



UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI INGEGNERIA

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA ELETTRONICA

**Analysis and Laboratory Verification of Bandgap Prototypes,  
Circuit Engineering, Optimization of Trimming Process**

RELATORE:

Prof. Gerosa Andrea

I CORRELATORE:

Ing. Piselli Marco

II CORRELATORE:

Dott. Zampieri Stefano

LAUREANDO:

Matteo Rigo 1012941

Padova, 16 Luglio 2013

Anno Accademico 2012/2013

*To my parents, for their encouragement, love and support.*



# Contents

<b>1</b>	<b>VOLTAGE REFERENCE</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Zero-order reference . . . . .	1
1.2.1	Forward-biased diode references . . . . .	1
1.2.2	Zener references . . . . .	2
1.3	Bandgap reference . . . . .	2
1.3.1	First-order reference . . . . .	2
1.3.2	Second-order reference . . . . .	3
1.3.3	The nonlinear reference: Diode Loop . . . . .	4
1.4	The diode loop spt6 . . . . .	5
<b>2</b>	<b>FIRST ORDER VOLTAGE REFERENCE</b>	<b>9</b>
2.1	First-order compensation . . . . .	9
2.2	Evaluation of $\eta$ and $V_{GO}$ . . . . .	13
2.3	PTAT stage . . . . .	18
2.3.1	Ideal schematic . . . . .	18
2.3.2	Real PTAT stage with BJT: first architecture . . . . .	19
2.3.3	Real PTAT stage with BJT: second architecture . . . . .	22
2.3.4	Real PTAT stage with BJT: third architecture . . . . .	25
2.3.5	Real PTAT stage with MOSFET . . . . .	30
2.3.6	The source degeneration . . . . .	31
2.3.7	Dominant pole compensation of PTAT stage . . . . .	34
2.4	CTAT stage . . . . .	37
2.4.1	Ideal schematic . . . . .	37
2.4.2	Real schematic . . . . .	39
2.4.3	Dominant pole compensation of CTAT loop . . . . .	40
2.5	First-order compensated bandgap output voltage . . . . .	43
2.5.1	Evaluation of output voltage . . . . .	43
2.5.2	Transient response with $V_{DD}$ steps . . . . .	44
<b>3</b>	<b>DIODE LOOP SUPPLY1</b>	<b>47</b>
3.1	Curvature correction . . . . .	47
3.2	Ideal Diode Loop . . . . .	48
3.3	Evaluation of current mismatch . . . . .	51
3.3.1	Output currents mirrored . . . . .	51

3.3.2	Currents mirrored to generate $I_{const}$	52
3.4	Improving current mirrors	55
3.5	Effects of base currents	61
3.6	Stabilization of non-linear order bandgap reference: dominant pole compensation	63
3.6.1	Transient response with $V_{DD}$ steps	67
<b>4</b>	<b>TRIMMING NETWORKS</b>	<b>69</b>
4.1	Analysis of Process-Induced Errors	69
4.1.1	Current mirror mismatch	69
4.1.2	Transistor mismatch	70
4.1.3	Early Voltage	71
4.1.4	Resistor process dependency	71
4.1.5	Resistors temperature coefficient	72
4.1.6	Package-shift effects	72
4.1.7	Threshold Voltage Mismatch	72
4.1.8	Current Factor Mismatch	73
4.1.9	Solutions implemented against mismatch effects	73
4.1.10	Monte Carlo Analysis	74
4.1.11	Corner Analysis	77
4.2	Trimming Techniques	79
4.2.1	Zener zap	79
4.2.2	Fusible links	79
4.2.3	Laser-trimmable resistors	79
4.3	Trimming options	80
4.3.1	Description of the Zener zap mechanism	80
4.3.2	First trimming option	82
4.3.3	Second trimming option	82
4.3.4	Third trimming option	83
4.3.5	Fourth trimming option	84
4.4	Trimming Resistances	84
4.4.1	Trimming $R_{NL}$	86
4.4.2	Trimming $R_{LOAD}$	90
<b>5</b>	<b>TRIMMING VERIFICATION</b>	<b>95</b>
5.1	Residual curvature of nominal run	95
5.2	Residual curvature of Monte Carlo Analysis	97
5.3	Residual curvature of Corner Analysis	99
5.4	Trimming Algorithm	101
5.5	Bandgap output voltage after double trimming	105
5.5.1	First trimming option	105
5.5.2	Second trimming option	107
5.5.3	Comparison of trimming options	109
5.6	Residual curvature after double trimming	110
<b>6</b>	<b>CONCLUSIONS</b>	<b>113</b>

# List of Figures

1.1	Diode voltage reference . . . . .	2
1.2	Zener voltage reference . . . . .	2
1.3	First-order bandgap reference . . . . .	3
1.4	First-order bandgap output voltage $V_{BG}$ . . . . .	3
1.5	Squared PTAT curvature-correction . . . . .	4
1.6	Squared PTAT curvature-correction bandgap output voltage . . . . .	4
1.7	Nonlinear curvature-correction reference . . . . .	5
1.8	Nonlinear curvature-correction bandgap output voltage . . . . .	5
1.9	Diode Loop of spt6 . . . . .	6
1.10	Simulation of bandgap output voltage references [16] . . . . .	6
1.11	Experimental points of a Diode Loop . . . . .	7
2.1	Ideal first-order compensation . . . . .	10
2.2	PTAT current generator . . . . .	11
2.3	CTAT current generator . . . . .	12
2.4	Theoretical circuit to evaluate $V_{G0}$ and $\eta$ . . . . .	15
2.5	Base current for 2.4 . . . . .	15
2.6	$\ln(I_X - I_{BASE})$ (magenta) and $\ln(I_X)$ (blu) . . . . .	16
2.7	Circuit used to get “experimental” values with Cadence . . . . .	16
2.8	$V_{BE}$ for a temperature-independent $I_C$ . . . . .	17
2.9	$V_{BE}$ for a PTAT $I_C$ . . . . .	17
2.10	PTAT stage: ideal schematic . . . . .	18
2.11	First PTAT stage with feedback loop . . . . .	19
2.12	PTAT current decreases at high temperatures . . . . .	20
2.13	$V_{BE}$ decreases at high temperatures . . . . .	20
2.14	Degeneration with resistance of emitter of T4 to avoid low drop voltage at high temperatures . . . . .	21
2.15	Mismatch collector current ( $I_{CT3} - I_{CT2}$ ) after balancing base currents . . . . .	21
2.16	Montecarlo Analysis of the first architecture PTAT stage $\bar{\mu} = 36.23V \times 10^{-2}$ $\sigma = 2.21 \times 10^{-2}$ . . . . .	22
2.17	Second PTAT stage with a different feedback loop . . . . .	23
2.18	Collector currents mismatch ( $I_{CT3} - I_{CT2}$ ) after base current compensation . . . . .	24
2.19	Montecarlo Analysis of the second architecture PTAT stage $\bar{\mu} = 36.71V \times 10^{-2}$ $\sigma = 2.06 \times 10^{-2}$ . . . . .	24
2.20	Comparison of different PTAT voltage . . . . .	25
2.21	Comparison of the derivate of PTAT voltage . . . . .	25
2.22	Third solution . . . . .	26
2.23	Mismatch collector currents T1-T2 . . . . .	27

2.24	Comparison of different PTAT voltage . . . . .	28
2.25	Comparison of the derivate of PTAT voltage . . . . .	28
2.26	Third solution to avoid Early Effect . . . . .	29
2.27	Mismatch collector currents T1-T2 will reduce . . . . .	29
2.28	Mismatch base currents after optimization of feedback transistors . . . . .	30
2.29	PTAT stage realized with mosfet . . . . .	30
2.30	Mismatch collector currents $I_{C_{M3}}$ and $I_{C_{M2}}$ . . . . .	31
2.31	Mismatch drain currents $I_{D_{T1}}$ and $I_{D_{T0}}$ . . . . .	31
2.32	Current mirror with source degeneration . . . . .	32
2.33	Current gain $\frac{I_{out}}{I_{in}}$ of fig. 2.32 if $R_e = 0$ . . . . .	33
2.34	Current gain $\frac{I_{out}}{I_{in}}$ of fig. 2.32 if $R_e \neq 0$ . . . . .	33
2.35	Gain margin and phase margin . . . . .	34
2.36	Gain loop of PTAT stage before compensation . . . . .	35
2.37	PTAT loop gain after dominant pole compensation . . . . .	36
2.38	Phase margin and Gain margin of PTAT loop at $-40^\circ\text{C}$ . . . . .	36
2.39	Phase margin and Gain margin of PTAT loop at $27^\circ\text{C}$ . . . . .	37
2.40	Phase margin and Gain margin of PTAT loop at $160^\circ\text{C}$ . . . . .	37
2.41	CTAT current generator . . . . .	38
2.42	Ideal CTAT current generator with feedback loop . . . . .	38
2.43	Ideal CTAT stage . . . . .	39
2.44	Real CTAT stage . . . . .	39
2.45	Current mismatch between M7 and M6 . . . . .	40
2.46	Current mismatch between M4 and M8 . . . . .	40
2.47	Real CTAT stage . . . . .	41
2.48	CTAT loop gain without dominant pole compensation . . . . .	41
2.49	CTAT loop gain with dominant pole compensation . . . . .	42
2.50	Phase margin and Gain margin of CTAT feedback loop at $-40^\circ\text{C}$ . . . . .	42
2.51	Phase margin and Gain margin of CTAT feedback loop at $27^\circ\text{C}$ . . . . .	43
2.52	Phase margin and Gain margin of CTAT feedback loop at $160^\circ\text{C}$ . . . . .	43
2.53	$V_{PTAT}$ , $V_{CTAT}$ and $V_{BG}$ . . . . .	44
2.54	Step response of the first-order compensated voltage bandgap generator at $T = -40^\circ\text{C}$ . . . . .	44
2.55	Step response of the first-order compensated voltage bandgap generator at $T = 27^\circ\text{C}$ . . . . .	45
2.56	Step response of the first-order compensated voltage bandgap generator at $T = 160^\circ\text{C}$ . . . . .	45
3.1	Diode-loop curvature ideal topology . . . . .	48
3.2	Ideal Diode Loop . . . . .	49
3.3	Output voltage . . . . .	50
3.4	Non linear voltage $V_{NL}$ across $R_{NL}$ . . . . .	50
3.5	Changing $R_{NL}$ output voltage “rotates” around $T = T_R$ . . . . .	51
3.6	Evaluation of real mirror ratio in $I_{OUT} = I_{PTAT} + I_{CTAT} + I_{NL}$ generation . . . . .	51
3.7	Mismatch between CTAT collector current of M14 and CTAT collector current of M6 (ideal $\frac{I_{C_{M14}}}{I_{C_{M6}}} = 1$ ) . . . . .	52
3.8	Mismatch between PTAT collector current of M13 and PTAT collector current of M3 (ideal $\frac{I_{C_{M13}}}{I_{C_{M3}}} = 1$ ) . . . . .	52
3.9	Wrong compensation temperature . . . . .	53

