T be

$$\phi_{I_{OUT}} = I_{OUT} - \frac{1}{T_{MAX} - T_{MIN}} \int_{T_{MIN}}^{T_{MAX}} I_{OUT}\, dT \tag{3.10}$$

$$\phi_T = T - \frac{T_{MAX} - T_{MIN}}{2} \tag{3.11}$$

Then, (3.8) is equivalent to

$$PMCC = \frac{\int_{T_{MIN}}^{T_{MAX}} \phi_{I_{OUT}} \phi_T\, dT}{\sqrt{\int_{T_{MIN}}^{T_{MAX}} \phi_{I_{OUT}}\, dT \int_{T_{MIN}}^{T_{MAX}} \phi_T\, dT}} \tag{3.12}$$

$$= \left. \frac{\int \phi_{I_{OUT}} \phi_T\, dT}{\sqrt{T_{MIN}\left(T_{MAX} - T_{MIN}\right) \int \phi_{I_{OUT}}\, dT}} \right|_{T=T_{MIN}}^{T=T_{MAX}} \tag{3.13}$$

The only metric critical to the assessment of the circuit thus far undiscussed is Monte Carlo standard deviation. Monte Carlo analysis is a widely-used method of determining the distribution profile of a probabilistic function. Repeated trials are conducted and the distribution is built as a histogram. In Cadence, the repeated trials disturb the technology process parameters by a small quantity each time and a histogram of the average value of the circuit current output is built.

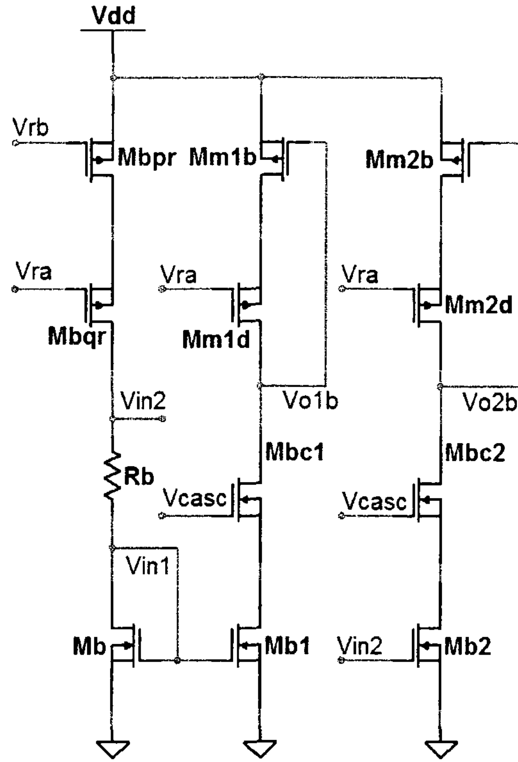


Figure 3.3: The biasing current generator.

The standard deviation of such a histogram is computed using the standard formula of (3.14),

$$\sigma=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(I_{OUT\,i}-\langle I_{OUT}\rangle\right)^2} \tag{3.14}$$

where N is the number of trials conducted. A greater N yields a more accurate distribution. A large value of σ indicates that the circuit is very sensitive to small variations in CMOS process characteristics, such as dopant concentration, and thus would have to have some form of calibration stage. Conversely, a small value of σ indicates that the probability distribution of the output is very narrow and that circuit behaviour remains essentially unchanged during such variation.

The final circuit design represents a compromise between all of these metrics.

3.2 Circuit Equations

This circuit's transistors operate predominantly in the subthreshold region. The simplified

Figure 3.4: The biasing current subtractor, adder and mirrors.

model used for subthreshold region calculations is Equation (A.1) of Appendix A., and is discussed in detail there.

The circuit attempts to implement an equation to directly generate a current proportional to the circuit temperature. As illustrated in Figure 3.3, Mbpr/Mbqr constitute a mirror output from the reference that generates a certain reference current I_{REF}. (Compare Figures 3.3 and 3.5 to Figure 2.8; design changes are described in Chapter 4) This reference current induces a voltage V_{IN2}-V_{IN1} over the resistor R_b and a gate-source voltage over Mb. Transistors Mb1 and Mb2 experience gate voltages separated by the small voltage over the biasing resistor. The resistor is ideally temperature-independent; if thermal variation of the resistor becomes a factor a temperature-invariant equivalent resistor such as in [19] or two temperature-invariant voltage references could be used at the cost of increased power consumption. The Mb1 and Mb2 branches respectively generate two currents, I_{D1} and I_{D2}, separated by a small value; and their mirror gate voltages V_{O1B} and V_{O2B}. If channel length modulation is neglected and transistor dimensions are matched, the ratio of I_{D2} to I_{D1} is given as (3.15) due to (A.1).

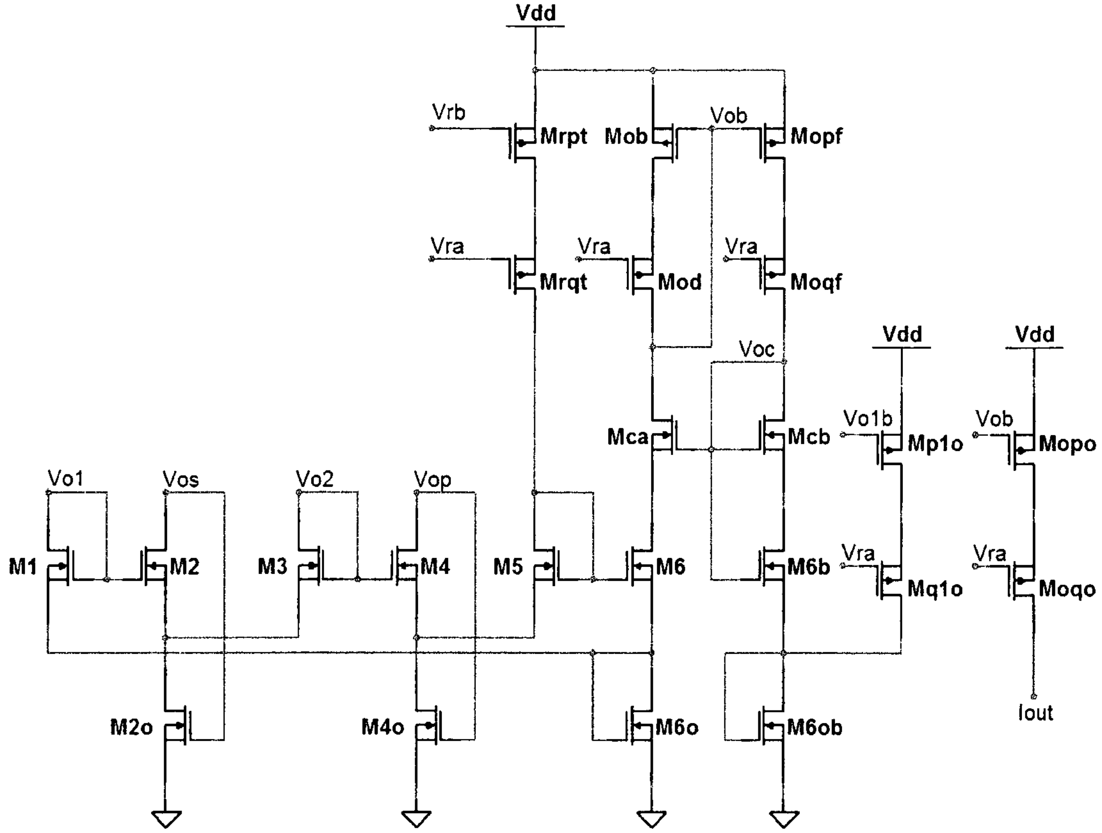


Figure 3.5: The translinear cell and output stage.

$$\frac{I_{D2}}{I_{D1}} = \exp\left(\frac{V_{IN2} - V_{IN1}}{\eta V_T}\right) \quad (3.15)$$

Then, the first term of the Taylor series expansion can give approximations (3.16) and (3.17):

$$\frac{I_{D2}}{I_{D1}} \approx 1 + \frac{V_{IN2} - V_{IN1}}{\eta V_T} \quad (3.16)$$

$$\frac{I_{D1}}{I_{D2}} \approx 1 - \frac{V_{IN2} - V_{IN1}}{\eta V_T} \quad (3.17)$$

This approximation is taken as a step toward calculating the PTAT function. To implement this function, both the sum and the difference of the two biasing currents must be generated. This necessitates the use of several current mirrors, shown in Figure 3.4: Two current mirrors for I_{D1} and I_{D2}, with outputs through Mp1m/Mq1m and Mp2m/Mq2m; two current mirrors to form a current subtractor $I_{D1} - I_{D2}$, with outputs through Mp2s/Mq2s and Mc1s/Mb1s; and two current mirrors to form a current adder $I_{D1} + I_{D2}$ with outputs through Mp1p/Mq1p and

Mp2p/Mq2p.

To arrive at (3.21), the translinear cell of Figure 3.5 must be used. Translinear cells are commonly used in both BJT-based circuits and subthreshold CMOS-based circuits to perform analog computation. In this case the translinear cell is comprised of M1/M2/M3/M4/M5/M6, with M2o/M4o/M6o for biasing. The principle of translinearity can be demonstrated by first applying Kirchoff's Voltage Law over the gate-source potentials of the six transistors, yielding (3.18):

$$\sum_n V_{GS[2n]} = \sum_n V_{GS[2n+1]} \tag{3.18}$$

Due to the exponential model of (A.1), this implies the drain current relationship of (3.19):

$$\prod_n I_{D[2n]} = \prod_n I_{D[2n+1]} \tag{3.19}$$

The drain currents through M1, M2, M3 and M4 are (ideally) I_{D1}, $I_{D2}+I_{D1}$, I_{D2}, and $I_{D2}-I_{D1}$, respectively. The drain current through M5 is I_{REF}, and the drain current through M6 is the output. Thus, the translinear mechanism implements the relationship of (3.20):

$$I_{D1} I_{D2} I_{REF} = (I_{D2}+I_{D1})(I_{D2}-I_{D1}) I_{OUT} \tag{3.20}$$

In terms of I_{OUT}, (3.20) is

$$I_{OUT} \propto \frac{I_{D2} I_{D1}}{(I_{D2}-I_{D1})(I_{D2}+I_{D1})} \tag{3.21}$$

Finally, substituting (3.16) and (3.17) into (3.21) yields

$$I_{OUT} \propto \frac{\eta V_T}{2(V_{IN2}-V_{IN1})} \tag{3.22}$$

and thus the output current is (ideally) a linear function of the silicon's temperature due to the thermal voltage term, as in (A.2).

4 Design

Design of the system required initial calculations according to the simplified model parameters in Appendix A., followed by fine tuning based on simulation. Initial calculations are shown in Appendix C..

4.1 Cascode Current Mirror

Following the theory of Section 3.2, a simple circuit was constructed for simulation in the Cadence software suite using a 180nm CMOS technology. Unfortunately, the circuit exhibited very poor power supply independence when the topology of Figure 2.8 was used with its simple current mirrors.

Higher PSRR requires the extensive application of cascoded current mirrors, reviewed in Section 2.3. An isolated example complete with cascode bias voltage generator is illustrated in Figure 4.1. This mirror requires two identical inputs sIin1 and sIin2. Note that the resistors are only present to demonstrate the performance of non-ideal current sources. The first current is used to bias M5 and M6 in order to generate the cascode biasing voltage at the gate of M6. This current need not be exactly equal to the main mirrored current; a current that is in the same general range as the mirrored current will function satisfactorily. Thus, in the final system design there is only one such branch whose cascode biasing voltage is reused throughout the circuit for all other instances of this mirror design.

All six transistors are typically matched except M5. The width-to-length ratio of M5 must be small to increase the voltage at its gate. Often this aspect ratio is made smaller than the aspect

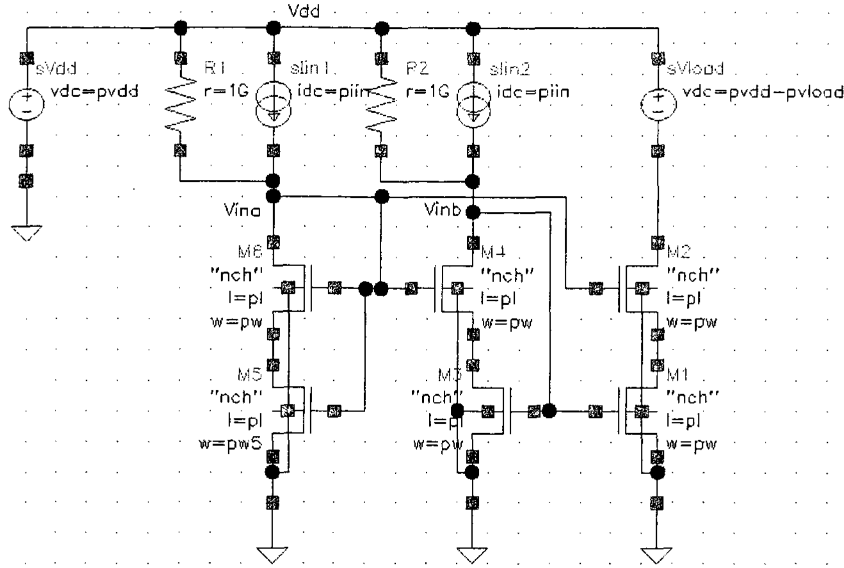


Figure 4.1: Isolated low-voltage cascode current mirror.

pw	20 μm
pw5	400 nm
pl	1 μm
piin	50 nA
pvdd	1 V
pvload	500 mV

Table 4.1: Cascode mirror test parameter values.

ratio of the other transistors by a factor of three to four, but in practise is made even smaller to bias M1 correctly and overcome a mismatch between the thresholds of M6 and M2 due to the body effect [20]. As long as M1 and M3 are matched, their drain-source voltages are equal, and the systematic gain error is nearly zero. However, their input branch voltages will not be equal. Whereas the main input branch has an input voltage equal to a single gate-source potential (M3), the input voltage of the cascode bias generator branch is the sum of the drain-source potential of M5 and the gate-source potential of M6. In the main circuit these potentials are referred to V_{DD} instead of ground because the mirrors are all PMOS-based instead of NMOS-based.