3.10	Evaluation of real mirror ratio in $I_{const}$ generation	53
3.11	Mismatch between collector current of M10 and collector current of M6 (ideal $\frac{I_{CM10}}{I_{CM6}} = 1$ )	54
3.12	Mismatch between collector of M11 and collector current of M3 (ideal $\frac{I_{CM11}}{I_{CM3}} = 0.65$ )	54
3.13	Mismatch between collector of M12 and collector current of M6 (ideal $\frac{I_{CM12}}{I_{CM6}} = 1$ )	54
3.14	Circuit used to evaluate $I_{DS} = f(V_{GS}, V_{DS}, \lambda)$	55
3.15	$I_{DS} = f(V_{GS}, V_{DS}, \lambda)$	56
3.16	Rectification of $I_{DS} = f(V_{GS}, V_{DS}, \lambda)$ with a trendline	56
3.17	Cascode	57
3.18	Bias for cascode	57
3.19	Nonlinear-order without base currents	58
3.20	Output bandgap voltage	59
3.21	Mismatch between collector current of M10 and collector current of M6 (ideal $\frac{I_{CM10}}{I_{CM6}} = 1$ )	59
3.22	Mismatch between collector current of M11 and collector current of M3 (ideal $\frac{I_{CM11}}{I_{CM3}} = 0.66$ )	59
3.23	Mismatch between collector current of M12 and collector current of M6 (ideal $\frac{I_{CM12}}{I_{CM6}} = 1$ )	60
3.24	Mismatch between the PTAT collector current of M14 and the PTAT collector current of M6 (ideal $\frac{I_{CM14}}{I_{CM6}} = 1$ )	60
3.25	Mismatch between the CTAT collector current of M13 and the CTAT collector current of M3 (ideal $\frac{I_{CM13}}{I_{CM3}} = 1$ )	60
3.26	Real second-order compensated bandgap voltage reference	61
3.27	Output voltage Non linear voltage $V_{NL}$ across $R_{NL}$	61
3.28	Base current of T0	63
3.29	Base current of T3	63
3.30	PTAT loop gain after dominant pole compensation	64
3.31	CTAT loop gain before dominant pole compensation	64
3.32	Phase margin of CTAT loop using $C_{comp} = 1pF$ and $R_{comp} = 66K\Omega$	65
3.33	CTAT loop gain after dominant pole compensation	65
3.34	Phase margin and Gain margin of CTAT feedback loop at $-40C$	66
3.35	Phase margin and Gain margin of CTAT feedback loop at $27^{\circ}C$	66
3.36	Phase margin and Gain margin of CTAT feedback loop at $160^{\circ}C$	67
3.37	Step response of the first-order compensated voltage bandgap generator at $T = -40^{\circ}C$	67
3.38	Step response of the first-order compensated voltage bandgap generator at $T = 27^{\circ}C$	68
3.39	Step response of the first-order compensated voltage bandgap generator at $T = 160^{\circ}C$	68
4.1	Monte Carlo Analysis of bandgap reference output voltage	74
4.2	Amplitude spread's distribution of bandgap output voltage $V_{BG}(T = 27^{\circ}C)$ -800mV $\mu_D = 2.8383 \times 10^{-4}$ $\sigma_D = 5.731 \times 10^{-3}$	75
4.3	Solpe of bandgap output voltage (nominal run)	76
4.4	Slope distribution of bandgap output voltage $V_{BG}(T = 100^{\circ}C) - V_{BG}(T = -40^{\circ}C)$ $\mu_S = 1.311 \times 10^{-4}V$ $\sigma_S = 2.192 \times 10^{-3}$	76
4.5	Corner Analysis of bandgap output voltage	77
4.6	Amplitude spread distribution of bandgap output voltage $V_{BG}(T = 27^{\circ}C)$ -800mV	78
4.7	Slopes distribution of bandgap output voltage $V_{BG}(T = 100^{\circ}C) - V_{BG}(T = -40^{\circ}C)$	78
4.8	Complete Diode Loop	80



4.9	Bipolar Zener Device . . . . .	81
4.10	Statistical distribution of $T = T_R$ . . . . .	82
4.11	Switches matrix used to trim the current which generates $I_{PTAT}^*$ . . . . .	83
4.12	Changing $R_{NL}$ , the bandgap output voltage will rotate around $T = T_R$ . The yellow line is the nominal output voltage ( $R_{NL} = 134K\Omega$ and $R_{LOAD} = 344K\Omega$ ) . . . . .	83
4.13	Distributions of $I_{PTAT}^*/I_{PTAT}$ ratio at $T = 27^\circ C$ $R_{NL1} = 130k\Omega$ $R_{NL2} = 134k\Omega$ $R_{NL3} = 138k\Omega$ . . .	84
4.14	Distributions of $I_{CTAT}^*/I_{CTAT}$ ratio at $T = 27^\circ C$ $R_{NL1} = 130k\Omega$ $R_{NL2} = 134k\Omega$ $R_{NL3} = 138k\Omega$ . . .	84
4.15	Initial spread of bandgap output voltage's slope $V_{BG}(T = 100^\circ C) - V_{BG}(T = -40^\circ C)$ $\bar{\mu}_S = 1.311 \times 10^{-4}V$ $\sigma_S = 2.192 \times 10^{-3}$ . . . . .	85
4.16	Output voltage's distribution $V_{BG}(T = 27^\circ C) - 800mV$ $\bar{\mu}_D = 2.8383 \times 10^{-4}$ $\sigma_D = 5.731 \times 10^{-3}$ . .	85
4.17	Cascaded trimming elements . . . . .	86
4.18	$R_{NL}^* = f(\Delta V, T_1, T_2)$ . . . . .	88
4.19	Comparison between $R_{NL}^* = f(\Delta V, T_1, T_2)$ and Cadence values . . . . .	88
4.20	Trimming network with 8 trimming words $R = 52.5K\Omega$ , $\Delta R_1 = 11K\Omega$ , $\Delta R_2 = 22K\Omega$ , $\Delta R_3 = 44K\Omega$	89
4.21	Trimming network with 4 trimming words $R = 52.5K\Omega$ , $\Delta R_1 = 26K\Omega$ , $\Delta R_2 = 52K\Omega$ . . . . .	89
4.22	$R_{LOAD}^* = f(R_{LOAD}, \Delta V_{BG}, T)$ . . . . .	91
4.23	Comparison between $R_{LOAD}^* = f(R_{LOAD}, \Delta V_{BG}, T)$ and Cadence's values . . . . .	91
4.24	Trimming network with 8 trimming words $R = 52.5K\Omega$ , $\Delta R_1 = 2.5K\Omega$ , $\Delta R_2 = 5K\Omega$ , $\Delta R_3 = 10K\Omega$	92
4.25	Trimming network with 16 trimming words $R = 52.5K\Omega$ , $\Delta R_1 = 1.1K\Omega$ , $\Delta R_2 = 2.2K\Omega$ , $\Delta R_3 = 4.4K\Omega$ , $\Delta R_4 = 8.8K\Omega$ . . . . .	93
5.1	The nominal bandgap output voltage $R_{PTAT} = 60k\Omega$ , $R_{CTAT} = 470k\Omega$ , $R_{NL} = 134k\Omega$ , $R_{LOAD} = 344k\Omega$	95
5.2	Fitting bandgap output voltage with $L = mx + q$ . . . . .	96
5.3	Residual curvature for nominal run of nonlinear-order compensated bandgap voltage reference . .	96
5.4	Comparison between residual curvatures of the nonlinear-order compensated voltage reference and a first-order compensated voltage reference . . . . .	97
5.5	Evaluation of the residual curvature of a first-order using Monte Carlo Analysis $\mu = 6.03 \times 10^{-3}V$ $var = 4.22 \times 10^{-8}$ . . . . .	98
5.6	Evaluation of the residual curvature of the nonlinear-order using Monte Carlo Analysis $\mu = 1.28 \times 10^{-4}V$ $var = 3.46 \times 10^{-9}$ . . . . .	98
5.7	Evaluation of the residual curvature of another nonlinear-order using Monte Carlo Analysis $\mu = 1.28 \times 10^{-4}V$ $var = 3.46 \times 10^{-9}$ . . . . .	99
5.8	Evaluation of residual curvature of a first-order using Corner Analysis . . . . .	100
5.9	Evaluation of residual curvature of nonlinear-order using Corner Analysis . . . . .	100
5.10	First step of trimming verification algorithm . . . . .	102
5.11	Second step of trimming verification algorithm . . . . .	103
5.12	Third step of trimming verification algorithm . . . . .	104
5.13	Verification of Monte Carlo seed . . . . .	105
5.14	Untrimmed statistical distribution of bandgap output voltage's slope $\Delta_S = V_{BG}(T = 100^\circ C) - V_{BG}(T = -40^\circ C)$ $\bar{\mu}_S = 1.311 \times 10^{-4}V$ $\sigma_S = 2.192 \times 10^{-3}$ . . . . .	105
5.15	Untrimmed statistical distribution of bandgap output voltage's amplitude $\Delta_A = V_{BG}(T = 27^\circ C) - 800mV$ $\bar{\mu}_A = 2.8383 \times 10^{-4}V$ $\sigma_A = 5.731 \times 10^{-3}$ . . . . .	106
5.16	Bandgap output voltage after double trimming (first trimmig option) . . . . .	106
5.17	Statistical distribution of $\Delta_S = V_{BG}(T = 100^\circ C) - V_{BG}(T = -40^\circ C)$ after double trimming $\bar{\mu}_S = -3.067 \times 10^{-5}V$ (first trimming option) . . . . .	107

5.18	Statistical distribution of $\Delta_A = V_{BG}(T = 27C) - 800mV$ ater double trimming $\bar{\mu}_A = -2.677 \times 10^{-4}V$ (first trimming option) . . . . .	107
5.19	Bandgap output voltage after double trimming (second trimmig option) . . . . .	108
5.20	Statistical distribution of $\Delta_S = V_{BG}(T = 100^\circ C) - V_{BG}(T = -40^\circ C)$ after double trimming $\bar{\mu}_S = -4.929 \times 10^{-6}V$ (second trimming option) . . . . .	108
5.21	Statistical distribution of $\Delta_A = V_{BG}(T = 27C) - 800mV$ ater double trimming $\bar{\mu}_A = -6.576 \times 10^{-6}V$ (second trimming option) . . . . .	109
5.22	Statistical distributions for first trimming option of $\Delta_{A_1} = V_{BG}(T_1 = -40C) - 800mV$ $\bar{\mu}_{A_1} = -2.190 \times 10^{-4}V$ (left) and of $\Delta_{A_2} = V_{BG}(T_2 = 100^\circ C) - 800mV$ $\bar{\mu}_{A_2} = -2.496 \times 10^{-4}V$ (right) . . . . .	109
5.23	Statistical distributions for first trimming option of $\Delta_{A_1} = V_{BG}(T_1 = -40C) - 800mV$ $\bar{\mu}_{A_1} = 2.450 \times 10^{-5}V$ (left) and of $\Delta_{A_2} = V_{BG}(T_2 = 100^\circ C) - 800mV$ $\bar{\mu}_{A_2} = 1.809 \times 10^{-5}V$ (right) . . . . .	110
5.24	Residual curvature of first-oder voltage reference: before trimming and after trimming . . . . .	110
5.25	Residual curvature of nonlinear-oder bandgap voltage reference (first trimming option): before trimming (left) and after trimming (right) . . . . .	111
5.26	Residual curvature of nonlinear-oder bandgap voltage reference (second trimming option): before trimming (left) and after trimming (right) . . . . .	111



# Sommario

Lo scopo di questa tesi è l'analisi e la progettazione di due riferimenti di tensione bandgap ad alta precisione e basso consumo, utilizzando una tecnologia economica. Il punto di partenza è lo studio di riferimenti di tensione simili, già progettati, implementati e testati. Durante la fase di progettazione ogni singolo stadio viene analizzato, ottimizzato e confrontando con altre possibili soluzioni, con lo scopo di ottenere la soluzione migliore. Vengono inoltre esaminati anche gli effetti del processo e del mismatch, per valutare la robustezza dei riferimenti di tensione progettati. Per migliorare la loro precisione vengono inoltre progettate due reti di trimming, la cui verifica viene ottenuta mediante un nuovo algoritmo.