This current mirror design exhibits very low error, very good power supply rejection ratio and very good output resistance. A test circuit was constructed in Cadence to verify the correct operation of this circuit. The test load, sVload, is configured to apply a voltage V_{LOAD} across the

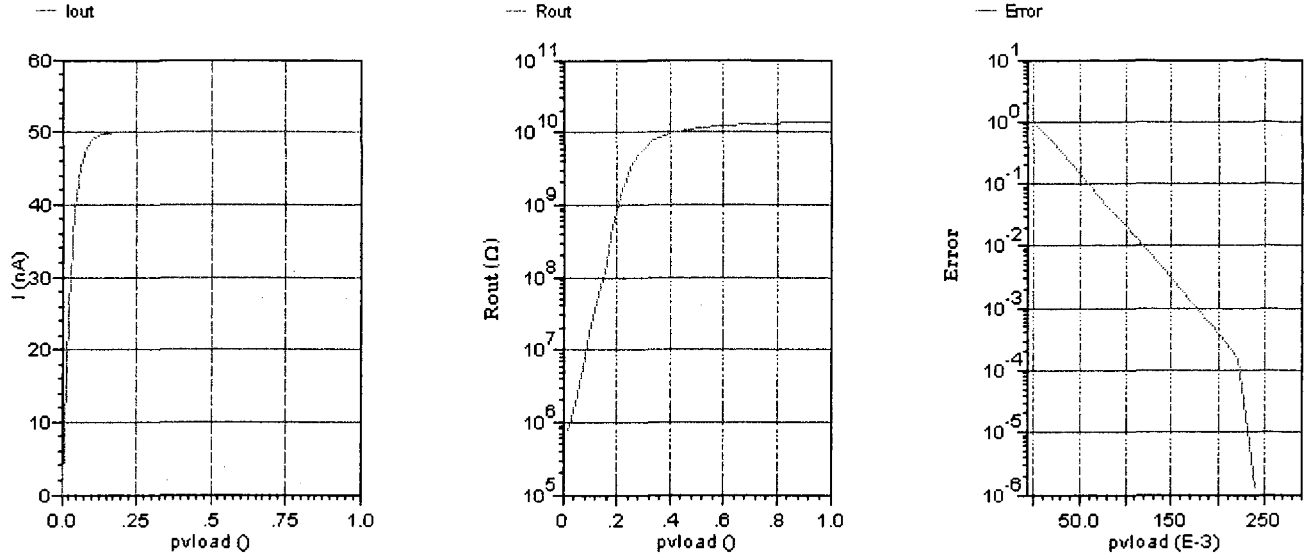


Figure 4.2: I_{OUT}, R_{OUT} and error versus variations in V_{LOAD}.

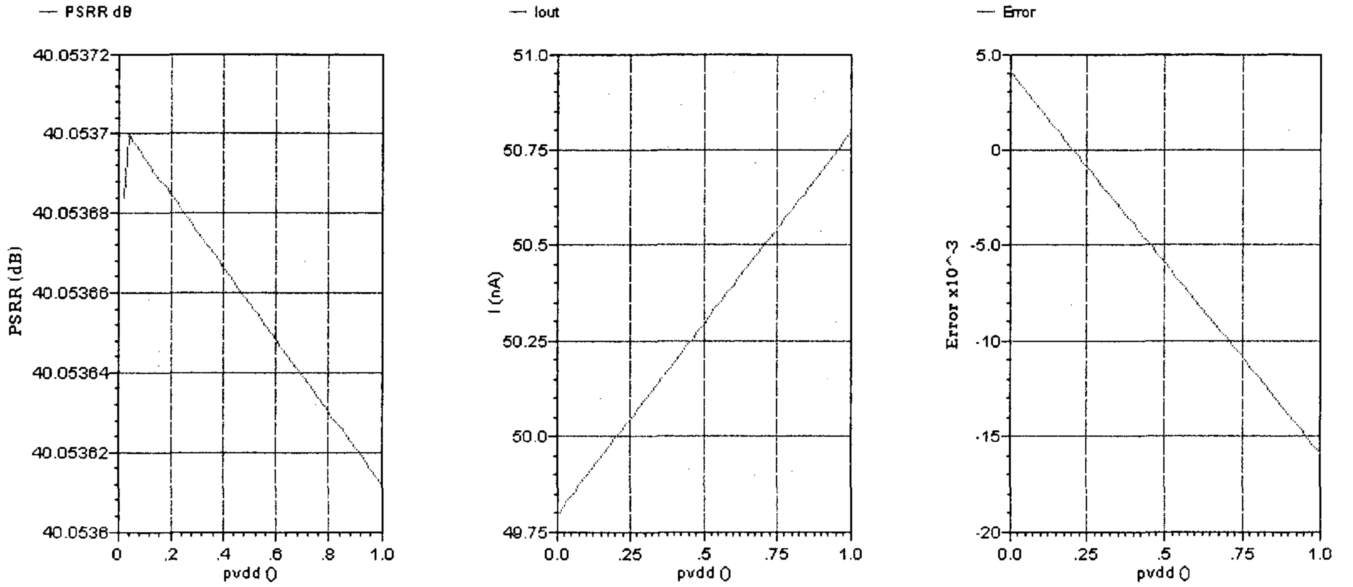


Figure 4.3: PSRR, I_{OUT} and error versus variations in V_{DD} for 1GΩ input sources.

output branch of M1/M2. Nominal values for the test parameters are shown in Table 4.1.

Figure 4.2 shows simulated variation in I_{OUT}, R_{OUT} and error versus variations in V_{LOAD}, calculated with (4.1) and (4.2):

$$R_{OUT} = \frac{\partial V_{LOAD}}{\partial I_{OUT}} \tag{4.1}$$

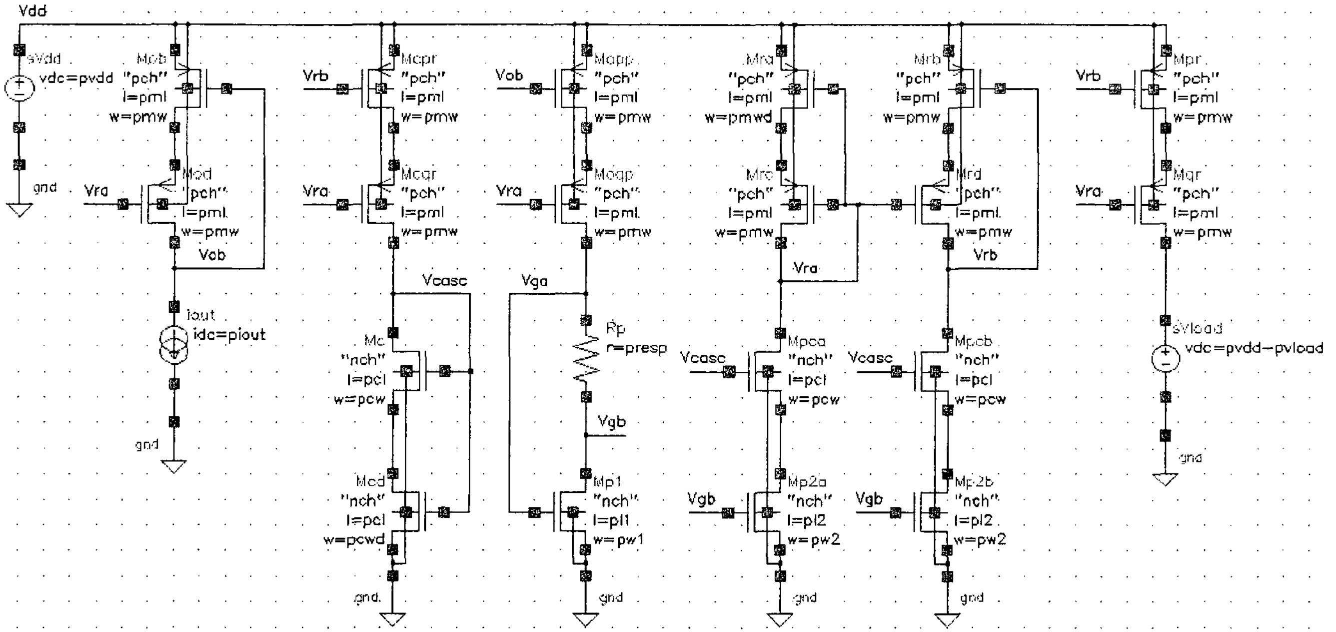


Figure 4.4: Isolated peaking current reference circuit.

$$error = -\frac{I_{OUT}}{50\,\text{nA}} + 1 \tag{4.2}$$

The error goes below 1% for a load voltage of approximately 120mV, and is effectively zero for higher voltages. The output resistance exceeds 10GΩ for a load voltage of approximately 400mV.

Figure 4.3 shows simulated variations in PSRR, I_{OUT} and error versus variations in V_{DD}. V_{DD} is set to vary between 0V and 1V. PSRR is calculated using (3.3). The PSRR in this case is strongly correlated with the quality of sIin1 and sIin2; for ideal current sources the PSRR is well above 100dB, and for current sources with a conservative parallel resistance of 1GΩ the PSRR is 40dB, as shown. The worst-case error is approximately 1.6%. Note that the setup of the two simulations implies that the R_{OUT} case imposes no biasing variation on the first two branches, and the PSRR case does not affect the load voltage; thus the two effects are isolated. In the ideal case that all three branches experience a current of 50nA and for the nominal supply of 1V, the power consumption will be 150 nW.

pvdd	1 V
pvload	500 mV
pl1, pl2, pml, pcl	1 µm
pw1	2 µm
pw2, pmw, pcw	10 µm
pcwd, pmwd	1 µm
piout	50 nA
presp	1 MΩ

Table 4.2: Peaking source parameters.

4.2 Peaking Current Reference

The peaking current reference introduced in Section 2.3 has been used for this thesis. To assess its behaviour independent of the entire system, the example isolated circuit of Figure 4.4 has been simulated.

The essential operation of the peaking current source rests on Mp1 and Rp. As the current through this branch increases between zero and the peak value, Mp1 is turned on. As Mp1 turns on, Vga increases and, since the voltage drop across Rp remains small due to a small current, Vgb is pulled up, in turn pulling up the gate voltages on Mp2a/Mp2b and increasing the output. As the drain current through Mp1 exceeds the value required for the output to peak, the voltage drop across Rp becomes non-negligible and forces Vgb to decrease, in turn decreasing the gate voltages on Mp2a/Mp2b and decreasing the output.

Mpca/Mpcb are cascode stabilization transistors added to improve the circuit's insensitivity to supply variation. Their drains are connected to the two inputs of the cascoded current mirror; note that if a ground-referred reference were necessary this would not be needed, but the rest of the system requires a V_{DD}-referred reference. Mra and Mrc constitute the mirror's cascode bias voltage generator, reused throughout the circuit so that the left portion of all other V_{DD}-referred mirrors may be omitted. Mrb and Mrd are the primary reference mirror input transistors. Mrc's source voltage Vra is used to bias two mirror inputs and three mirror outputs in this example circuit. Mrd's source voltage Vrb is used to bias two reference mirror outputs: the first, to the ground-referred cascode bias voltage generation branch of Mc/Mcd; and the second, to the test

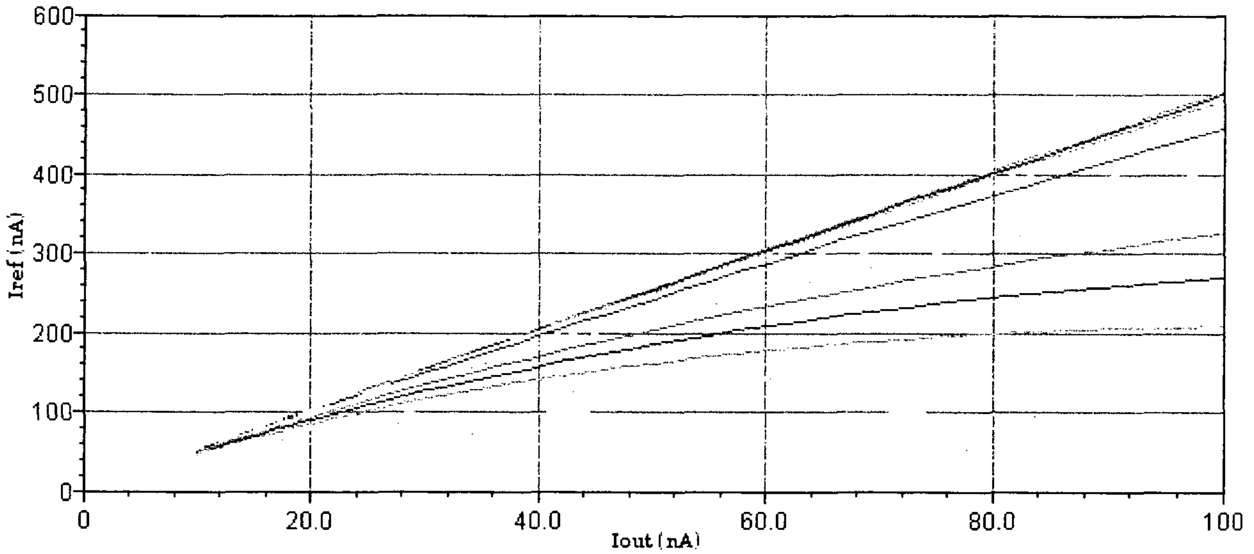


Figure 4.5: Iref versus Iout for Rp from 10kΩ to 1MΩ.