# Chapter 1

## VOLTAGE REFERENCE

### 1.1 Introduction

Voltage references are essential to the accuracy and performance of analog systems. They are used in a lot of analog circuits for signal processing, such as, analog to digital or digital to analog converters, smart sensors and linear regulators. Of all the types of references, only bandgap references are suited to operate at a very low supply voltages. They can be used in order to have a precisely regulated supply. In addition, voltage references are needed in the design of products which must be accurate such as: voltmeters, ohmmeters and ammeters. Diodes, current mirrors and current references comprise the necessary building blocks used in the synthesis of the most voltage reference topologies.

A voltage reference can be categorized into different performance levels: zero-order, first-order, second order and nonlinear order.

The zero-reference is the most simple: this type of reference is typically not temperature-compensated. In other words there is no effort on the part of the designer to improve the precision of the given voltage, which could be derived from a Zener or a forward-biased diode.

First-order references instead are temperature-compensated at the first-order term of the polynomial relationship with respect to temperature. In order to compensate the first-order term (or higher-order terms) a Taylor-series expansion of the voltage can be useful.

Second-order references among compensating the first-order terms attempt to cancel approximately the nonlinear component of the temperature-dependent voltage, leaving only third-order and higher-order components.

### 1.2 Zero-order reference

#### 1.2.1 Forward-biased diode references

The most basic method to generate a voltage reference is to force a current through a  $p-n$  junction. It can be also replaced by a junction field-effect transistor (JFET) in order to optimize area and current overhead.

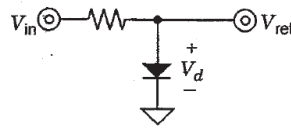


Figure 1.1: Diode voltage reference

Forward-biased diode references are very simple structures, however, their accuracy performance is degraded if the input voltage changes. The biasing current of the diode, in fact, changes with input voltage variations: in this way, also the output voltage changes.

### 1.2.2 Zener references

Another simple and cheap voltage reference is the zener diode reference. It is a diode, which operates in the reverse-bias region, where current begins to flow at a set voltage and increases dramatically as the voltage increases. If the current is forced to flow into the cathode, the diode goes into the reverse-breakdown region. In order to use it as a reference, a constant current has to be provided. This is achieved with a resistor from a higher supply voltage.

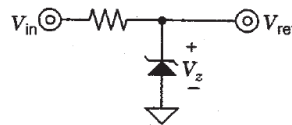


Figure 1.2: Zener voltage reference

One feature of zener diodes is that in the operating region of 6V, the zener becomes very stiff against current changes and simultaneously achieves a zero temperature coefficient. Zener references are somewhat difficult to use: the voltage tolerance is poor, they are noisy, the zener voltage depends on current and temperature. Moreover, in a lot of low voltage application, Zener Diodes can not be used because of their high operating voltages.

## 1.3 Bandgap reference

The other popular voltage reference is the bandgap reference. Bandgap references are more accurate and more suitable for low voltage operation than Zener diodes. A bandgap voltage reference is a temperature independent voltage reference circuit widely used in integrated circuits. Usually it has an output voltage around 1.2 V, close to the theoretical bandgap voltage of silicon at 0K. This reference involves the creation of a voltage with a positive temperature coefficient and of a voltage with negative temperature coefficient. The voltages have opposite temperature coefficients, so when they are summed together, the resulting voltage has a zero temperature coefficient. This concept was first published by David Hiber in 1964, Paul Brokaw [1] and Bob Widlar [14].

### 1.3.1 First-order reference

The first-order compensation is obtained by summing a proportional-to-absolute-temperature (PTAT) voltage and a complementary-to-absolute-temperature (CTAT) voltage. PTAT voltage has a positive temperature coefficient,

while CTAT voltage has a negative temperature coefficient. PTAT voltage increases linearly with temperature: in this way it can efficiently cancel the effects of the negative linear temperature dependence of the CTAT voltage.

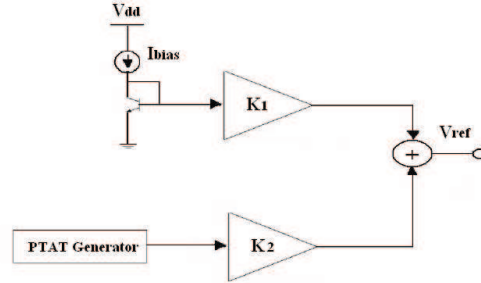


Figure 1.3: First-order bandgap reference

CTAT voltage is obtained by forcing a current through a base-emitter junction (diode voltage). Base-emitter voltage can be expressed using Tsividis' equation [12]:

$$V_{BE} = V_{G0} - \frac{V_{G0} - V_{BE}(T_R)}{T_R} T - (\eta - x) \frac{k_B}{q} T \ln\left(\frac{T}{T_R}\right) \quad (1.1)$$

$V_{G0}$  is the bandgap voltage at 0°K,  $T$  is the absolute temperature,  $T_R = 338^\circ\text{K}$  is a reference temperature,  $V_{BE}(T_R)$  is the voltage across the base-emitter junction at temperature  $T_R$ ,  $\eta$  is a temperature-independent and process dependent constant and  $x$  refers to the temperature dependence of the current forced through the base-emitter junction.

First-order bandgap voltage references compensate the linear component of CTAT voltage, but ineffectively compensate the nonlinear component of the CTAT voltage: in this way  $V_{BG}$  still shows the effects of the logarithmic behaviour of the diode voltage.

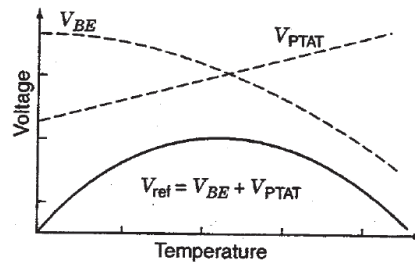


Figure 1.4: First-order bandgap output voltage  $V_{BG}$

### 1.3.2 Second-order reference

In addition to the first-order terms compensation, curvature-corrected bandgap references attempt to approximately cancel also the nonlinear component of the base-emitter voltage [11]. The classical method for this kind of compensation is by the addition of a squared PTAT term ( $PTAT^2$ ) to the output voltage relation of first-order bandgap reference.



The idea is to compensate the negative temperature dependence of the logarithmic term in base-emitter voltage with a positive parabolic term:

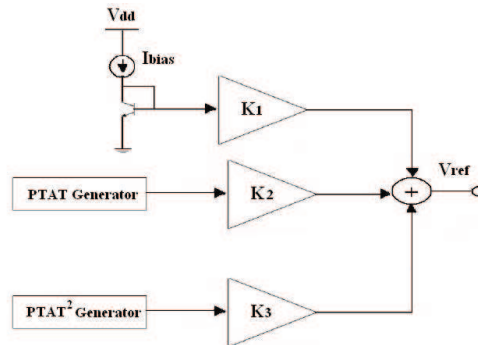


Figure 1.5: Squared PTAT curvature-correction

As a result, the first half of the temperature range exhibits the curvature of a first-order bandgap reference. The squared PTAT voltage term becomes considerably large as the temperature increases. This behaviour is used in order to cancel the increasingly negative temperature dependence of base-emitter voltage at higher temperatures.

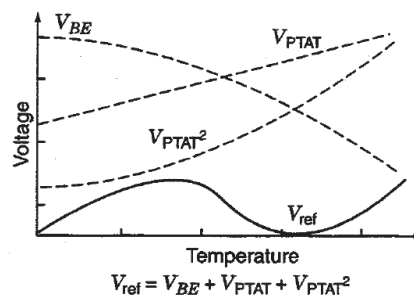


Figure 1.6: Squared PTAT curvature-correction bandgap output voltage

### 1.3.3 The nonlinear reference: Diode Loop

The squared PTAT curvature-correction method is not the only technique used to cancel the higher-order terms of the diode voltage [6, 11]. The curvature-correcting component of a high-order bandgap reference can also be effectively generated through the use of different temperature-dependent currents and a diode voltage loop. In this way it is possible to generate a nonlinear voltage  $V_{NL}$  with a logarithmic behaviour.

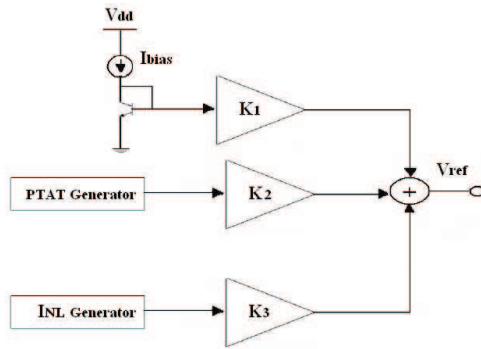


Figure 1.7: Nonlinear curvature-correction reference

The nonlinear temperature dependence of voltage  $V_{NL}$  is designed to cancel the effects of the higher-order components introduced by the base-emitter voltage.

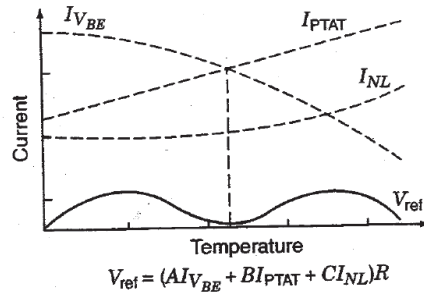


Figure 1.8: Nonlinear curvature-correction bandgap output voltage

## 1.4 The diode loop spt6

The work [16] focused on studying the different implementations of sub-bandgap structures and all the related source of errors, exploiting the structures of a diode loop. The nonlinear bandgap voltage reference has been designed, tested and implemented on silicon in order to evaluate its strenght and its weakness. The experimental results has been compared with the Cadence simulations in order to verify the specifications like the current consumption, the precision over the temperature range and the stability of the voltage reference in different operational conditions.