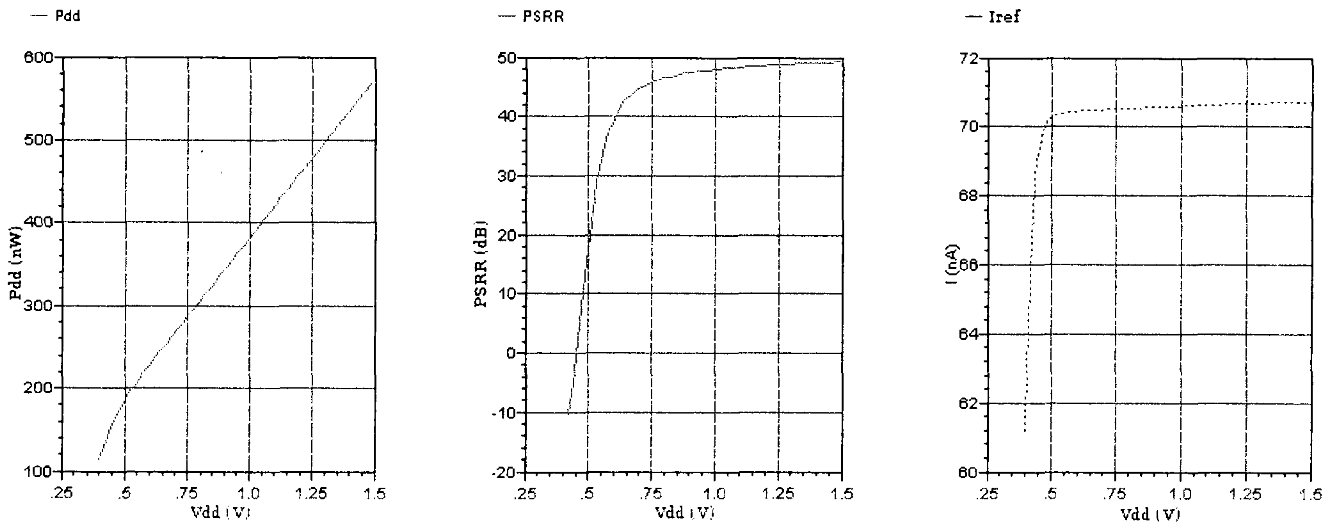


Figure 4.6: Power consumption, PSRR and reference current as a function of supply voltage.

load sVload.

Iout is an idealized current source that models the linearly varying output of the main PTAT circuit. It is connected to the two mirror input transistors Mob/Mod, whose function is to mirror Iout into the input of the peaking source. The only topological differences between this isolated example circuit and the circuit within the main system are the use of Iout and sVload.

The main drawback to this peaking source is a reliance on a large Rp. Figure 4.5 shows Iref versus Iout for a parametric sweep of Rp varying logarithmically between 10kΩ and 1MΩ, with

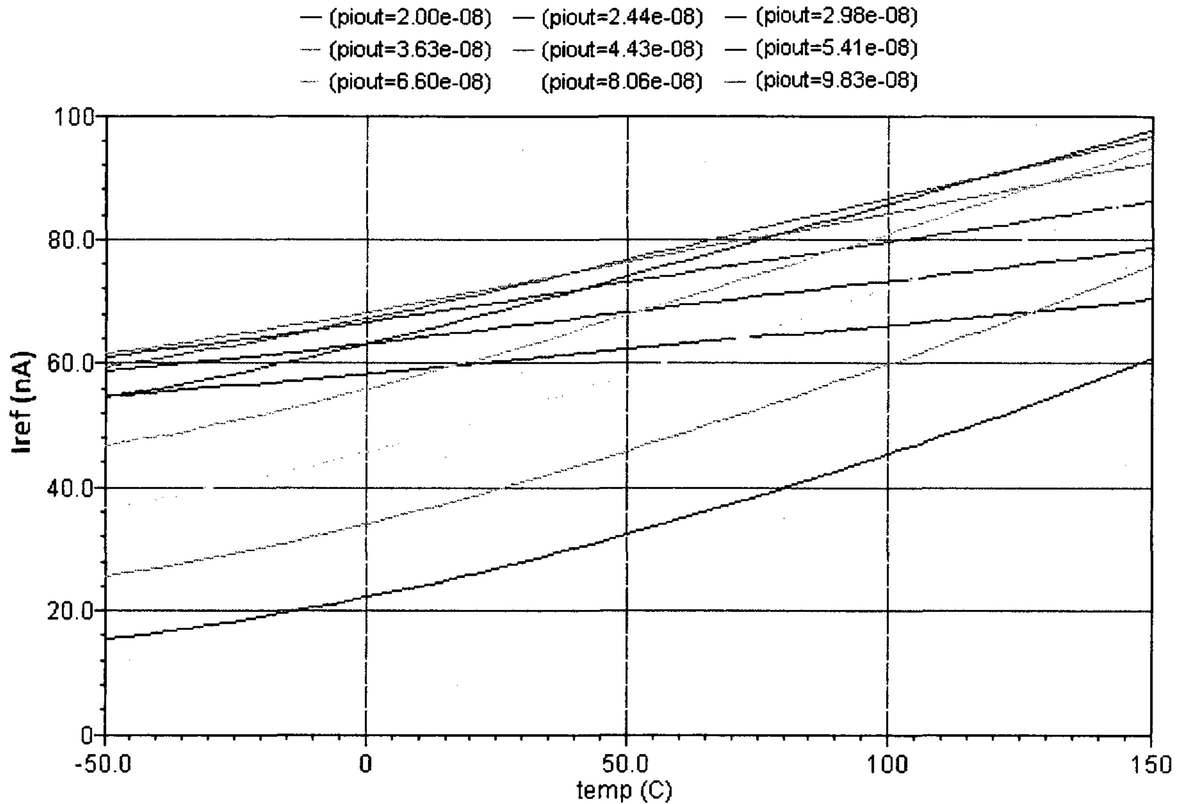


Figure 4.7: Peaking reference current as a function of temperature.

higher resistances producing lower current curves. For currents on the scale of tens to hundreds of nanoamps such as those used to implement the sensor as a low-power circuit, Rp must be on the order of 1MΩ for peaking source behaviour. Such a resistor would usually occupy great area in a VLSI design. Fortunately, there exist simple topologies for implementing high-value, low-error resistors in CMOS; thus, the use of such a resistor value is not unjustifiable [21].

The lowest curve of Figure 4.5 corresponds with the peaking behaviour of the source for the parameters of Table 4.2. Figure 4.6 shows the power consumption (in nW), PSRR, and reference current output as a function of a varying voltage supply, simulated with the same parameters. The minimal operative supply voltage is approximately 500mV. The PSRR approaches an adequate value of 50 dB.

Figure 4.7 shows the behaviour of the peaking current source as temperature varies. Each of the curves represents the reference current produced from the peaking source as a function of temperature for one constant value of the PTAT output current fed to the input of the source.

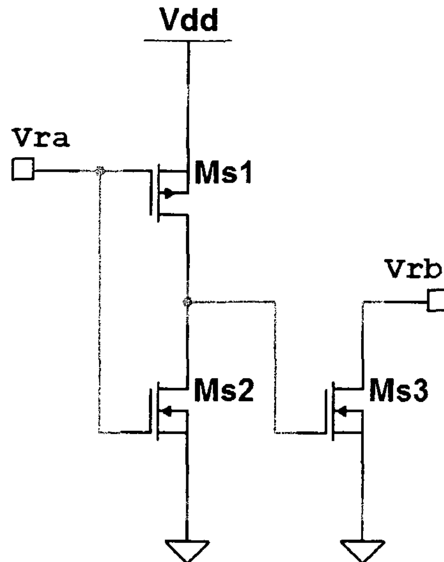


Figure 4.8: The startup circuit.

The peak for each temperature is the curve with the highest Iref. At the lowest temperature shown, a peak reference current of 62 nA is produced from a PTAT current of approximately 36 nA; at the highest temperature shown, a peak reference current of 98 nA is produced from a PTAT current of approximately 54 nA.

4.3 Startup Circuit

The system is a feedback loop through the main sensor circuit and the peaking current source. This loop has two points of stability, one where the output exhibits expected characteristics and one where most circuit branches are off. To ensure that the system operates in the former (active) region, the small startup circuit of Figure 4.8 is used.

The startup circuit is a simple and common design. Ms1 and Ms2 comprise an inverter. The input to the inverter is the cascode mirror bias voltage Vra. When this voltage goes below a certain threshold value, the inverter goes high and Ms3 is activated. Ms3 then pulls down Vrb, the reference mirror voltage, to activate the circuit. When Vra is above the inverter threshold, Vstart is low, Ms3 is inactive and the system's behaviour is unaffected.

To illustrate the operation of the startup circuit, Figure 4.9 shows a contrived startup situation where Vdd is seen to increase linearly over a period of 1 ms. Initially, Vra follows Vdd

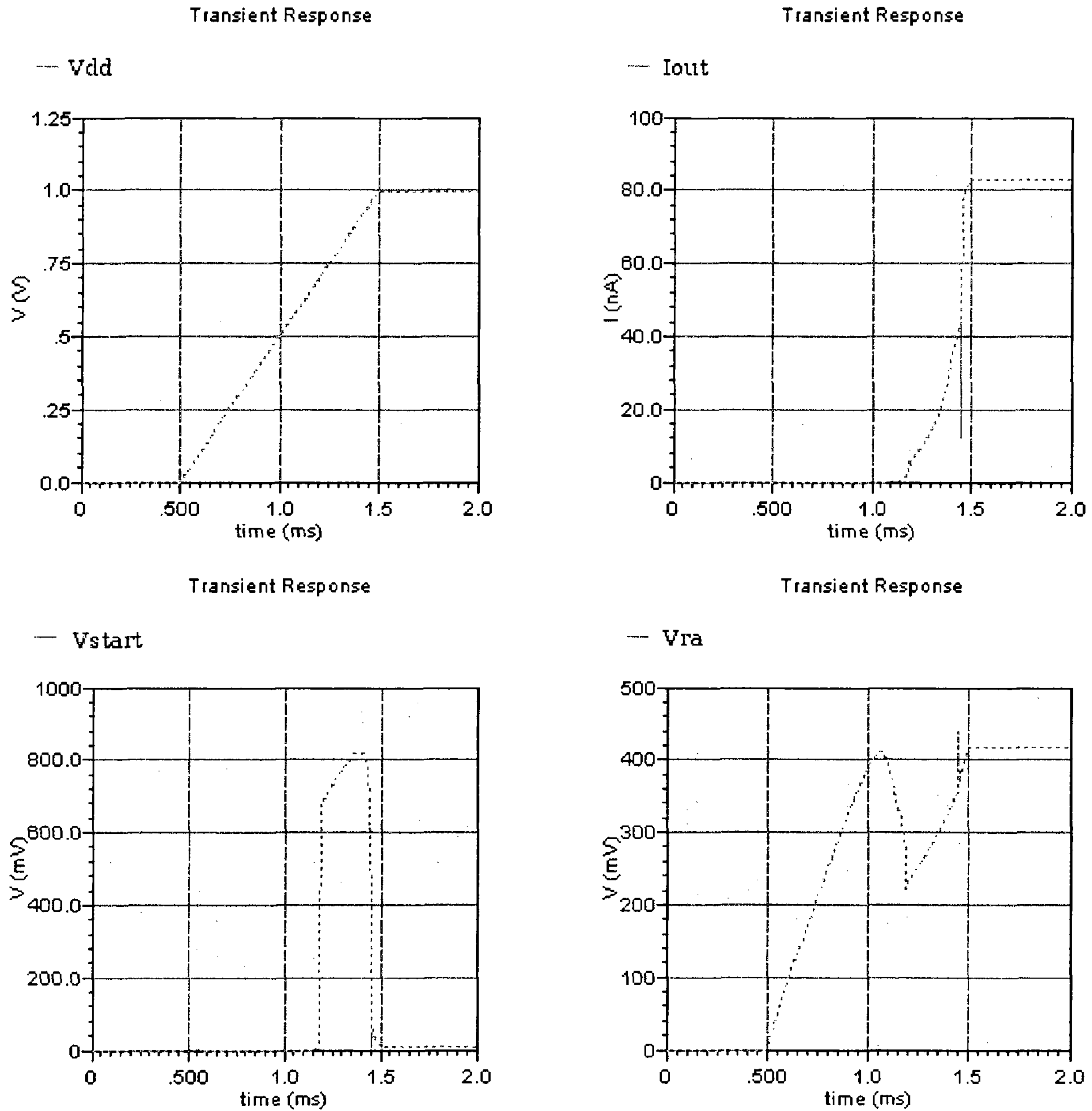


Figure 4.9: Vdd, Iout, Vstart and Vra transients during startup.

indicating that the reference circuit is off, and Iout is zero. After approximately 1.2 ms, the startup inverter has sufficient voltage to activate, pulling down Vrb and successfully starting the system.

Figure 4.10: The main thermal biasing section.

Figure 4.11: The ground-referred cascode bias branch and peaking source.

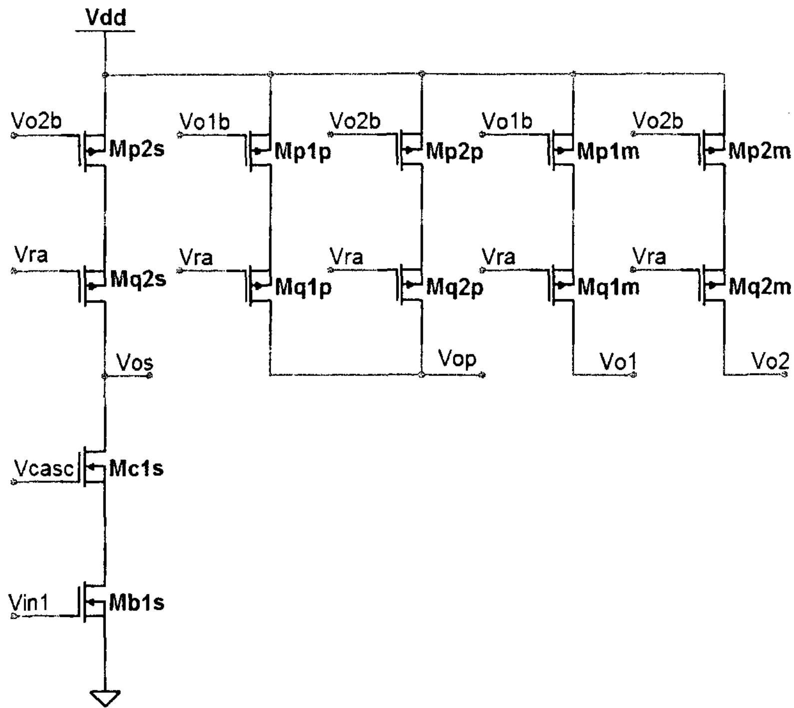


Figure 4.12: Thermal current subtractor, adder and mirrors.

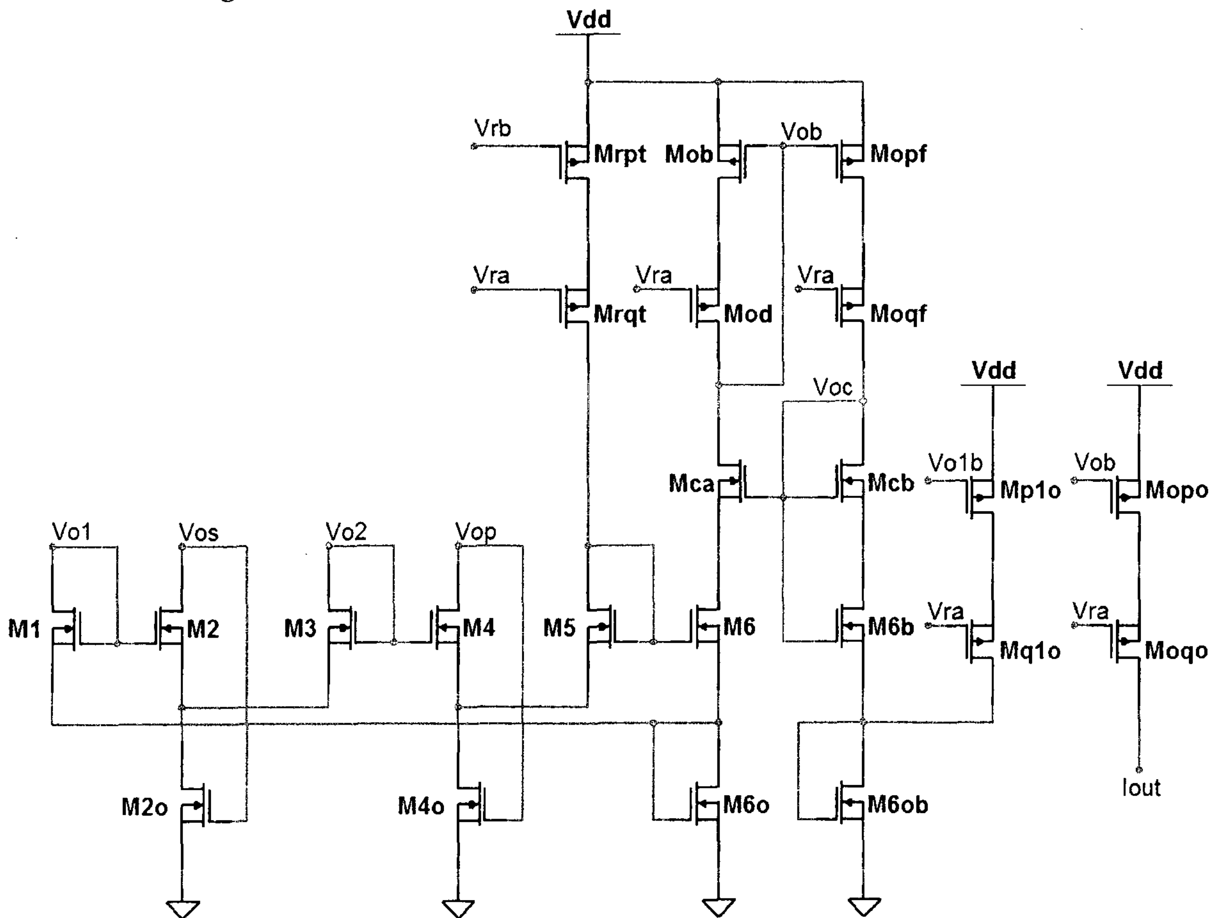


Figure 4.13: The translinear cell and output block.

4.4 Complete Topology

The complete system topology includes the startup circuit of Figure 4.8; the ground-referred cascode bias branch and the peaking source shown in Figure 4.11; the main thermal biasing section shown in Figure 4.10; the thermal current subtractor, adder and mirrors shown in Figure 4.12; and the translinear cell and output block shown in Figure 4.13.

The output block is designed to be self-biased and to avoid interference with the operation of the translinear cell while still preserving voltage supply insensitivity. Mca and Mcb are cascode transistors to improve stability. M6ob is biased to match the operating point of M6o by receiving the sum of Iout (from Mopf/Moqf) and Id1 (from Mp1o/Mq1o), just as M6o receives the sum of Iout (controlled by M6) and Id1 (from Mp1m/Mq1m). M6 and M6b are matched. Mcb generates the cascode bias voltage for Mca.

4.5 Final Design Parameters

Transistor	W (nm)	L (nm)
Ms1	3000	1000
Ms2	10000	180
Ms3	20000	1000
Mcpr, Mcqr, Mopp, Moqp, Mbpr, Mbqr, Mp2s, Mq2s, Mp1p, Mq1p, Mp2p, Mq2p, Mp1m, Mq1m, Mp2m, Mq2m, Mrpt, Mrqt, Mopf, Moqf, Mp1o, Mq1o, Mopo, Moqo	20000	180
Mc, Mpca, Mpcb, Mbc1, Mbc2, Mc1s, Mca	20000	1000
Mcd	1000	1000

Transistor	W (nm)	L (nm)
Mp1	1000	500
Mp2a, Mp2b	4000	1000
Mra	500	1000
Mrb, Mrc, Mrd, Mm1b, Mm1d, Mm2b, Mm2d, Mob, Mod	20000	1000
Mb, Mb1, Mb2, Mb1s	10000	1000
M1, M2, M3, M4, M5, M6, M6b	2500	290
M2o, M4o, M6o, M6ob	3000	2500
Mcb	1000	5000

Table 4.3: Final transistor dimensions used.

Table 4.3 shows the dimensions of all transistors in the final design. Note the scaling factor of 180:1000 on most of the major mirrors; this is done to improve gain and to ensure that the translinear cell is correctly biased.

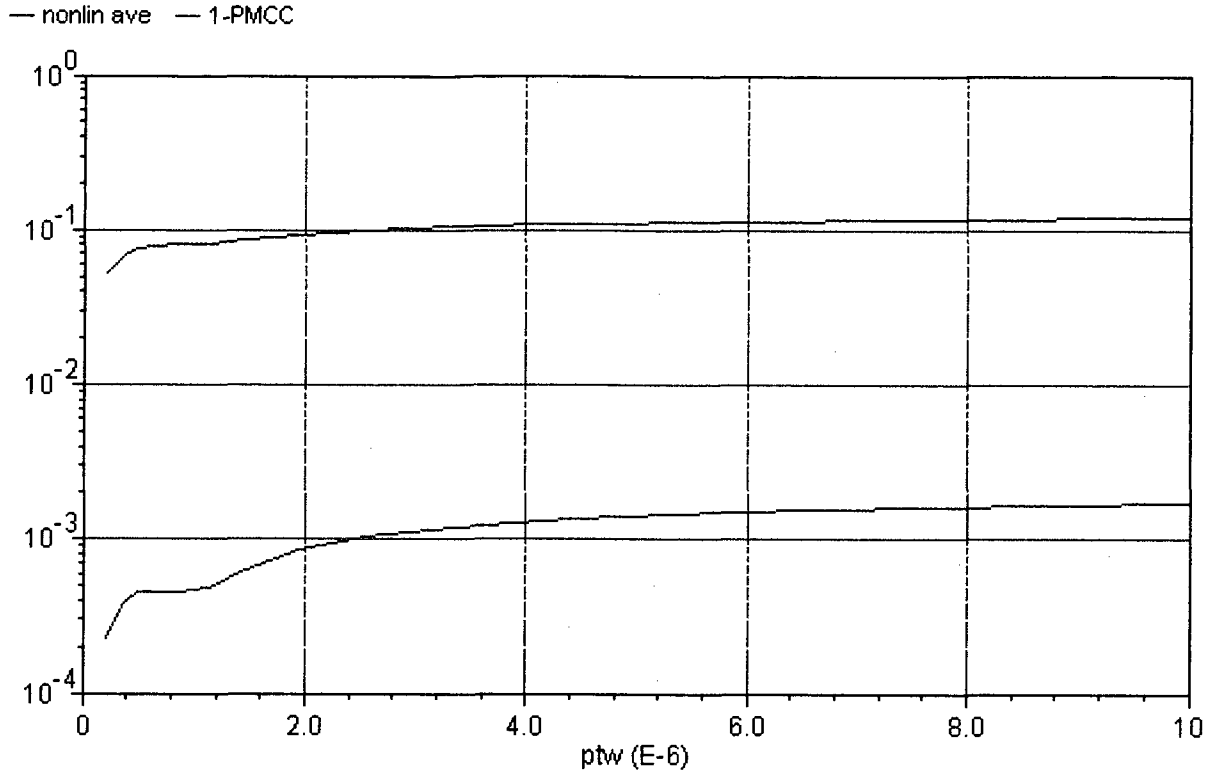


Figure 4.14: An example parametric sweep of δ and PMCC against M1-M6 width.

These dimensions require significant fine-tuning to optimize the system's performance. Such fine-tuning is performed with parametric sweeps of design variables, monitoring the resulting value of performance metrics. Figure 4.14 shows an example of optimization for linearity where the widths of M1 through M6 are swept between 220 nm and 10 μm, and the PMCC and δ metrics are plotted. Each coordinate of transistor width and PMCC/δ represents a sub-simulation in which the temperature is varied between -50°C and 150°C and the linearity metrics are taken from the resulting I_{OUT} curve. Thus, the parametric sweeps are based on the surface function of $I_{OUT}(W, T)$, producing the two curve functions PMCC(W) and δ(W). Note that in the graph PMCC is expressed as 1 – PMCC to facilitate logarithmic plotting.

5 Simulation Results

5.1 Thermal Variation

Given the relationship of (3.15), η may be found with (5.1):

$$\eta = \frac{V_{IN2} - V_{IN1}}{V_T \ln(I_{D2} - I_{D1})} \tag{5.1}$$

The extent to which (3.15) is true may be measured by the extent to which (5.1) is constant over a varying temperature. This is demonstrated in the graph of Figure 5.1. The change in η is small enough to be negligible. Note that this is not a graph of the true η parameter, as it would be ideally perfectly constant; the graph illustrates second-order effects that are not accounted for in the simplified subthreshold transistor behaviour model.

Figure 5.2 shows the graphs of I_{D1} and I_{D2}. Figure 5.3 shows the final PTAT output I_{OUT}. The output is highly linear. Linearity depends on the chosen temperature range. Table 5.1 shows three example temperature ranges and their corresponding linearity metrics when V_{DD} = 1V. Narrower ranges select for regions of the output function that have better linearity. The first range shown may be considered the maximally wide range, as the output current approaches zero for temperatures below -90°C, and the function becomes highly nonlinear for temperatures above 150°C.

5.2 Power Supply Variation

The circuit is acceptably insensitive to power supply variation, with a PSRR exceeding 40dB

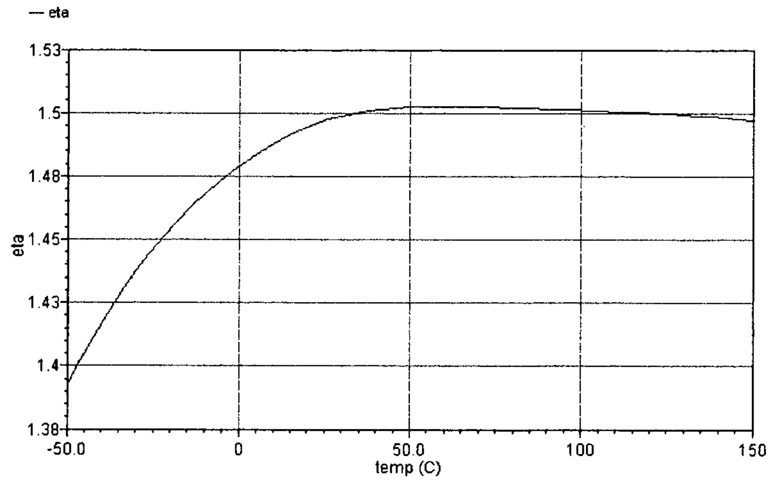


Figure 5.1: η as derived from the ratio of I_{D2} to I_{D1}.

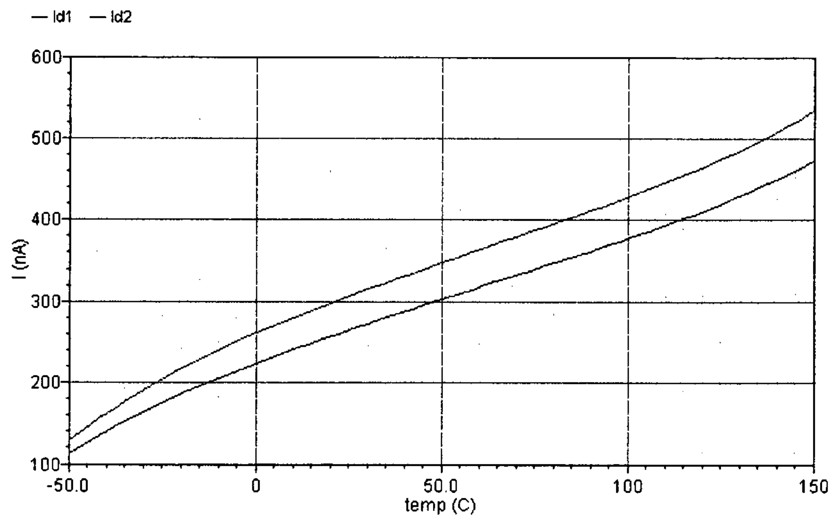


Figure 5.2: I_{D1} and I_{D2}.

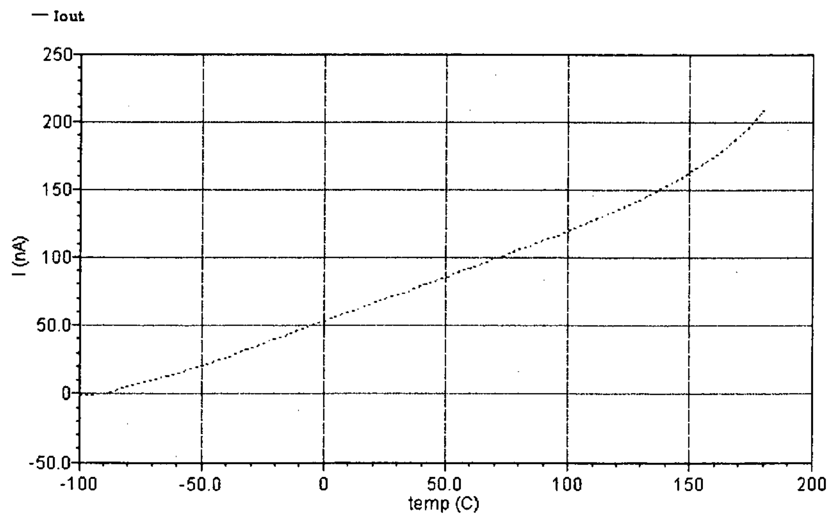


Figure 5.3: The PTAT output current.