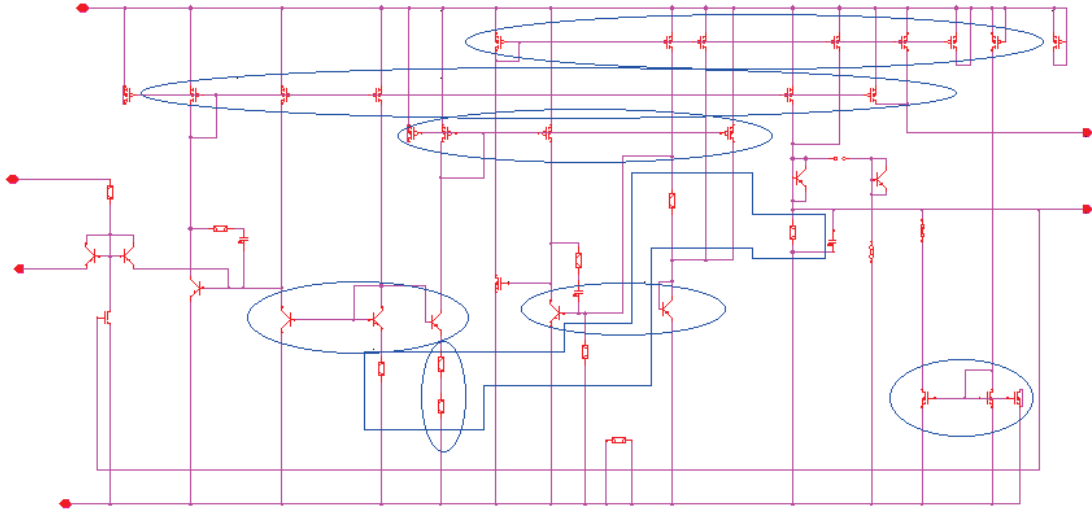


Figure 1.9: Diode Loop of spt6

The aim of the following chapters is to design a new first-order and a new diode loop using a different cheaper technology, trying to retain two accurate and stable bandgap voltage references.

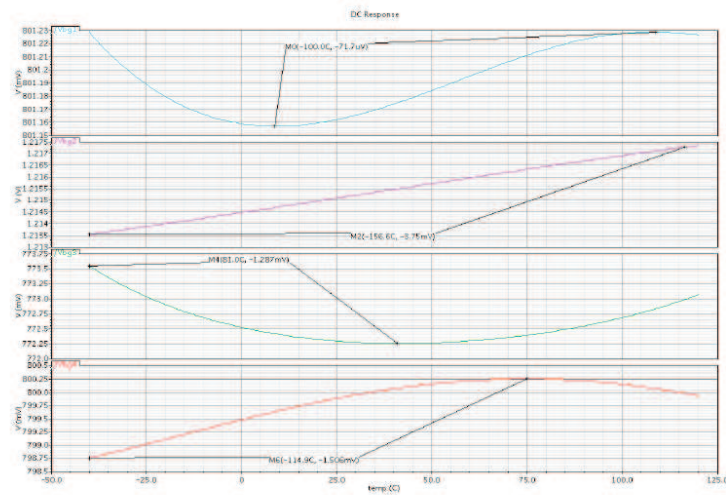


Figure 4.2: Nominal output voltage.

Figure 1.10: Simulation of bandgap output voltage references [16]

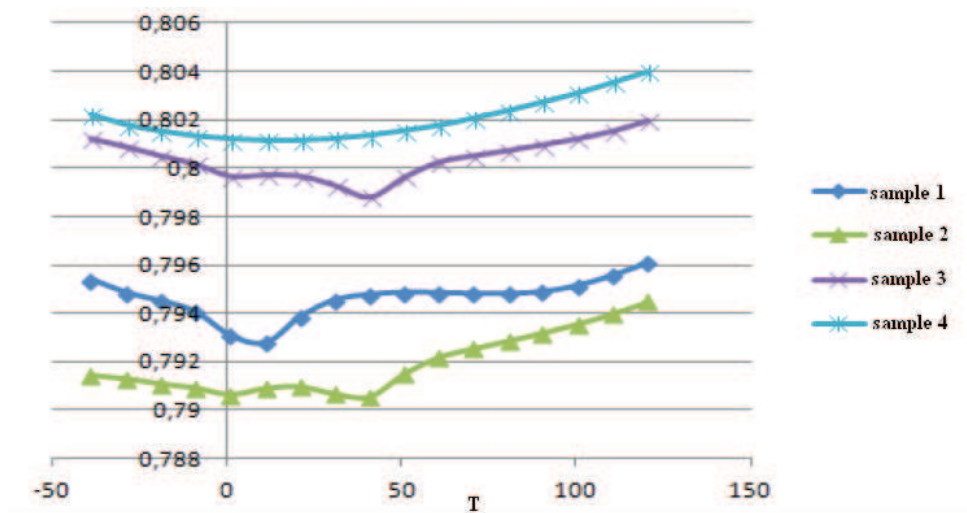


Figure 1.11: Experimental points of a Diode Loop



## Chapter 2

# FIRST ORDER VOLTAGE REFERENCE

### 2.1 First-order compensation

An ideal voltage reference should provide a constant output voltage that is independent of temperature, power supply voltage and line noise. Precision voltage reference are always in great demand in many applications as power converters: to achieve this precision high-order temperature compensation techniques are used. In order to implement these advanced compensation techniques with great accuracy, it is necessary to have a good knowledge of the technological devices used and a good knowledge of their strenghts and weaknesses. It is necessary to get an accurate bandgap voltage reference, a low current consumption and low voltage supply.

The simplest bandgap voltage reference is a first-order compensated. The principle of the circuit is shown in fig. 2.1

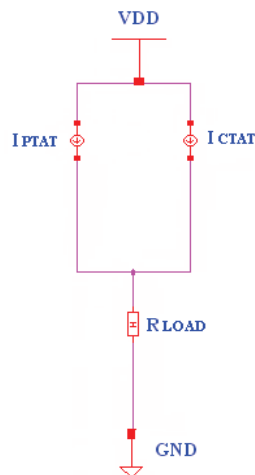


Table 2.1: Current mode topology

The reference described in fig. 2.1 uses a current mode topology: a current mode output stage is obtained by summing temperature-dependent currents into a resistor. Currents must be generated from a certain voltage over a resistance:

a current is generated by a PTAT voltage (“Proportional to absolute temperature”) and a current is generated by a CTAT voltage (“Complementary to absolute temperature”).

The generated current are summed into a load resistor, in this way they have a low sensitivity to the temperature coefficient of resistors. This happens because the transfer function between the voltage that generated that current and its effect on the output voltage is a resistor ratio. If the resistor behave similarly over temperature the ratio of resistor will be temperature independent. The resulting voltage is, to the first order, independent of temperature.

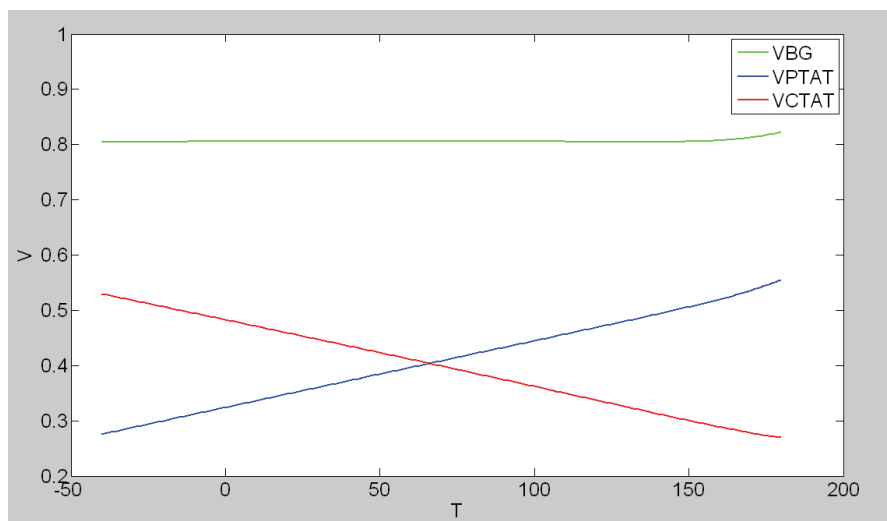


Figure 2.1: Ideal first-order compensation

CTAT current is generated starting from a base emitter voltage of a BJT, while PTAT current is generated with different circuits. The simplest circuit that implements a PTAT current generator is shown in fig. 2.2

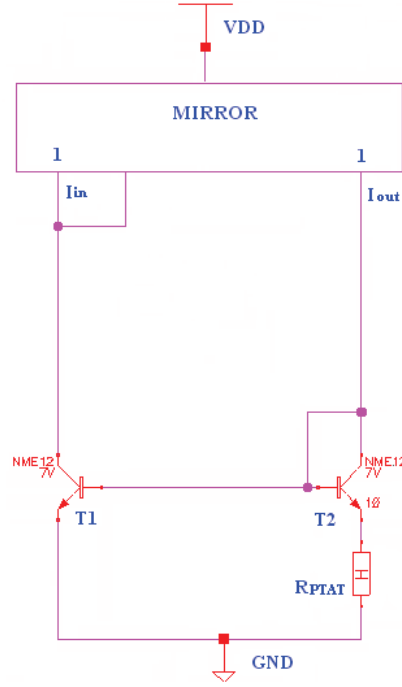


Figure 2.2: PTAT current generator

In the circuit of fig. 2.2 it is possible to identify:

- a current mirror with unity gain current;
- a Kirchhoff Voltage Loop (T1-T2);

referring to the Kirchhoff Voltage Loop, it is possible to write:

$$V_{BE1} - V_{BE2} - I_{OUT}R_{PTAT} = 0 \quad (2.1)$$

The relation between the collector current and the voltage of a BJT can be written as

$$I_C = I_s \exp\left(\frac{V_{BE}q}{k_B T}\right) = I_s \exp\left(\frac{V_{BE}}{V_t}\right) \quad (2.2)$$

$q$  is the electron charge,  $k_B$  is the Boltzmann  $1.38065 \times 10^{-23} JK^{-1}$ .  
 $I_s$ , which is a function of absolute temperature  $T$ , can be written as:

$$I_s(T) = \frac{qA_e n_i^2(T) D(T)}{W_b N_b} = J_s A_e \quad (2.3)$$

where  $A_e$  represents the base-emitter junction area,  $n_i^2$  represents the intrinsic carriers concentration,  $D(T)$  represents the effective minor constant diffusion in the base,  $W_b$  represents the effective width of the base and  $N_b$  is the total number of doping atoms per unit area in base.