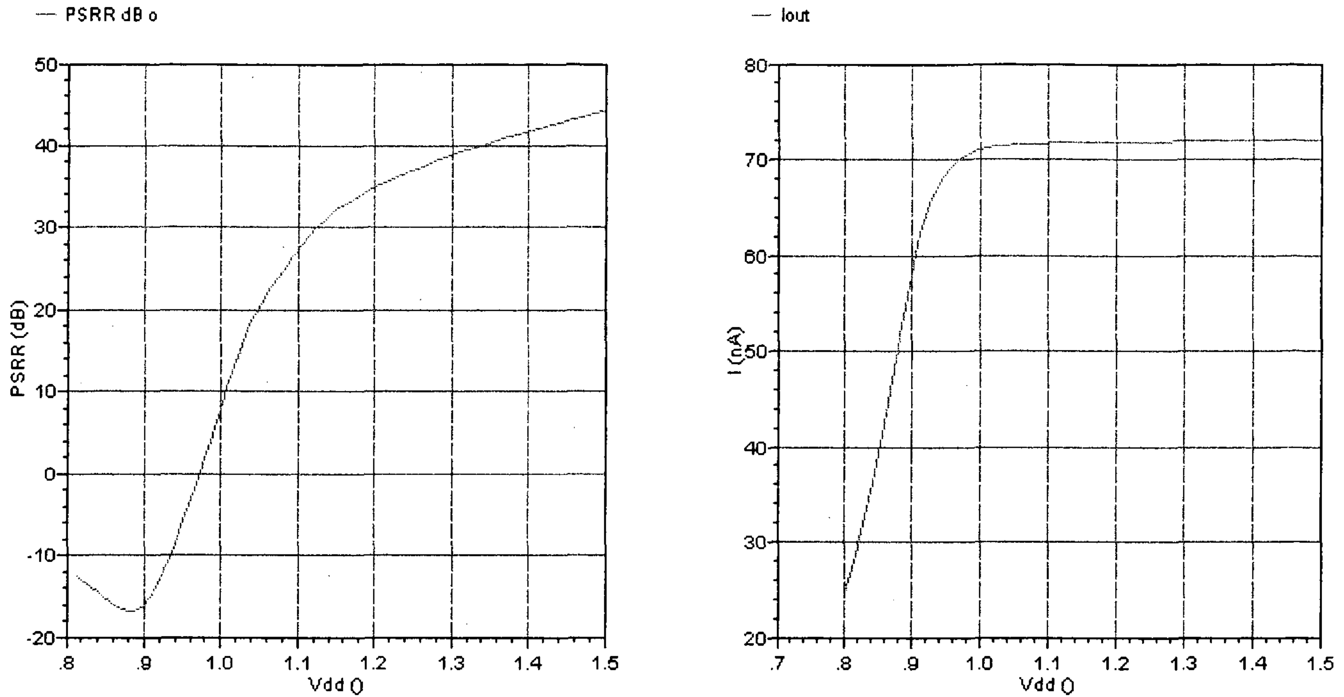


Figure 5.4: PSRR and I_{OUT} as a function of V_{DD}.

T_{MIN} (°C)	T_{MAX} (°C)	δ	PMCC
-90	150	11.06%	0.998731
-75	125	6.126%	0.999686
-50	100	2.213%	0.999971

Table 5.1: Temperature range and corresponding linearity metrics.

for upper values of V_{DD}. Figure 5.4 shows PSRR and I_{OUT} as functions of the supply voltage for the nominal temperature of 300°K. The minimal supply voltage is approximately 1V.

To ensure that the circuit output remains linear over a wide range of supply voltages, a parametric series of simulations constructing the surface function of I_{OUT}(T, V_{DD}) was performed, yielding PMCC(V_{DD}). As shown in Figure 5.5, the output remains highly linear for all values of V_{DD} above the minimal operative voltage.

5.3 Monte Carlo Analysis

Monte Carlo analysis is useful to assess the stability of the circuit over variations in the technology process parameters ("inter-die" variations) and matching. The most informative way

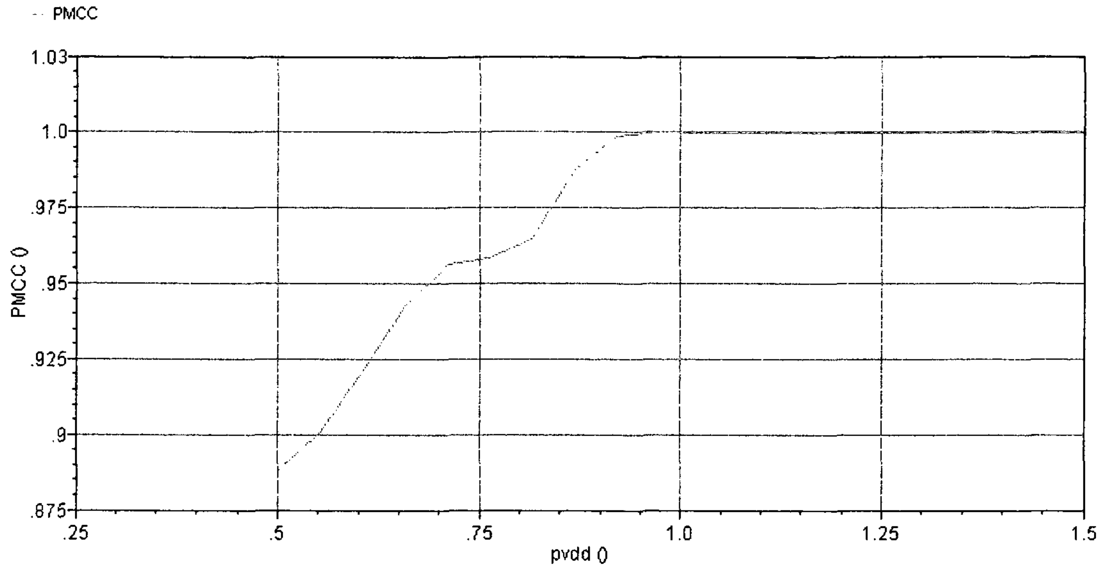


Figure 5.5: PMCC as a function of V_{DD}.

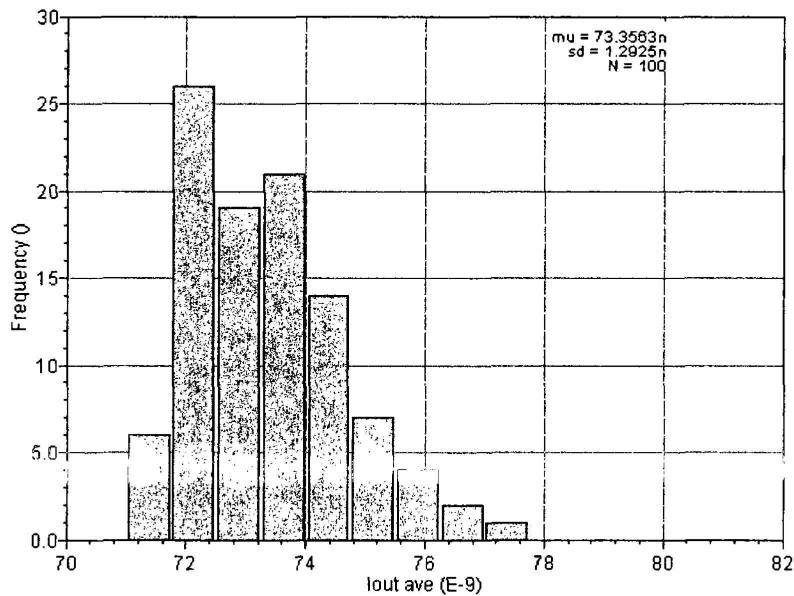


Figure 5.6: Monte Carlo histogram of average I_{OUT} for process and mismatch variation.

to do this is to set Cadence to run many simulations, slightly changing the process parameters each time. When the outputs are single values (rather than functions of the simulated independent variable, such as T or V_{DD}), Cadence will display a histogram of these outputs along with basic statistical information. This histogram is shown in Figure 5.6. The standard deviation is approximately 1.29 nA, or 1.76% on the mean of 73.36 nA shown in the figure.

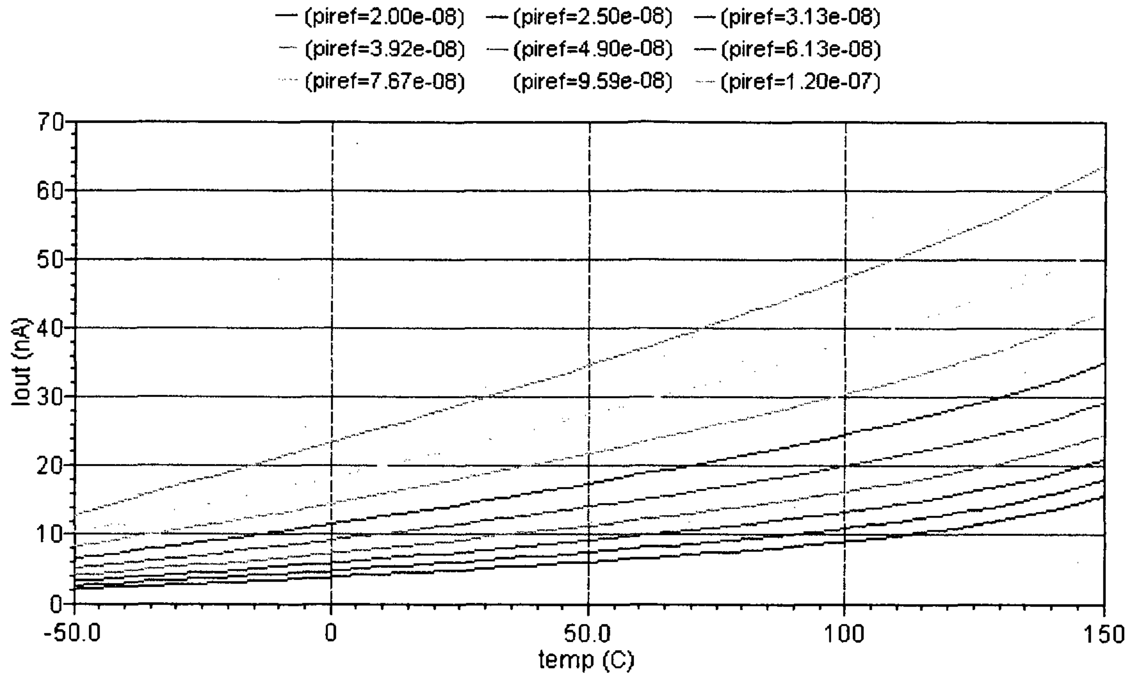


Figure 5.7: Output current as a function of temperature and reference current.

5.4 Reference Behaviour

An additional simulation was performed wherein the output from the peaking current source was discarded, and the reference current is varied parametrically to examine the effect on the current output. Figure 5.7 shows a set of curves of Iout versus temperature, each for a specific constant value of Iref. Comparing any curve in this temperature-independent reference experiment to a curve of equal average gain in the final system with the peaking source included, the latter has better linearity in the high end.

5.5 Comparison with Literature Review

It is important to compare the system's performance characteristics to those assessed in the Literature Review. Table 5.1 is a reiteration of Table 2.1 but with the first row displaying metrics of the current system.

Src.	V_{DD}	Power	PSRR	T range	Sensitivity	Linearity	Area	Technology	Monte Carlo deviations
	> 1V	**12.95 µW ave**	**>30dB [V_{DD}>1.12] >40dB [V_{DD}>1.34]**	**-90°C - 150°C**	**673.1 pA/°C ave**	**Excellent**	**>520µm²**	**180nm CMOS**	**1.76%**
[2]				30°C - 150°C	~1.8mV/°C	Good		50nm CMOS predictive model	~10%
[4]				10°C - 90°C	0.95 – 1.15 °C/bit	Excellent		65nm CMOS	
[6]	5V				2.7V/°C			1.2µm BiCMOS	
[7]				-23°C - 157°C				500nm CMOS	
[8]		> 16nW		25°C - 350°C	~ 37Ω/°C	Acceptable	>1200µm²	Custom SOI	
[5]	3.3V	< 50µW	32.2dB	0°C - 75°C	1.2V/°C	Good	1600µm²	180nm CMOS	0.43% or 6.14%
[9]	>950 mV	< 5µW	>60dB	0°C - 50°C	220µV/°C, 1.13 nA/°C	Good	0.05mm²	1.2µm and 350nm CMOS	<~5%
[10]	1.3V	80µW	>100dB	-50°C - 130°C				180nm CMOS	0.5%
[11]	2.5V (1.8V-3.3V)	300µW-2.2mW		20°C - 60°C			< 1mm²	250nm CMOS	Calibrated
[12]	1.5V	<5.8µW		-20°C - 100°C	113pA/°C, 21.4Hz/°C	Good		350nm CMOS	
[13]		28.53µW		-55°C - 170°C	264µV/°C	Excellent	1260µm²	130nm and 180nm CMOS	

Table 5.2: Comparison of metrics from the reviewed PTAT designs with the current design.

The minimal voltage supply, nominal power consumption, temperature range, sensitivity, area and Monte Carlo stability are all quite competitive in comparison to other designs in the literature review. The area shown for this design is simply the sum of the transistor areas but of course in practise once layout has been performed, this figure will increase. The PSRR is not

among the best but is adequate for a variable supply context. Depending on the temperature range considered, the linearity can be said to exceed that of all other circuits whose linearity is quantified: This design's potential PMCC of 0.999971 is greater than the next best circuit, [4], having one instance with a PMCC of 0.9996.

6 Conclusions and Future Work

6.1 Conclusions

The design goals for this system have been met. The described circuit exhibits very low power consumption, very high linearity, acceptable power supply variation insensitivity, and acceptable characteristics in the other areas described in Section 5.5. The circuit is competitive with other reviewed designs, and exceeds the performance of other designs in many aspects.

The most significant advantage to this design is a very highly linear output for a wide range of supply voltages. This feature is especially important in DVS contexts where the sensor must exhibit consistent behaviour despite changes in the supply. The linearity may be attributed to the unique formulation of the PTAT output relying on subthreshold transistor characteristics, as modified from [12]. This formulation has the added advantage of being inherently largely independent of process parameters.

The greatest disadvantage to this design is its complexity. The circuit relies on many subtle mechanisms to function correctly; thus debugging, and presumably transferral to other technologies or simulation environments, is difficult and error-prone.

6.2 Future Work

CMOS layout was not done in this thesis. In the future, layout may be performed and the sensor may be fabricated for testing and comparison against simulated results. The fabrication process is normally quite expensive, but there are some opportunities for academic discounts or

complete waiving of fabrication fees through corporations such as CMC.

An analog-to-digital converter (ADC) may be added to the sensor for application in digital system design. Such an ADC would have to adopt the same design objectives of adequate power supply insensitivity and power consumption, with added metrics of conversion speed and quantization error (among others).

The circuit may be reimplemented in a smaller, more modern technology such as a 45 nm CMOS process, widely available since 2007.

This work made the assumption that the main biasing resistor is temperature-independent. Future work could be done to examine the effect of temperature variation on the biasing resistor.

Appendix A. Technology Parameter Extraction

As discussed in the Introduction, it would be very difficult to work with BSIM equations directly. They are very complex, and are implemented with the use of many piecewise functions, each tailored to a specific operating region of an actual fabricated CMOS transistor and tuned to match its real-life characteristics. Since this thesis uses transistors predominantly in the subthreshold region, it is justifiable to reduce the model to a more simple expression. The simplified expression for drain current is a function of the transistor's width, length, a constant I_{ON} assumed to be a parameter of the CMOS process, the gate-source potential, the threshold potential V_{TH} and constant of proportionality η also assumed to be process parameters, the output resistance parameter λ, and the drain-source potential. It is given as (A.1):

$$I_D \approx \frac{W\, I_{ON}}{L(1-\lambda V_{DS})} \exp\left(\frac{V_G - V_{TH}}{\eta V_T}\right) \quad \text{(A.1)}$$

V_T is the thermal voltage, given as (A.2)

$$V_T = \frac{k\, T}{q_e} \quad \text{(A.2)}$$

where k is Boltzmann's constant, T is the absolute temperature, and q_e is the electron charge.

The simple circuit of Figure A.1 is simulated in Cadence. Three different simulations are performed: drain current as a function of temperature, drain current as a function of gate-source voltage, and drain current as a function of drain-source voltage. A table of values is recorded from each of these and inserted into a Matlab file (refer to Appendix A.).

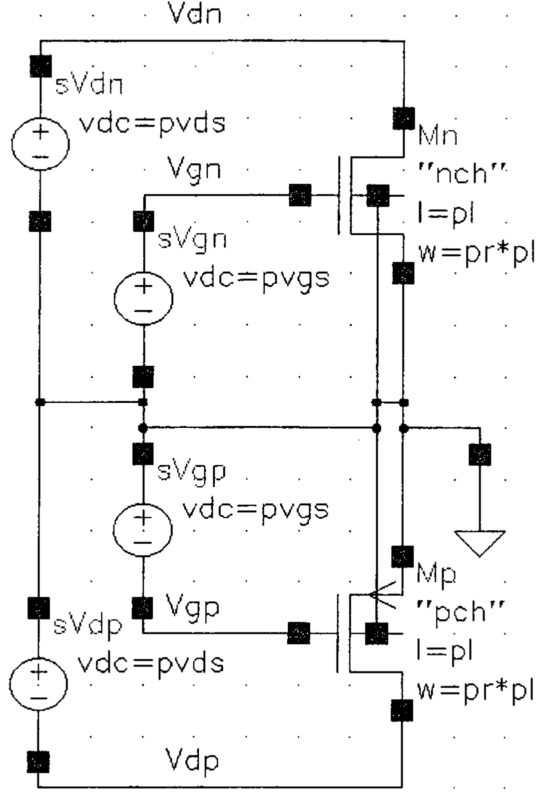


Figure A.1: The parameter extraction circuit.

The Matlab script constructs the function of (A.1) and performs a four-dimensional regression over the parameters of I_{ON}, V_{TH}, η, and λ. Matlab has no direct regression feature. Instead, it has a multidimensional minimization function *fminsearch*. The minimization function repeatedly calls a custom error function until the error is considered minimized or until the alloted number of execution cycles has elapsed. In the case of this regression, the error function has been written to use the simple relative least-squares formula of (A.3):

$$\text{Relative error} = \left(\frac{I_{\text{Regressed}}}{I_{\text{Simulated}}} - 1 \right)^2 \tag{A.3}$$

Table A.1 shows the numerical results of the regression. The listed "nominal" values are employed when a certain variable is not being swept during regression. For the given transistor dimensions and biasing voltages, and for the nominal temperature of 300°K that Cadence uses, a viable set of technology parameters have been extracted. With the exception of η, these parameters are subject to variation based on changes in the transistor dimensions, and these variations are difficult to predict. Nonetheless, having an approximate model facilitates design

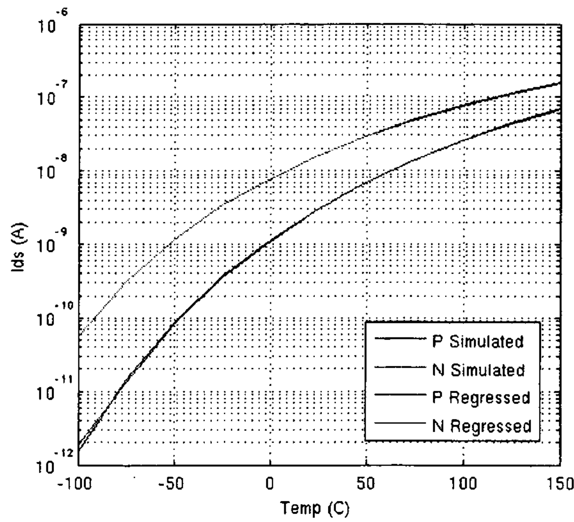
Dimensions	5μm/1μm	
V_{GS} nominal	250 mV	
V_{DS} nominal	500 mV	
T nominal	300 °K (26.85 °C)	
Extracted parameter	**PMOS**	**NMOS**
I_{ON}	22.838 μA	7.991 μA
V_{TH}	604.913 mV	515.468 mV
η	1.305	1.315
λ	0.09662 V^{-1}	0.09989 V^{-1}

Table A.1: Assumed and extracted Matlab regression values.

calculations that would not otherwise be feasible with the BSIM model.

Figure A.2 illustrates the close correspondence between the simulated and regressed temperature dependence curves. Similarly, Figure A.3 illustrates the correspondence between the simulated and regressed gate-source-voltage-to-drain-current transfer function. Figure A.4 illustrates the output resistance effect. Note that simulated data are not included for the triode region as this would harm the regression quality. All three error figures are plots of (A.3).

In terms of practical in-circuit behaviour, I_{ON} denotes the drain current experienced when the gate-source and threshold potentials are equal and the drain-source potential is at a value that does not induce any output resistance effect. V_{TH} denotes the potential below which a gate-source voltage puts the transistor in the subthreshold region. λ controls the partial derivative of the drain current with respect to drain-source voltage, and thus controls the output resistance effect. In most of the Theoretical Operation section this effect is neglected, but it becomes important in the Design section, especially when considering PSRR.



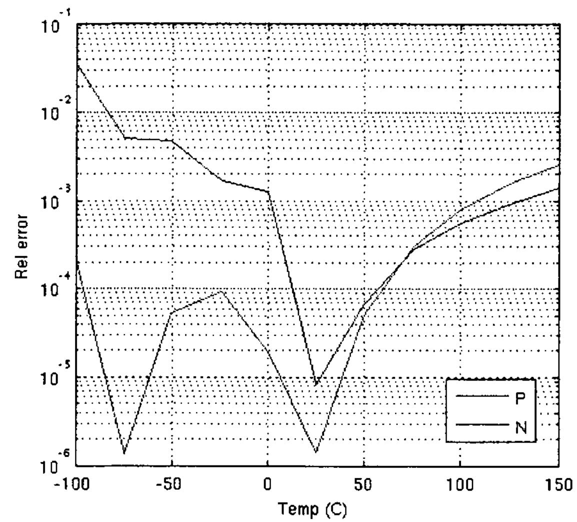
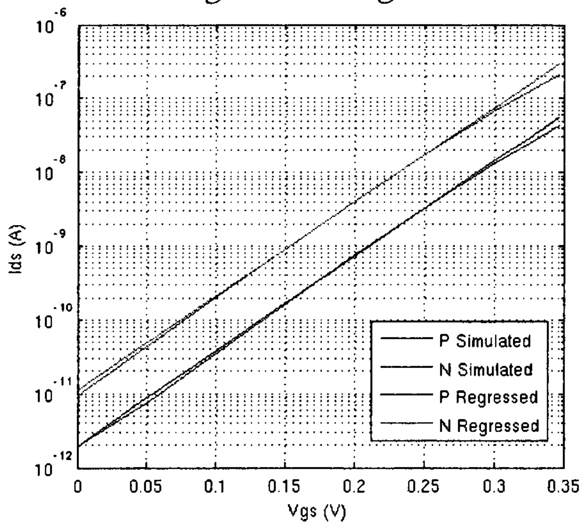


Figure A.2: Regression on temperature and corresponding error



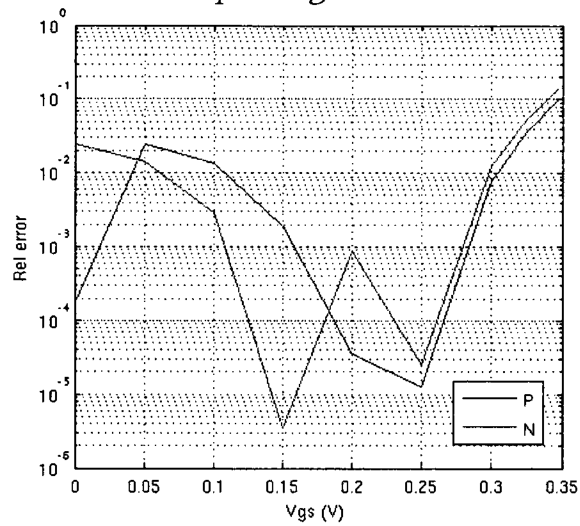
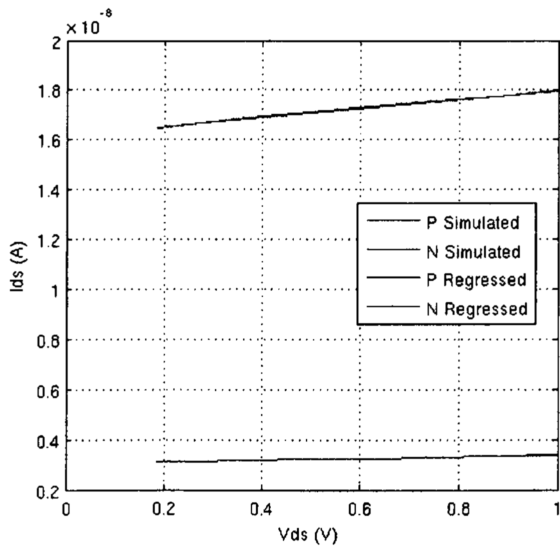