If the collector currents of T1 and T2 are equal  $I_{C1} = I_{C2}$ , it is possible to write the following equations:



$$\begin{cases} V_{BE1} = V_t \ln\left(\frac{I_{C1}}{J_s A_{e1}}\right) \\ V_{BE2} = V_t \ln\left(\frac{I_{C2}}{J_s A_{e2}}\right) \end{cases} \quad (2.4)$$

Now using eq. 2.1 and eq. 2.4 it is possible to get the following equation

$$I_{OUT} = \frac{1}{R_{PTAT}} V_t \ln \frac{A_{e2}}{A_{e1}} = \frac{1}{R_{PTAT}} V_t \ln Q \quad (2.5)$$

where  $Q = \frac{A_{e2}}{A_{e1}}$

The simple circuit that implements a CTAT current generator is shown in fig. 2.3

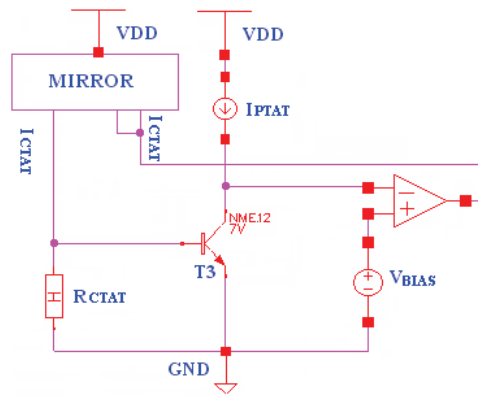


Figure 2.3: CTAT current generator

The temperature characteristics of  $V_{BE}$  are studied extensively by [12]. This work suggests that the analytical form of  $V_{BE}$  is

$$V_{BE} = V_{G0} - \frac{V_{G0} - V_{BE}(T_R)}{T_R} T - (\eta - x) \frac{k_B}{q} T \ln\left(\frac{T}{T_R}\right) \quad (2.6)$$

$V_{G0}$  is the bandgap voltage at 0°K,  $T$  is the absolute temperature,  $T_R = 338^\circ\text{K}$  the chosen reference temperature,  $V_{BE}(T_R)$  is the voltage across the base-emitter junction at temperature  $T_R$ ,  $\eta$  is a temperature-independent and process dependent constant ranging from 3.6 and 4.5 and  $x$  refers to the temperature dependence of the current forced through the base-emitter junction ( $x$  is 0 for a constant temperature current and equals 1 for PTAT current, i.e a proportional to absolute temperature current) [11].

To design a first order compensated bandgap voltage reference we assume that the current forced through base-emitter junction is a PTAT current ( $x = 1$ ) and we use a first-order series' Taylor approximation of  $V_{BE}$ :

$$V_{BE} = \left( V_{G0} + \frac{k_B}{q} T_R (\eta - 1) \right) - \frac{T}{T_R} \left( V_{G0} - V_{BE}(T_R) + \frac{k_B}{q} T_R (\eta - 1) \right) + o(T) \quad (2.7)$$

The CTAT current is given by

$$I_{CTAT} = \frac{V_{BE}}{R_{CTAT}} \quad (2.8)$$

In order to obtain the bandgap output voltage we sum the PTAT current with the CTAT current through a resistance  $R_{OUT}$ :

$$V_{BG} = R_{OUT}(I_{PTAT} + I_{CTAT}) = \frac{R_{OUT}}{R_{PTAT}} \frac{k_B}{q} T \ln Q + \frac{R_{OUT}}{R_{CTAT}} \left[ \left( V_{G0} + \frac{k_B}{q} T_R (\eta - 1) \right) - \frac{T}{T_R} \left( V_{G0} - V_{BE}(T_R) + \frac{k_B}{q} T_R (\eta - 1) \right) \right] \quad (2.9)$$

To have the first-order compensation the linear part of  $V_{BG}$  has to be zero at  $T = T_R$ . Our output voltage  $V_{BG}$  has to satisfy the following equation:

$$\frac{\partial V_{BG}}{\partial T} = \frac{R_{OUT}}{R_{PTAT}} \frac{k_B}{q} \ln Q - \frac{R_{OUT}}{R_{CTAT}} \left[ \frac{1}{T_R} \left( V_{G0} - V_{BE}(T_R) + \frac{k_B}{q} T_R (\eta - 1) \right) \right] = 0 \quad (2.10)$$

$$R_{CTAT} = \frac{\left( V_{G0} - V_{BE}(T_R) + \frac{k_B}{q} T_R (\eta - 1) \right) R_{PTAT}}{T_R \frac{k_B}{q} \ln Q} \quad (2.11)$$

## 2.2 Evaluation of $\eta$ and $V_{G0}$

As we have seen in the previous section it is necessary to know the correct value of  $V_{G0}$  and  $\eta$  to get the correct value of  $R_{CTAT}$ .

Measurements of the energy gap of several semiconductors versus temperature have been published by several authors. According to [12], an expression of  $V_G(T)$  is

$$V_G(T) = V_G(0) - \frac{\alpha T^2}{T + \beta} \quad (2.12)$$

For silicon, the values for the constants in the equations are, according to [12],  $\alpha = 7.021 \times 10^{-4} \text{V/K}$  and  $\beta = 1108 \text{K}$ .  $V_{G0}$  is given in [13] as 1.1557, but as stated there, it should be corrected by the dissociation energy of the exciton, which at the time had not been accurately measured; using the values reported in [12] for the evaluation, it is possible to get  $V_{G0} = 1.1704$ . In a later work 2.12 has been compared to other results and it was suggested that the constants in it should instead have the values  $\alpha = 4.73 \times 10^{-4} \text{V/K}$ ,  $\beta = 636 \text{K}$  and  $V_{G0} = 1.170 \text{V}$ .

In [12] Y. P. Tsividis suggests that  $V_G(T)$  for  $150 \text{K} < T < 300 \text{K}$  can be described by the following empirical equation:

$$V_G(T) = a - bT - cT^2 \quad (2.13)$$

and the three constants are

$$\begin{aligned} a &= 1.1785 \text{V} \\ b &= 9.025 \times 10^{-5} \text{V/K} \\ c &= 3.05 \times 10^{-7} \text{V/K}^2 \end{aligned}$$

Now we will describe a method to evaluate directly  $V_{G0}$  and  $\eta$  from experimental data obtained with Cadence [8].

J. W. Slobboom and H. C. de Graff [8] in 1976 found this equation to describe the collector current of a BJT as a function of  $V_{BE}$ ,  $V_{G0}$  and  $T$ .

$$I_c = f(V_{BE}, V_{G0}, T) = CT^\eta \exp\left(\frac{V_{BE} - V_{G0}}{\frac{k_B T}{q}}\right) \quad (2.14)$$

In order to obtain 2.14, it is necessary to start from 2.3 and write  $n_i^2$  and  $\bar{D}$  as a function of absolute temperature [11]:

$$\begin{aligned}\bar{D} &= f(T) = V_t \bar{\mu} = V_t B T^{-n} \\ n_i^2 &= g(T) = A T^3 \exp\left(\frac{-V_{G0}}{V_t}\right)\end{aligned}\quad (2.15)$$

$\bar{\mu}$  represents the average mobility for minority carries in the base, while  $A$ ,  $B$  and  $n$  are temperature independent constants.

Using 2.3 and 2.15 we get:

$$I_s = \frac{q A T^3 \exp\left(\frac{-V_{G0}}{V_t}\right) (V_t B T^{-n}) A_e}{W_b N_b} = C T^{(4-n)} \exp\left(\frac{-V_{G0}}{V_t}\right)\quad (2.16)$$

Substitute eq. 2.16 in eq. 2.2, it is possible to get eq. 2.14.

In [8] the method to determine  $\eta$  and  $V_{G0}$  is shown for a temperature independent constant collector current  $I_c(T) = I$ . This method can be analyzed and then it will adjust for a PTAT current.

Equation 2.14 contains three independent parameters:  $V_{G0}$ ,  $\eta$  and  $C$ ; therefore  $V_{BE}$  has to be measured at least three different temperatures  $T_1$ ,  $T_2$  and  $T_3$ .

$$I = C T_1^\eta \exp\left(\frac{V_{BE}(T_1) - V_{G0}}{\frac{k_B T_1}{q}}\right) \quad I = C T_2^\eta \exp\left(\frac{V_{BE}(T_2) - V_{G0}}{\frac{k_B T_2}{q}}\right) \quad I = C T_3^\eta \exp\left(\frac{V_{BE}(T_3) - V_{G0}}{\frac{k_B T_3}{q}}\right)$$

After removing the exponential functions, it is possible to obtain a system of equations with two unknown quantities:

$$\begin{cases} \eta T_1 T_2 \frac{k_B}{q} \ln\left(\frac{T_2}{T_1}\right) = T_2 V_{BE}(T_1) - T_1 V_{BE}(T_2) - V_{G0}(T_2 - T_1) \\ \eta T_2 T_3 \frac{k_B}{q} \ln\left(\frac{T_3}{T_2}\right) = T_3 V_{BE}(T_2) - T_2 V_{BE}(T_3) - V_{G0}(T_3 - T_2) \end{cases}\quad (2.17)$$

The theoretical circuit used to find the values to solve 2.17 is drawn in 2.4. The goal is to find for a BJT the values of  $V_{BE}$  (and in the next step the values of  $I_c$  too) when it is biased with a known current. However, using eq. 2.4 base currents gives an error.

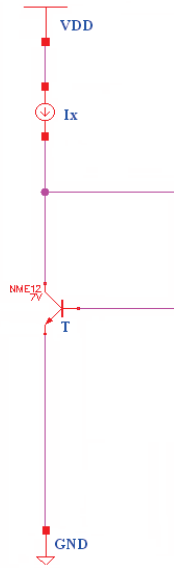


Figure 2.4: Theoretical circuit to evaluate  $V_{G0}$  and  $\eta$

Analyzing the circuit and the the base currents, it is possible to observe that the real collector current of T is a bit different from the current  $I_X$ :

$$I_C = I_X - I_{BASE} = J_S A_e \exp\left(\frac{V_{BE}}{V_t}\right) \tag{2.18}$$

So it is possible to write  $V_{BE}$  as a function of  $I_X - I_{BASE}$ :

$$V_{BE} = V_t \ln\left(\frac{I_X - I_{BASE}}{J_S A_e}\right) \tag{2.19}$$

For a temperature independent constant collector current  $I_c(T) = I = 1\mu A$ , a base current  $I_{BASE}(T = 27^\circ C) = 9.477nA$ , at room temperature, is found.