Figure A.3: Regression on V_{GS} and corresponding error



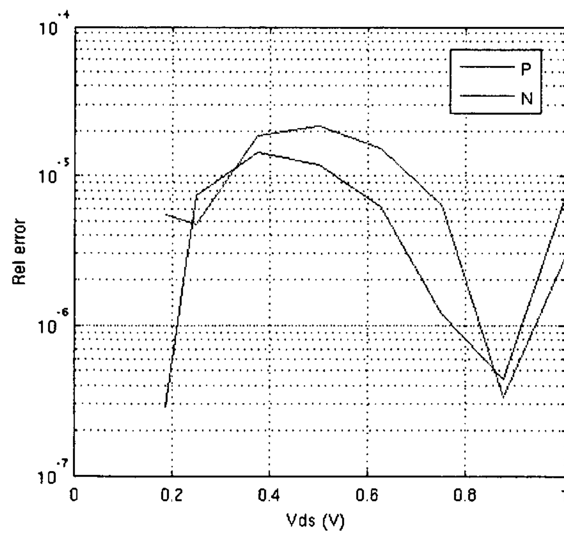


Figure A.4: Regression on V_{DS} and corresponding error

The following is Matlab code used to perform the regression and graph the results.

```
function estimate
  % Set up data from simulation
  clc; clear all; format short eng;
  global ids_temp; global ids_vgs; global ids_vds;
  global     temp; global     vgs; global     vds;
  global params;
  % temp, id pmos, id nmos;
  ids_temp = [
     -100.00,    1.7874e-12,     55.023e-12;
     -74.952,    13.527e-12,     308.45e-12;
     -50.385,    78.608e-12,      1.1457e-9;
     -24.374,    355.05e-12,      3.4494e-9;
    89.02e-3,     1.1115e-9,      7.9026e-9;
      25.241,     3.0403e-9,      16.373e-9;
      49.807,     6.8758e-9,      29.505e-9;
      75.337,     14.193e-9,      49.688e-9;
      99.904,     25.953e-9,      76.548e-9;
      124.95,     44.508e-9,      112.39e-9;
      150.00,     71.731e-9,      157.52e-9; ];
  % vgs, id pmos, id nmos;
  ids_vgs = [
    0.00000,    1.9223e-12,     9.4098e-12;
    49.904e-3,  7.3919e-12,     42.270e-12;
    99.807e-3,  33.676e-12,     195.02e-12;
    0.14971,    158.09e-12,     893.93e-12;
    0.19961,    728.67e-12,     3.9905e-9;
    0.24952,    3.1929e-9,      16.902e-9;
    0.29942,    12.837e-9,      65.647e-9;
    0.32505,    25.116e-9,      126.12e-9;
    0.34663,    43.147e-9,      213.24e-9; ];
  % vds, id pmos, id nmos;
  ids_vds = [
    0.18644,    3.1292e-9,      16.474e-9;
    0.24862,    3.1554e-9,      16.654e-9;
    0.37385,    3.1984e-9,      16.907e-9;
    0.49908,    3.2380e-9,      17.135e-9;
    0.62431,    3.2765e-9,      17.351e-9;
    0.74954,    3.3146e-9,      17.561e-9;
    0.87477,    3.3525e-9,      17.767e-9;
    1.00000,    3.3904e-9,      17.973e-9; ];

  % pmos, nmos; (starting values)
  start = [
    23e-6, 7.9e-6;    % ion
    0.60,  0.52;      % vth
    1.30,  1.31;      % eta
    0.094, 0.100 ];   % lambda
  % Do regression
  params = fminsearch(@subth_err, start)

  %Plot results
  format short eng;
  legtext = { 'P Simulated', 'N Simulated', ...
              'P Regressed', 'N Regressed' };

  temp = 26.85; vgs = ids_vgs(:,1); vds = 0.5;
  figure(1); clf; subplot(1,2,1);
  semilogy(vgs*[1,1,1,1], [ids_vgs(:,2:3), subth]);
  grid on; legend(legtext,'Location','Best');
  xlabel('Vgs (V)'); ylabel('Ids (A)');

  subplot(1,2,2);
  semilogy(vgs*[1,1], rerr(ids_vgs));
  grid on; legend('P','N','Location','Best');
  xlabel('Vgs (V)'); ylabel('Rel error');

  temp = ids_temp(:,1); vgs = 0.25; vds = 0.5;
  figure(2); clf;
  subplot(1,2,1);
  semilogy(temp*[1,1,1,1],[ids_temp(:,2:3),subth]);
  grid on; legend(legtext,'Location','Best');
  xlabel('Temp (C)'); ylabel('Ids (A)');

  subplot(1,2,2);
  semilogy(temp*[1,1], rerr(ids_temp));
  grid on; legend('P','N','Location','Best');
  xlabel('Temp (C)'); ylabel('Rel error');

  temp = 26.85; vgs = 0.25; vds = ids_vds(:,1);
  figure(3); clf; subplot(1,2,1);
  plot(vds*[1,1,1,1], [ids_vds(:,2:3), subth]);
  grid on; legend(legtext,'Location','Best');
  xlabel('Vds (V)'); ylabel('Ids (A)');

  subplot(1,2,2);
  semilogy(vds*[1,1], rerr(ids_vds));
  grid on; legend('P','N','Location','Best');
  xlabel('Vds (V)'); ylabel('Rel error');
end

function fids = subth
  % Return pmos/nmos Id matrix
  k   = 1.3806505e-23;
  qe  = 1.60217653e-19;
  wid = 5e-6; len = 1e-6;

  global  temp; global vgs; global vds;
  templ = temp; vgsl = vgs; vdsl = vds;
  if (length(temp)~=1)
    ninputs=length(temp); templ=temp*[1,1]; end
  if (length (vgs)~=1)
    ninputs=length (vgs);  vgsl= vgs*[1,1]; end
  if (length (vds)~=1)
    ninputs=length (vds);  vdsl= vds*[1,1]; end
  mcol = ones(ninputs,1);

  global params;
  ion   = mcol*params(1,:);
  vth   = mcol*params(2,:);
  eta   = mcol*params(3,:);
  lambda = mcol*params(4,:);

  fids = wid/len.*ion.*exp( ...
        (vgsl-vth)*qe/k./eta./(templ+273.15) ...
     ) ./ (1 - lambda.*vdsl);
end

function rerr = rerr(idmatrix)
  rerr = (subth./idmatrix(:,2:3) - 1).^2;
end

function err = subth_err(regparams)
  global ids_temp; global ids_vgs; global ids_vds;
  global     temp; global     vgs; global     vds;
  global params;
  params = regparams;

  % Sum up errors from this round of regression
  vgs = 0.250; vds = 0.500; temp = ids_temp(:,1);
  err = sum(rerr(ids_temp));

  vgs = ids_vgs(:,1); vds = 0.500; temp = 26.85;
  err = err + sum(rerr(ids_vgs));

  vgs = 0.250; vds = ids_vds(:,1); temp = 26.85;
  err = sum(err*0.01 + sum(rerr(ids_vds)));
end
```

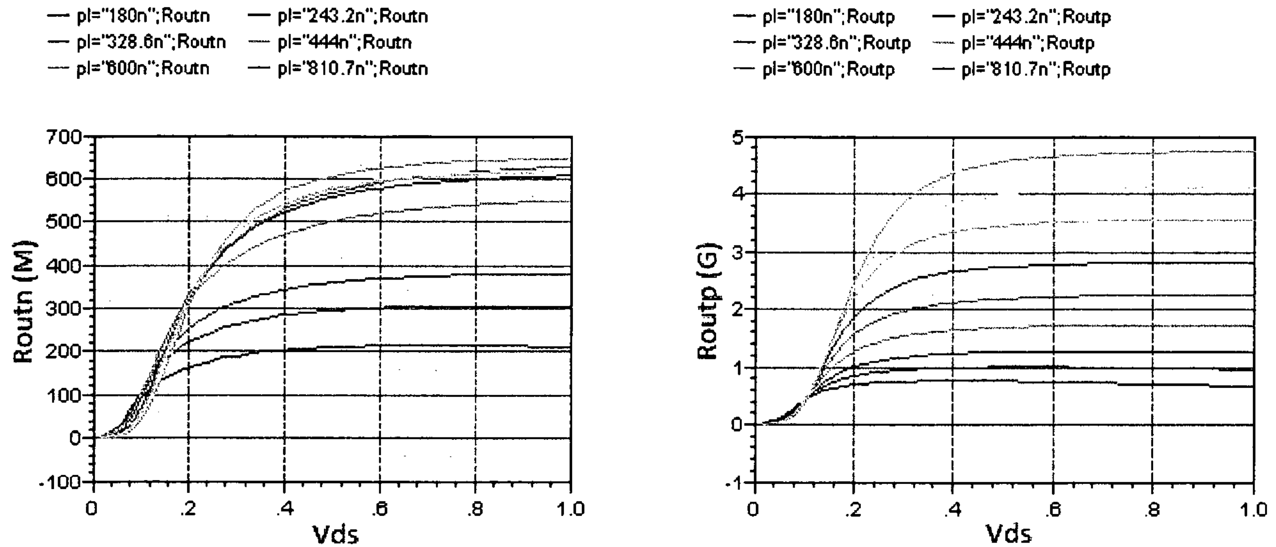


Figure A.5: Output resistance (NMOS and PMOS) versus V_{DS} for length from 180nm to 2μm.

The channel length modulation parameter λ is of course a function of the transistor dimensions. Figure A.5 shows the effect of varying length on the output resistance, where the output resistance is taken as the partial derivative of V_{DS} with respect to I_{DS}. The aspect ratio is held at 5. The large discrepancy between NMOS and PMOS resistances is due to there being a much smaller current, and much smaller current variation, through the PMOS transistor for a Vgs of 250 mV.

Traditional models (such as Shichman-Hodges) define the channel length modulation effect using λ, in turn determined by transistor length. An important difference of this model from the Shichman-Hodges model is that the channel length modulation term of (A.1) is a subtraction in the denominator (instead of an addition in the numerator). Attempting to apply the traditional channel length modulation model gives the illusion of λ being a function of both length and width. The selected model fits the behaviour of Cadence simulation much better because in the former case, output resistance is assumed to be independent of current, but in the latter case output resistance depends strongly on current. Output resistance can be found with (A.4) through (A.6):

$$I_{base}=\frac{W\,I_{ON}}{L}\exp\left(\frac{V_G-V_{TH}}{\eta V_T}\right) \tag{A.4}$$

$$I_{DS} \approx \frac{I_{base}}{1-\lambda V_{DS}} \tag{A.5}$$

$$\frac{\partial V_{DS}}{\partial I_{DS}} = \frac{\left(1-\lambda V_{DS}\right)^2}{\lambda I_{base}} \tag{A.6}$$

Based on the values of λ in Table A.1 obtained with a length of 1 μm, the following approximations may be made.

$$\lambda_P \approx \frac{\left(0.09662\,\mathrm{V}^{-1}\right)\left(1\,\mu m\right)}{L} \tag{A.7}$$

$$\lambda_N \approx \frac{\left(0.09989\,\mathrm{V}^{-1}\right)\left(1\,\mu m\right)}{L} \tag{A.8}$$

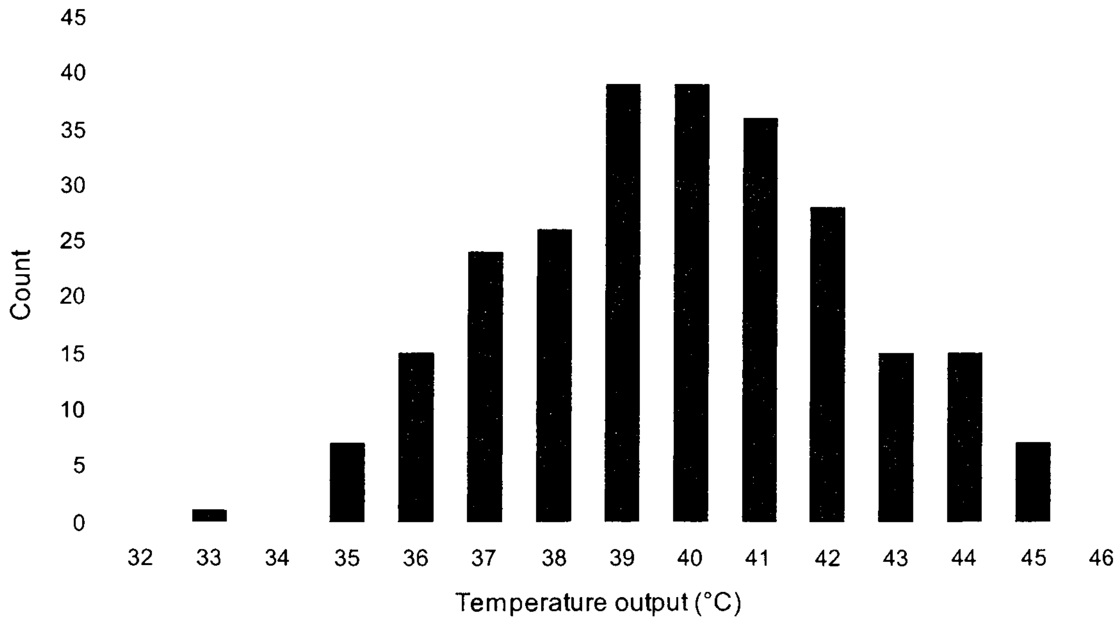
Appendix B. Standard deviation evaluation for the histogram of [5]

This is the derivation of a standard deviation for the histogram of circuit [5]. It uses frequency-weighted formulae for the mean and standard deviation.

T_i	n_i	$T_i n_i$	$(T_i-\mu)^2 n_i$
32	0	0	0
33	1	33	47.18
34	0	0	0
35	7	245	165.95
36	15	540	224.54
37	24	888	197.55
38	26	988	90.83
39	39	1521	29.45
40	39	1560	0.67
41	36	1476	46.05
42	28	1176	127.15
43	15	645	147.04
44	15	660	255.97
45	7	315	184.29
46	0	0	0
$\sum$	252	10047	1516.68

$$\mu=\frac{\sum T_i n_i}{\sum n_i} \qquad 39.87$$

$$\sigma=\sqrt{\frac{\sum (T_i-\mu)^2 n_i}{\sum n_i}} \qquad 2.45$$



Appendix C. Initial Dimension Calculations

The following initial calculations needed to be performed to acquire starting values for the dimensions of transistors. The final values shown in Table 4.3 reflect modifications to these dimensions based on simulation results. Refer to the schematics in Chapter 4 for the names of transistors and voltage nodes in the following discussion.

Assume that V_{DD} = 1V and T = 300°K. The thermal voltage at this temperature is given by (C.1):

$$V_{T(300^\circ K)} = \frac{\left(1.381\times10^{-23}\,J/^\circ K\right)\left(300^\circ K\right)}{1.602\times10^{-19}C} = 25.852\,mV \tag{C.1}$$

For this temperature, we will adopt a target current of 75 nA for Iout. This is a mid-range current that allows for good gain while still maintaining a relatively low power consumption. The bias currents must be significantly higher than the output current due to stability issues with power supply variation and Monte Carlo variation. For this reason, take 250 nA for Iref, 250 nA for Id1 and 300 nA for Id2. A 50 nA difference between Id1 and Id2 is reasonable; smaller values start to incur significant error in the subtraction circuit.

First consider the current mirrors. It is very important for the input half of a current mirror to have high output resistance, so all mirror input transistors (Mrb, Mrc, Mrd, Mm1b, Mm1d, Mm2b, Mm2d, Mob, Mod) have a length of 1 µm; any higher and these transistors would take up too much area. The width of these transistors must be high because these transistors are very vulnerable to Monte Carlo variations, so let their widths be 20 µm. Since the reference and

biasing currents must be relatively large, let there be a 1000:180 scaling factor on the mirror outputs (Mcpr, Mcqr, Mopp, Moqp, Mbpr, Mbqr, Mp2s, Mq2s, Mp1p, Mq1p, Mp2p, Mq2p, Mp1m, Mq1m, Mp2m, Mq2m, Mrpt, Mrqt, Mopf, Moqf, Mp1o, Mq1o, Mopo, Moqo) such that all of these transistors have a length of 1 μm. These transistors are also sensitive to Monte Carlo variation, so let their widths also be 20 μm.

For the same stability and output resistance reasons, let the cascode transistors (Mc, Mpca, Mpcb, Mbc1, Mbc2, Mc1s, Mca) be 20 μm / 1 μm. The cascode biasing transistors Mra and Mcd are more difficult to determine; they need to have much smaller aspect ratios to correctly bias their respective cascode sections.

Mrb and Mrd experience drain currents of Iref = 250 nA. Let the minimum drain-source voltage for Mrb and Mrd be 150 mV. Then Vd of Mrb = 1V - 150 mV = 850 mV and Vd of Mrd = 850 mV - 150 mV = 700 mV. Solving (A.1) for Vgs gives (C.2):

$$V_{GS}=V_{TH}+\eta V_T \ln\left(\frac{I_D L(1-\lambda V_{DS})}{I_{ON} W}\right) \tag{C.2}$$

$$\lambda_{(180\,\mathrm{nm})}=\frac{(0.09662\,\mathrm{V}^{-1})(1\,\mu\mathrm{m})}{180\,\mathrm{nm}}=0.5368\,\mathrm{V}^{-1} \tag{C.3}$$

$$V_{GS(\mathrm{Mrd})}=604.913\,\mathrm{mV}+1.305(25.852\,\mathrm{mV})\ln\left(\frac{(250\,\mathrm{nA})(180\,\mathrm{nm})(1-(0.5368\,\mathrm{V}^{-1})(150\,\mathrm{mV}))}{(22.838\,\mu\mathrm{A})(20\,\mu\mathrm{m})}\right)=290.850\,\mathrm{mV} \tag{C.4}$$

Mra and Mrc experience a ratio of 180/1000 of 250 nA, equal to 45 nA. (C.5) through (C.7) find the width of Mra:

$$V_{GS(\mathrm{Mra})}=150\,\mathrm{mV}+290.850\,\mathrm{mV}=440.850\,\mathrm{mV} \tag{C.5}$$

$$W=\frac{I_D L(1-\lambda V_{DS})}{I_{ON}}\exp\left(\frac{V_{TH}-V_{GS}}{\eta V_T}\right) \tag{C.6}$$

$$W_{(\mathrm{Mra})}=\frac{(45\,\mathrm{nA})(1\,\mu\mathrm{m})(1-(0.09662\,\mathrm{V}^{-1})(150\,\mathrm{mV}))}{22.838\,\mu\mathrm{A}}\exp\left(\frac{604.913\,\mathrm{mV}-440.850\,\mathrm{mV}}{1.305(25.852\,\mathrm{mV})}\right)=251.302\,\mathrm{nm} \tag{C.7}$$

For a constant current, as the width of Mra increases, the gate voltage decreases, forcing Mrb out of the active region; thus smaller widths allow for a safer operating range. In practice a slightly larger width is used to mitigate Monte Carlo variation issues.

Next we consider the biasing section. Mb has a drain current of 250 nA. Mb1 and Mb2 branches should have equal dimensions to correctly generate the two biasing currents. The voltage drop over Rb should not be too large (requiring a large resistance) or too small (generating an inadequate Id2-Id1); assume a voltage drop of 5mV and a resistor of 5mV / 250nA = 20kΩ. Since the branch of Mbpr/Mbqr/Mb only has three transistors, there is more room for biasing safety on the drain-source potential of Mb. Let Vds of Mb be 300mV, of course also equal to Vgs because it is diode-connected. Let the length be 1 μm for good output resistance. Then the width of Mb is found as follows:

$$W_{(Mb)}=\frac{(250\,\text{nA})(1\,\mu\text{m})\left(1-(0.09989\,\text{V}^{-1})(300\,\text{mV})\right)}{7.991\,\mu\text{A}}\exp\left(\frac{515.468\,\text{mV}-300\,\text{mV}}{1.315(25.852\,\text{mV})}\right)$$
$$=17.169\,\mu\text{m} \tag{C.8}$$

Let Mb1, Mb2 and Mb1s all be matched to Mb1; they are all cascode-protected.

Now consider the peaking current source. Rp and Mp1 receive a current of 75nA from the PTAT output. Rp should be at least ~1MΩ for peaking to occur; otherwise, the peak occurs at a current that is far too high for the range of this circuit. Such a resistance creates a voltage drop of 75mV. Similar to the biasing section, allot 300mV for the drain-source potential of Mp1. Let the length be 1 μm for adequate output resistance: the function of the peaking source relies on the drain voltage of Mp1. Since Vgs for Mp2a/Mp2b are close to that of Mp1, let Mp2a/b be matched to Mp1. The width is found by:

$$W_{(Mp1)}=\frac{(75\,\text{nA})(1\,\mu\text{m})\left(1-(0.09989\,\text{V}^{-1})(300\,\text{mV})\right)}{7.991\,\mu\text{A}}\exp\left(\frac{515.468\,\text{mV}-300\,\text{mV}}{1.315(25.852\,\text{mV})}\right)$$
$$=5.155\,\mu\text{m} \tag{C.9}$$

It should be noted that, after experimentation, this width had to be decreased substantially due to second-order effects resulting from the peaking topology moving the peak from its

expected value.

Mcpr and Mcqr mirror the reference current Iref of 250 nA to Mc and Mcd that create a cascode voltage reference, similar to Mra and Mrc. Mcd is similar to Mra in that it needs a small aspect ratio to correctly bias the cascode sections. To find the dimensions of this transistor, first examine the branch of Mp2a/Mpca. Let Mp2a and Mpca have a minimal drain-source potential of 150mV. The drain current through Mpca is 250nA*180/1000=45nA. Vgs is found by:

$$V_{GS\,(\mathrm{Mpca})} = 515.468\,\mathrm{mV} + 1.315(25.852\,\mathrm{mV})\ln\left(\frac{(45\,\mathrm{nA})(1\,\mu\mathrm{m})(1-(0.09989\,\mathrm{V}^{-1})(150\,\mathrm{mV}))}{(7.991\,\mu\mathrm{A})(20\,\mu\mathrm{m})}\right) = 237.038\,\mathrm{mV} \quad (C.10)$$

$$V_{GS\,(\mathrm{Mcd})} = 150\,\mathrm{mV} + 237.038\,\mathrm{mV} = 387.038\,\mathrm{mV} \quad (C.11)$$

$$W_{(\mathrm{Mcd})} = \frac{(250\,\mathrm{nA})(1\,\mu\mathrm{m})(1-(0.09989\,\mathrm{V}^{-1})(150\,\mathrm{mV}))}{7.991\,\mu\mathrm{A}}\exp\left(\frac{515.468\,\mathrm{mV} - 387.038\,\mathrm{mV}}{1.315(25.852\,\mathrm{mV})}\right) = 1.347\,\mu\mathrm{m} \quad (C.12)$$

The translinear block is perhaps the most complex and difficult part of the circuit for which initial calculations must be made. Due to the drain nodes of M2o/M4o/M6o affecting the biasing of two transistors each, it is more important than usual to have a very high output resistance. Let each of these three transistors have a length of 2μm. M1 through M6 have drain currents of 250nA, 300-250=50nA, 300nA, 300+250=550nA, 250nA and 75*180/1000=13.5nA, respectively. M2o, M4o and M6o have drain currents of 50+300=350nA, 550nA+250nA=800nA, and 13.5nA+250nA=263.5nA, respectively. The low current near the last stage of the circuit in the M6 branch is especially desirable due to the effect of current on output resistance (see (A.6)).

Start with M6, since it is diode-connected. Its drain current and length are now known. Let its drain-source potential be 300mV, as with other diode-connected transistors in the circuit. The width is found by

$$W_{(\mathrm{M6o})} = \frac{(263.5\,\mathrm{nA})(2\,\mu\mathrm{m})(1-(0.09989\,\mathrm{V}^{-1})(300\,\mathrm{mV}))}{7.991\,\mu\mathrm{A}}\exp\left(\frac{515.468\,\mathrm{mV} - 300\,\mathrm{mV}}{1.315(25.852\,\mathrm{mV})}\right) = 36.192\,\mu\mathrm{m} \quad (C.13)$$

Let M2o, M4o and M6o be matched for simplicity Also let M1 through M6 be matched with each other (but not with M2o/M4o/M6o). The dimensions for M1 through M6 must be selected such that the most stringent restriction on biasing is met. M1, M3 and M5 are not an issue due to being diode-connected, so M2, M4 and M6 must have biasing verified. Examine the gate voltages at M2o, M4o and M6o (neglecting the channel length modulation effect due to high length):

$$V_{GS(\mathrm{M2o})} = 515.468\,\mathrm{mV} + 1.315(25.852\,\mathrm{mV})\ln\left(\frac{(350\,\mathrm{nA})(2\,\mu\mathrm{m})}{(7.991\,\mu\mathrm{A})(36.192\,\mu\mathrm{m})}\right)$$
$$= 310.686\,\mathrm{mV} \qquad \text{(C.14)}$$

$$V_{GS(\mathrm{M4o})} = 515.468\,\mathrm{mV} + 1.315(25.852\,\mathrm{mV})\ln\left(\frac{(800\,\mathrm{nA})(2\,\mu\mathrm{m})}{(7.991\,\mu\mathrm{A})(36.192\,\mu\mathrm{m})}\right)$$
$$= 338.789\,\mathrm{mV} \qquad \text{(C.15)}$$

$$V_{GS(\mathrm{M6o})} = 300\,\mathrm{mV} \qquad \text{(C.16)}$$

It would be productive to impose a minimal 150 mV of drain-source potential for M2, M4 and M6. One would base the bias restriction on M2 (or M4) because its drain node is biased by the translinear cell.

$$V_{GS(\mathrm{M2o})} - \left(300\,\mathrm{mV} + V_{GS(\mathrm{M1})} - V_{GS(\mathrm{M2})}\right) = 150\,\mathrm{mV} \qquad \text{(C.17)}$$

Interestingly, in subtracting the gate voltages in (C.17) the width and length are cancelled:

$$V_{GS(\mathrm{M2o})} - \left(300\,\mathrm{mV} + \eta V_T \ln\left(\frac{I_{D(\mathrm{M1})}}{I_{D(\mathrm{M2})}}\right)\right) = 150\,\mathrm{mV} \qquad \text{(C.18)}$$

As long as the dimensions over transistors M1 through M6 are matched, the translinear cell should operate in theory.

References

[1] S. Sharifi, L. ChunChen, T.S. Rosing. "Accurate Temperature Estimation for Efficient Thermal Management". *9th International Symposium on Quality Electronic Design*, pp. 137-142, Mar 2008.

[2] Q. Chen, M. Meterelliyoz, K. Roy. "A CMOS Thermal Sensor and its Applications in Temperature Adaptive Design". *Proc. 7th International Symposium on Quality Electronic Design*, pp. 242-248, Mar 2006.

[3] B.H. Calhoun and A.P. Chandrakasan. "Ultra-Dynamic Voltage Scaling (UDVS) Using Sub-Threshold Operation and Local Voltage Dithering". *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 238-245, Jan 2006.

[4] D.E. Duarte, G. Geannopoulos, U. Mughal, et al.. "Temperature Sensor Design in a High Volume Manufacturing 65nm CMOS Digital Process". *IEEE Custom Integrated Circuits Conference*, pp. 221-224, Sept 2007.

[5] A. Syal, V. Lee, A. Ivanov, et al.. "CMOS Differential and Absolute Thermal Sensors". *Journal of Electronic Testing*, vol. 18, pp. 295-304, June 2002.

[6] J. Altet, A. Rubio, S. Dilhaire et al.. "BiCMOS thermal sensor circuit for built-in test purposes". *IET Journal of Electronics Letters*, vol. 34, pp. 1307-1309, June 1998.

[7] G. Meijer, G. Wang, F. Fruett. "Temperature Sensors and Voltage References Implemented in CMOS Technology". *IEEE Sensors Journal*, vol. 1, pp. 225-234, Oct 2001.

[8] B. Li, P. Lai, J. Sin et al.. "Ultra-low-power thermal sensor with silicon-on-insulator (SOI) structure for high-temperature applications". *Proc. IEEE Conference on Sensors*, pp. 1419-1422, Nov 2002.

[9] F. Serra-Graells, J. Huertas. "Sub-1-V CMOS Proportional-to-Absolute Temperature References". *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 84-88, Jan 2003.

[10] M. Danaie, R. Lotfi. "A low-voltage high-PSRR CMOS PTAT & constant-gm reference circuit". *Proc. 48th Midwest Symposium on Circuits and Systems*, pp. 1807-1810, Aug 2005.

[11] H. Lee, C. Hsu, C. Luo. "CMOS thermal sensing system with simplified circuits and high accuracy for biomedical application". *Proc. IEEE International Symposium on Circuits and Systems*, pp. 4367-4370, Sept 2006.

[12] K. Ueno, T. Hirose, T. Asai, et al.. "Ultralow-Power Smart Temperature Sensors with Subthreshold CMOS Circuits". *Proc. International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 546-549, Dec 2006.

[13] J. Tsai, H. Chiueh. "High Linear Voltage References for on-chip CMOS Temperature Sensor". *Proc. 13th IEEE International Conference on Electronics, Circuits and Systems*, pp.

216-219, Dec 2006.

[14] S. Yan, E. Sanchez-Sinencio. "Low Voltage Analog Circuit Design Techniques: A Tutorial". *IEICE Trans. Analog Integrated Circuits and Systems*, vol. E00-A, pp. 6-7, Feb 2000.

[15] D. Kerns. "Optimization of the peaking current source". *IEEE Journal of Solid-State Circuits*, vol. 21, pp. 587-590, Aug 1986.

[16] D. Kerns. "Enhanced peaking current reference". *IEEE Journal of Solid-State Circuits*, vol. 23, pp. 869-872, June 1988.

[17] C. Kwok. "Low-Voltage Peaking Complementary Current Generator". *IEEE Journal of Solid-State Circuits*, vol. 20, pp. 816-818, June 1985.

[18] P. Gray, P. Hurst, S. Lewis, et al.. *Analysis and Design of Analog Integrated Circuits*, Fourth Ed., John Wiley & Sons, Inc., pp. 303-305, 2001.

[19] C.A. Papazoglu and C.A. Karybakas. "Electronically tunable floating CMOS resistor independent of the MOS parameters and temperature". *The 6th IEEE International Conference on Electronics, Circuits and Systems*, pp. 311-314, Sept 1999.

[20] P. Gray, P. Hurst, S. Lewis, et al.. *Analysis and Design of Analog Integrated Circuits*, Fourth Ed., John Wiley & Sons, Inc., pp. 273-274, 2001.

[21] A. Tajalli, Y. Leblebici, E. Brauer. "Implementing ultra-high-value floating tunable CMOS resistors". *IET Eletronics Letters*, vol. 44, pp. 349-350, Feb 2008.