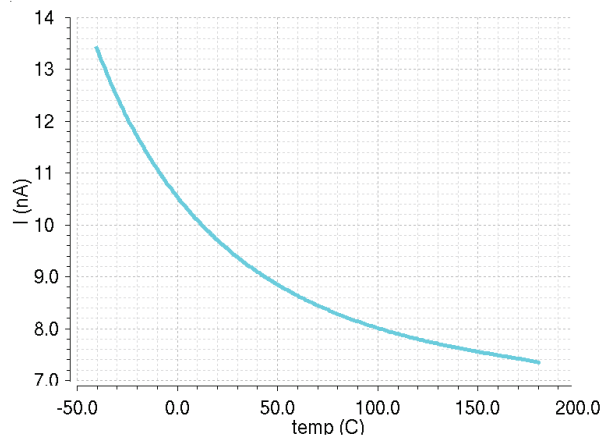


Figure 2.5: Base current for 2.4

It is also possible to evaluate also  $\ln(I_X - I_{BASE})$  (magenta) and  $\ln(I_X)$ (blue) with  $I_X = I = 1\mu A$ .

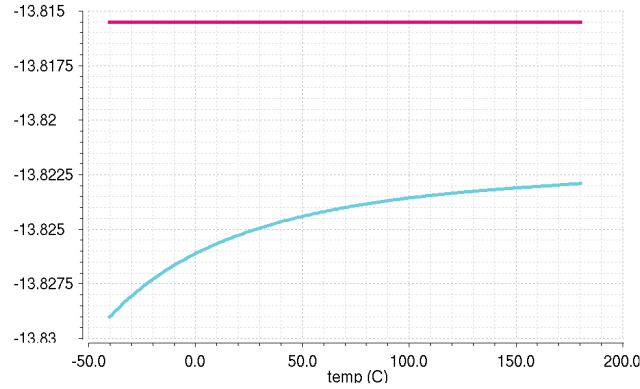


Figure 2.6:  $\ln(I_X - I_{BASE})$ (magenta) and  $\ln(I_X)$ (blue)

In order to avoid this error due to base currents it is better to use the scheme in 2.7. In this case the base current is delivered by an ideal amplifier (realized with the Cadence E-gain block).

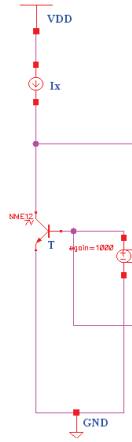


Figure 2.7: Circuit used to get “experimental” values with Cadence

Cadence can be used in order to to get “experimental” values ( $I = 1\mu A$ )

$$T_1 = -40^\circ\text{C} \quad T_2 = 27^\circ\text{C} \quad T_3 = 65^\circ\text{C}$$

$$V_{BE}(T_1) = 788.586\text{mV} \quad V_{BE}(T_2) = 664.073\text{mV} \quad V_{BE}(T_3) = 590.766\text{mV}$$

It is possible to calculate the paramters:  $\eta = 4.514$   $V_{G0} = 1.1189\text{V}$

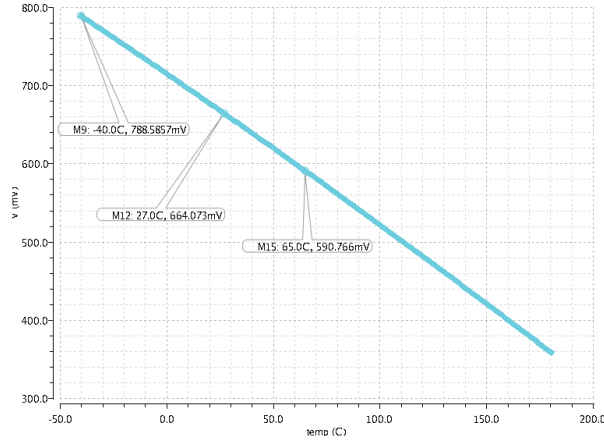


Figure 2.8:  $V_{BE}$  for a temperature-independent  $I_C$

If a PTAT current is forced through the base-emitter junction the equation 2.17 will be changed:

$$\begin{cases} T_1 T_2 \frac{k_B}{q} \ln\left(\frac{I_{C1}}{I_{C2}}\right) + \eta T_1 T_2 \frac{k_B}{q} \ln\left(\frac{T_2}{T_1}\right) = T_2 V_{BE}(T_1) - T_1 V_{BE}(T_2) - V_{G0}(T_2 - T_1) \\ T_2 T_3 \frac{k_B}{q} \ln\left(\frac{I_{C2}}{I_{C3}}\right) + \eta T_2 T_3 \frac{k_B}{q} \ln\left(\frac{T_3}{T_2}\right) = T_3 V_{BE}(T_2) - T_2 V_{BE}(T_3) - V_{G0}(T_3 - T_2) \end{cases} \quad (2.20)$$

Cadence will give the “experimental” values. In this case, it is necessary to measure also the collector current at the three different temperatures:

$$\begin{array}{lll} T_1 = -40^\circ\text{C} & T_2 = 27^\circ\text{C} & T_3 = 65^\circ\text{C} \\ V_{BE}(T_1) = 782.707\text{mV} & V_{BE}(T_2) = 663.234\text{mV} & V_{BE}(T_3) = 593.385\text{mV} \\ I_{C1} = I_C(T_1) = 747.121\text{nA} & I_{C2} = I_C(T_2) = 968.221\text{nA} & I_{C3} = I_C(T_3) = 1093.621\text{nA} \end{array}$$

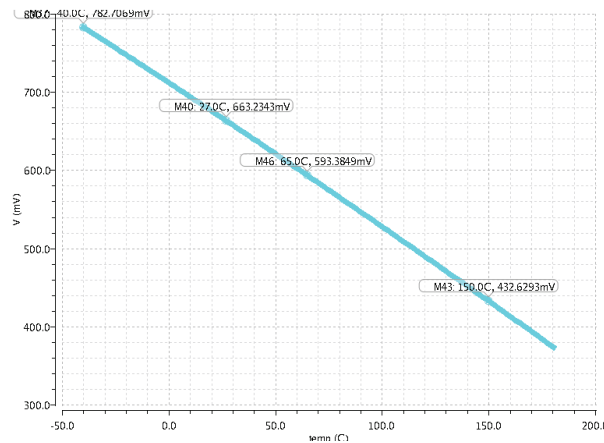


Figure 2.9:  $V_{BE}$  for a PTAT  $I_C$

It is possible to calculate the parameters:  $\eta = 4.504$   $V_{G0} = 1.1191\text{V}$

## 2.3 PTAT stage

### 2.3.1 Ideal schematic

In order to design a first-order compensated voltage reference, the first brick to build is the PTAT stage. PTAT current is a “proportional to absolute temperature” current. If the absolute temperature grows up, PTAT current follows it linearly. If the goal is to design a first order voltage reference, it is necessary to design an accurate PTAT stage. The choice of the components of PTAT stage influences the precision and the current consumption of bandgap voltage generator. PTAT stage needs a feedback loop to control the bias point and so it is interesting to evaluate the different architectures of PTAT stages, analysing their features. The analysis will start with an ideal schematic to implement fig. 2.2: in order to have a perfect mirror ratio, the current mirrors are realized with a CCCS (current control current source).

The realization of an ideal PTAT stage is shown in the following figures:

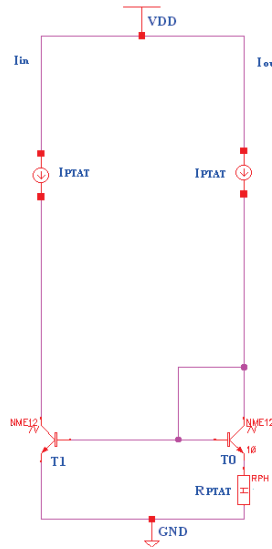


Figure 2.10: PTAT stage: ideal schematic

The ideal current mirror is implemented with two ideal current generators.

If  $I_{IN} = I_{OUT}$  it is possible to join eq. (2.1) and eq. (2.2), in order to get the equation to evaluate  $R_{PTAT}$ :

$$R_{PTAT} = \frac{V_t}{I_{OUT}} \ln Q \quad (2.21)$$

$Q$  represents the ratio between the emitter area of the two BJT. At room temperature (300°K)  $V_t = 25.9mV$ , moreover  $A_{e2} = 10A_{e1}$  and  $Q = 10$ .

As previously stated PTAT stage needs a feedback loop to control the bias point. It is fundamental to use a second circuit in order to force the currents of the circuit to be equal. In order to realize this feedback loop it is necessary to use an error amplifier. The feedback loop can be designed with different architectures using mosfet or BJT.

The goal is to get an accurate and low voltage PTAT stage. In order to evaluate the robustness of the circuit it is necessary to evaluate:

- the current mismatch of the real current mirrors;
- the headroom and the current consumption of transistors;
- the influence of the temperature dependent base currents and base-emitter voltage (if BJT are used to implement current mirrors);
- the stability at all the temperatures to avoid the circuit from oscillating.

### 2.3.2 Real PTAT stage with BJT: first architecture

It is possible to design three different PTAT stages with three different feedback loops to implement the ideal schematic. During the analysis the performances of all the realization will be compared and finally the best solution will be chosen. The first architecture will use BJT to implement feedback loops and current mirrors: BJT have a good matching in this technology .

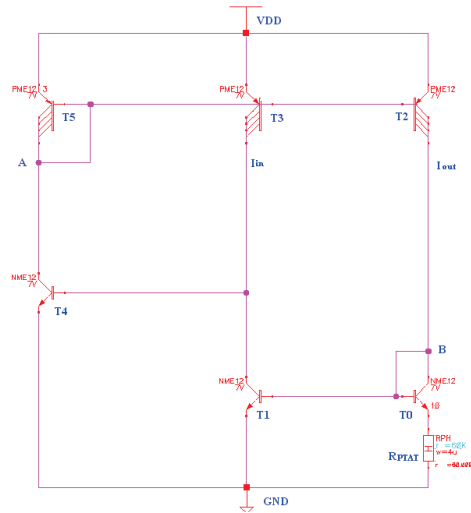


Figure 2.11: First PTAT stage with feedback loop

In this case we realize the feedback loop in order to regulate the bias point with a npn BJT (T4) and with a pnp BJT (T5). T4 sense the error current while T5 pnp force the correct bias point:  $I_{IN} = I_{OUT}$  and  $I_{IN} = I_{C_{T1}}$ .

It is possible to analyze the qualitative behaviour of PTAT loop by applying a small AC signal to the circuit. This small AC signal is superimposed on the circuit containing a large signal and it perturbs the voltages and the currents by small values around the correct bias point. In this way it is possible to evaluate the effect of a small signal perturbation  $i_{in}$  on the loop superimposed to the large signal  $I_{IN}$  (i.e.  $I_{in} = I_{IN} + i_{in}$ ).

If  $i_{in} > i_{C_{T1}}$  the error current  $i_e = i_{in} - i_{C_{T1}} > 0$  is sunk by the base of T4. The base current ( $i_{b_{T4}}$ ), the collector current ( $i_{c_{T4}}$ ) and the base-emitter voltage ( $v_{be_{T4}}$ ) of T4 will increase, pulling-down node A. This effect will increase  $i_{c_{T5}}$  and  $i_{c_{T2}}$ , so node B will pull-up. In this way  $v_{be_{T0}}$ ,  $i_{c_{T0}}$  and  $i_{c_{T1}}$  will increase.

If  $i_{in} < i_{C_{T1}}$  the error current  $i_e = i_{in} - i_{C_{T1}} < 0$  is pushed by the base of T4. Base current ( $i_{b_{T4}}$ ), the collector current ( $i_{c_{T4}}$ ) and the base-emitter voltage ( $v_{be_{T4}}$ ) of T4 will reduce, while node A will pull-up by T5. This effect will reduce  $i_{c_{T5}}$  and  $i_{c_{T2}}$ , while node B will pull-down by T0. In this way  $v_{be_{T0}}$ ,  $i_{c_{T0}}$  and  $i_{c_{T1}}$  will reduce.

Analyzing this solution it is possible to observe that the DC PTAT current ( $I_{PTAT}$ ) has a strange behaviour: PTAT current decreases at high temperatures.



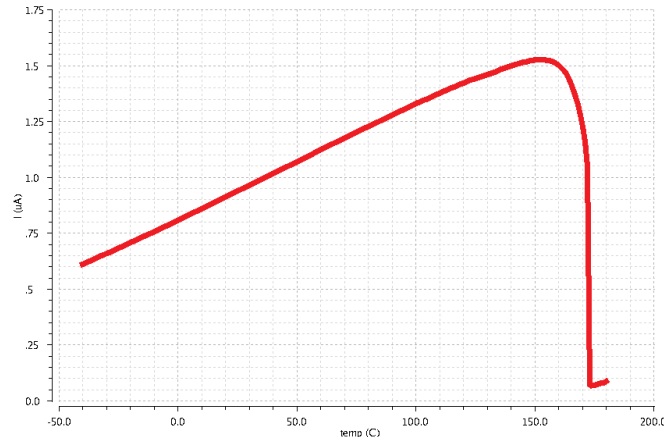
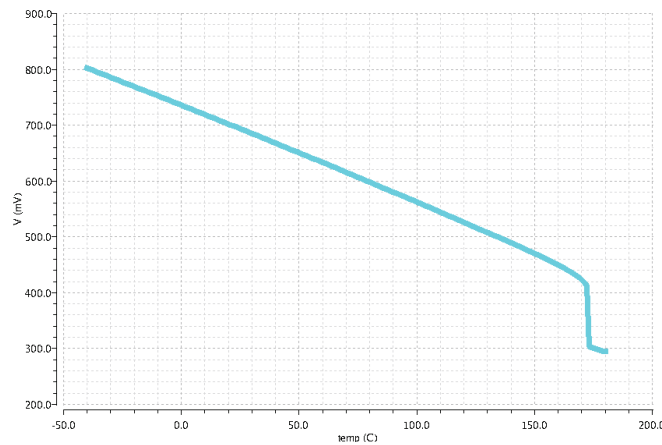


Figure 2.12: PTAT current decreases at high temperatures

This behaviour of PTAT current can be explained by observing base-emitter voltage of T4:

Figure 2.13:  $V_{BE}$  decreases at high temperatures

$V_{BE}$  decreases at high temperatures and so T1 hasn't enough headroom. In order to avoid this behaviour, it is possible to degenerate with a resistance ( $R_{DEG}$ ) the emitter of T4 fig. (2.14): in this way a higher voltage drops at high temperatures.

BJT have a good marching in this technology, however they introduce errors with parasitic temperature-dependent base currents. Analyzing the circuit of fig. (2.14), it is possible to see that the DC collector currents of T1 and T0 are different:

$$\begin{cases} I_{C_{T1}} = I_{C_{T3}} - I_{C_{T4}} = I_{PTAT} - I_B \\ I_{C_{T0}} = I_{C_{T2}} - I_{C_{T0}} - I_{B_{T1}} = I_{PTAT} - 2I_B \end{cases} \quad (2.22)$$

It is necessary to compensate the errors due to base currents: our PTAT current has to be as much accurate as possible. It is necessary to design the BJT feedback loop, in order to improve the precision of the stage.

The current absorbed by base of T4 has to be equal to the sum of base currents absorbed by T0 and T1:

$$I_{B_{T4}} = I_{B_{T0}} + I_{B_{T1}} \tag{2.23}$$

Base current can be written as a function of collector current and current gain, i.e.  $I_B = f(I_C, \beta) = \frac{I_C}{\beta}$ . In order to get eq.(2.23), it is necessary to satisfy the following equation for collector currents:

$$I_{C_{T4}} = 2I_{C_{T1}} = 2I_{C_{T0}} \tag{2.24}$$

In order to reach the purpose of improving the precision of PTAT feedback loop, the area of base-emitter junction of T5 has to be changed: collector current of T4 is the collector current of T3 mirrored by T5. In order to satisfy eq. (2.23), it is necessary to set  $A_{e_{T5}} = 2A_{e_{T3}} = 2A_{e_{T2}}$ .

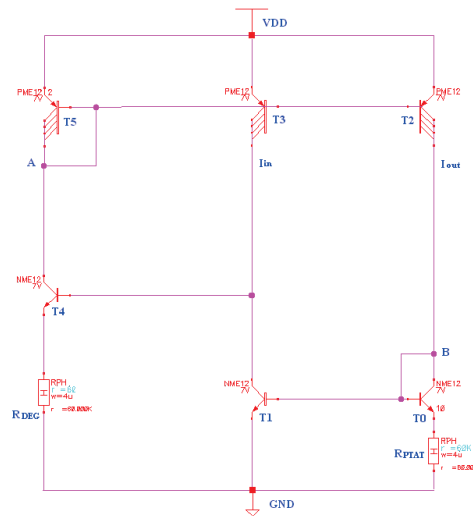


Figure 2.14: Degeneration with resistance of emitter of T4 to avoid low drop voltage at high temperatures

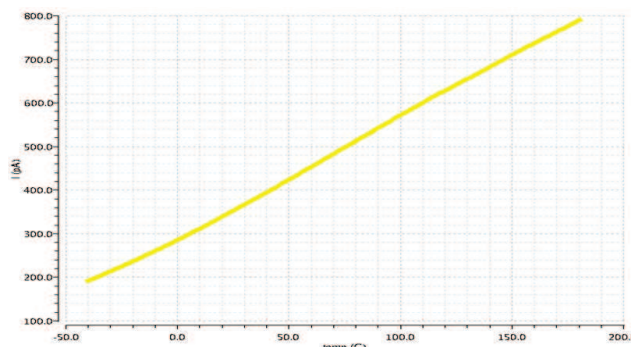


Figure 2.15: Mismatch collector current ( $I_{C_{T3}} - I_{C_{T2}}$ ) after balancing base currents

It is also interesting to investigate the effects of process tolerance and device mismatch, i.e. random variations of physical quantities of technological devices which should be nominally equal. Device parameter variations can limit the accuracy the reference voltage design can achieve, hence they need careful attention. Current-mirror

mismatch, followed by  $V_{BE}$  spread, package shift and resistor mismatch are the most important sources of random errors in bandgap reference circuits.

The errors due to current-mirror mismatch is originated by the deviation in the ratio of current mirrors. This deviation may be originated by various factors, like  $W/L$  mismatch, threshold voltage mismatch, lambda effects of MOS devices and area mismatch of bipolar devices. The spread in the base-emitter voltage of bipolar transistor is a considerable source of error and it is critical because it directly translates to an error in the bandgap reference voltage. Package shift is a post-package error that can only be effectly eliminated by post-package trimming techniques. Process variations can give a large deviation of resitors values (often as large as 20%): this variation changes the PTAT current flowing in the circuit.

Variation of the parameters are described by gaussian random variables and the impact of these random parameter variations on circuit behaviour can be studied with Monte Carlo simulation by analyzing a large set of circuit realisation with randomly varied devices. The tool generates a specified number of modified netlists with randomly varied device characteristics. Each circuit device is modified according to a mismatch model for the device type. All netlists are simulated and their results are collected: these results reflect the behaviour of a circuit to mismatch and process errors.

Referring to fig. (2.14), it is possible to evaluate (at room tempaure  $T = 27^{\circ}\text{C}$ ) the statistical distribution of PTAT current through a  $R_{PH} = 380\text{K}\Omega$  (i.e. the distribution of a PTAT voltage).

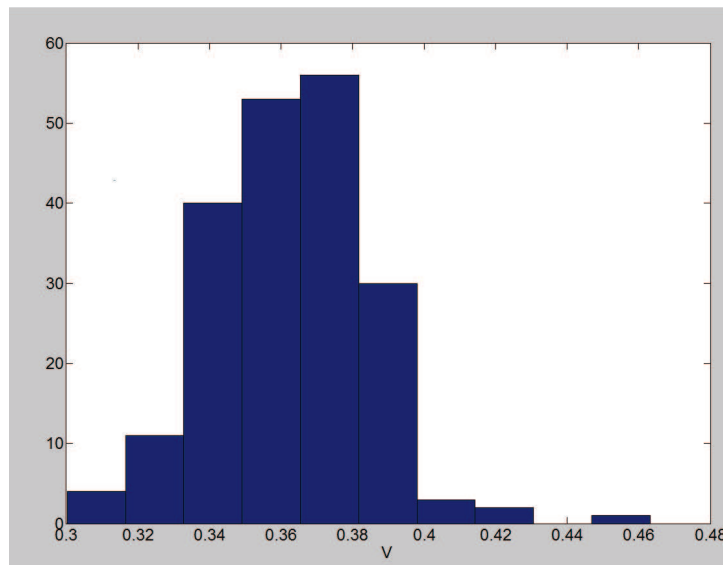


Figure 2.16: Montecarlo Analysis of the first architecture PTAT stage  $\bar{\mu} = 36.23\text{V} \times 10^{-2}$   $\sigma = 2.21 \times 10^{-2}$

### 2.3.3 Real PTAT stage with BJT: second architecture

It is possible to design an other PTAT stage with a similar feedback loop the get the correct bias point.

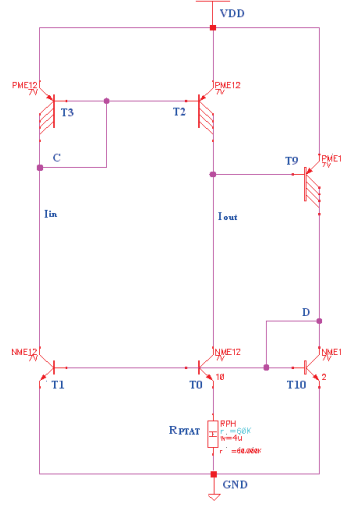


Figure 2.17: Second PTAT stage with a different feedback loop

In this architecture, a pnp BJT (T9) is used to sense the error current and a npn (T10) is used to mirror a current proportional to the sensed error. The feedback loop has to force the correct bias point:  $I_{IN} = I_{OUT}$  and  $I_{CT2} = I_{OUT}$ .

It is possible to analyze the qualitative behaviour of PTAT loop by applying a small AC signal to the circuit. This small AC signal is superimposed on the bias point. In this way it is possible to evaluate the effect of a small signal perturbation  $i_{in}$  on the loop superimposed to the large signal  $I_{IN}$  (i.e.  $I_{in} = I_{IN} + i_{in}$ ).

If  $i_{CT2} > i_{out}$  the error current  $i_e = i_{CT2} - i_{out} > 0$  is sunk by the base of T9 and this effect will reduce the collector current of T9 ( $i_{CT9}$ ) and its  $v_{be}$ . Node D will pull-down by T10, so  $v_{be}$  of T10 and collector current of T1,  $i_{CT1}$ , will reduce. In this way node C will pull-up by T3, reducing  $v_{beT3}$ ,  $i_{CT3}$  and  $i_{CT2}$ .

If  $i_{CT2} < i_{out}$  the error current  $i_e = i_{CT2} - i_{out} < 0$  is pushed by the base of T9 and this effect will increase the collector current of T9 ( $i_{CT9}$ ) and its  $v_{be}$ . Node D will pull-up by T9, so  $v_{be}$  of T10 and collector current of T1,  $i_{CT1}$ , will increase. In this way node C will pull-down by T1, increasing  $v_{beT3}$ ,  $i_{CT3}$  and  $i_{CT2}$ .

The precision of this stage can be also improved with the compensation of the current mismatch due to base currents: in order to minimize the current mismatch it is possible to change the dimension of T10.

By analyzing, the circuit it is possible to observe that the DC collector currents of T0 ( $I_{CT0}$ ) and T1 ( $I_{CT1}$ ) are different:

$$\begin{cases} I_{CT1} = I_{CT3} + I_{BT3} + I_{BT2} = I_{PTAT} + 2I_B \\ I_{CT0} = I_{CT2} + I_{BT9} = I_{PTAT} + I_B \end{cases} \quad (2.25)$$

In order to get  $I_{CT0} = I_{CT1}$ , it is necessary that  $I_{BT9} = I_{BT3} + I_{BT2}$ , i.e. the base current of T9 has to be equal to the sum of base currents of T3 and T2. If the base currents are similar (i.e.  $I_{BT3} \simeq I_{BT2}$ ), in order to improve the precision of PTAT loop it is necessary to satisfy the following equation:

$$I_{BT9} = 2I_{BT3} \quad (2.26)$$

If T9 and T3 have the same current gain  $\beta$ , eq. (2.26) can be written using collector currents:

$$\frac{I_{CT9}}{\beta} = \frac{2I_{CT3}}{\beta} \quad (2.27)$$

$I_{CT9}$  is the collector current of T3 mirrored by T1-T10; in order to satisfy (2.27) it is necessary that  $A_{eT10} = 2A_{eT1}$ .

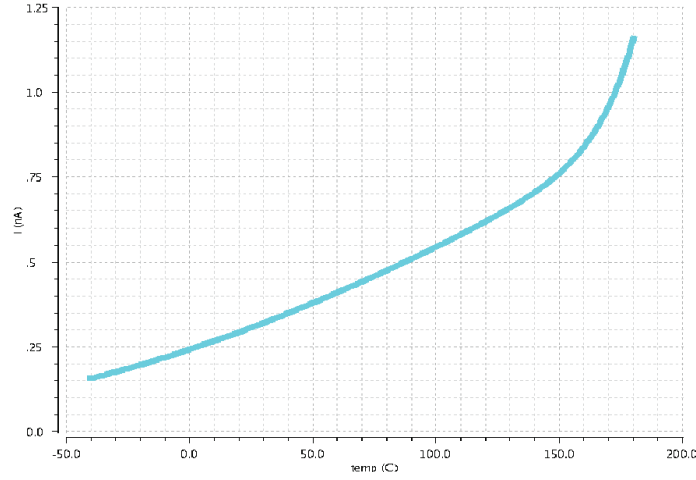


Figure 2.18: Collector currents mismatch ( $I_{C_{T3}} - I_{C_{T2}}$ ) after base current compensation

Now it is possible to compare the two architectures with Montecarlo Analysis: in this way, the effects of process and mismatch are evaluated by observing the spread at  $65^\circ\text{C}$  of PTAT current, through a resistance of  $340\text{K}\Omega$ .

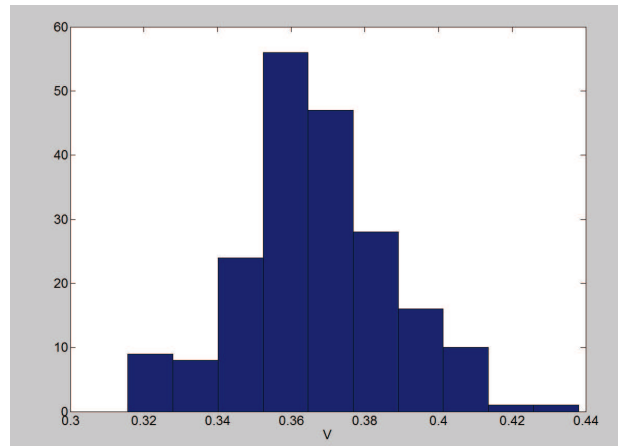


Figure 2.19: Montecarlo Analysis of the second architecture PTAT stage  $\bar{\mu} = 36.71\text{V} \times 10^{-2}$   $\sigma = 2.06 \times 10^{-2}$

The two PTAT stage have a very similar architecture, however the second is better because the Montecarlo resulting distribution is narrower.

Moreover it is possible to compare the precision of the generated PTAT voltage: it is possible to compare the PTAT voltage generated by the first architecture and the PTAT voltage generated by the second architecture with an ideal PTAT voltage generated by Matlab. referring to fig. (2.20) “first Vptat” rappresents the PTAT voltage generated by the first achitecture, “second Vptat” rappresents the PTAT voltage generated by the second architecture, while “ideal Vptat” rappresents the PTAT voltage generated by Matlab.

The PTAT voltages are generated by the PTAT currents across a resistance  $R_P$  ( $R_P = 340\text{k}\Omega$ ) :

$$V_{PTAT} = I_{PTAT}R_P = \frac{R_P}{R_{PTAT}} \frac{k_b}{q} T \ln Q \quad (2.28)$$

A ratio of resistances will be temperature independent: in this way it is possible to evaluate also the derivative of  $V_{PTAT}$  fig. (2.21):

$$\frac{\partial V_{PTAT}}{\partial T} = \frac{R_P}{R_{PTAT}} \frac{k_b}{q} \ln Q \quad (2.29)$$

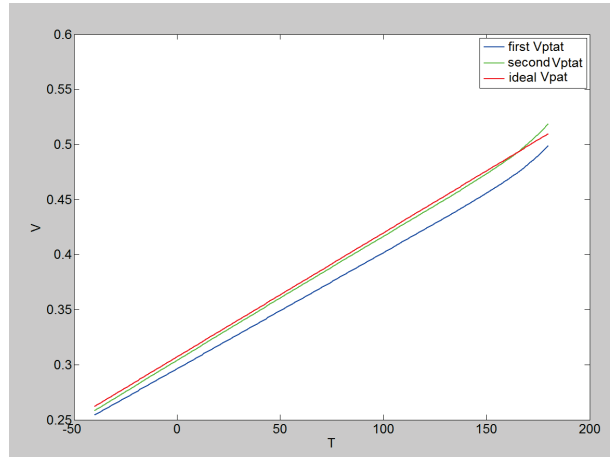


Figure 2.20: Comparison of different PTAT voltage

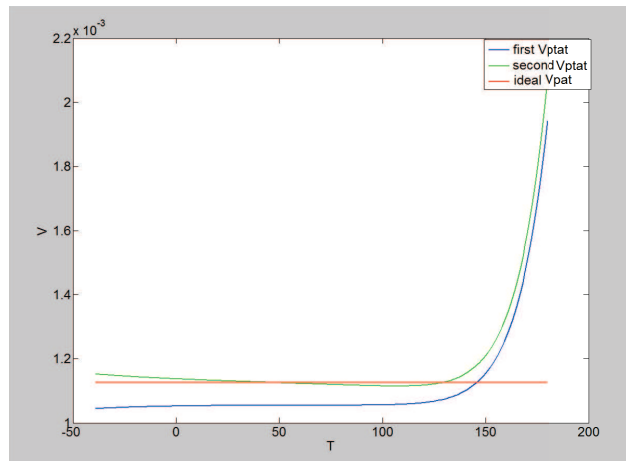


Figure 2.21: Comparison of the derivate of PTAT voltage

Fig. (2.20) and fig. (2.21) show that the PTAT voltage generated by the second architecture is closer to the ideal PTAT voltage. Fig. (2.21) shows that real PTAT voltage rises up very quickly if  $T > 120^\circ\text{C}$ : this is due to the leakage.

### 2.3.4 Real PTAT stage with BJT: third architecture

In [13] is described a new schematic to implement PTAT stage fig. 2.22. In this case a different structure to obtain PTAT current is used. The core of this architecture is the KVL, formed by  $T_0$ - $R_{PTAT}$ - $T_1$ :

$$V_{BE_{T0}} - V_{BE_{T1}} = I_{PTAT} R_{PTAT} \quad (2.30)$$

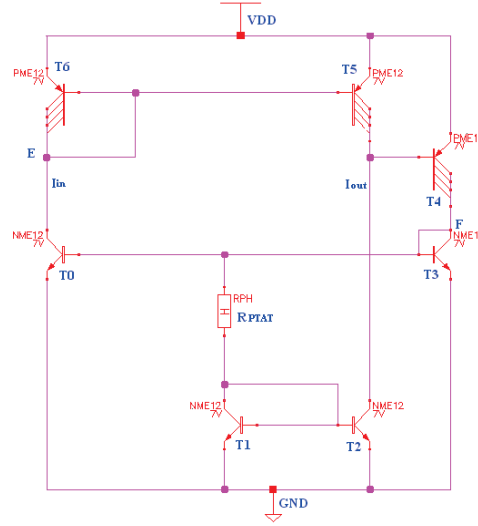


Figure 2.22: Third solution

The PTAT current is mirrored with npn current mirror T1-T2. A feedback loop is necessary to force the correct bias point  $I_{CT5} = I_{OUT}$ . It is possible to implement this feedback loop, as in the previous architecture: with T4 senses the error ( $I_E = I_{CT5} - I_{OUT}$ ), while T3 mirrors a current proportional to the sensed error.

It is possible to analyze the qualitative behaviour of PTAT loop by applying a small AC signal to the circuit. This small AC signal is superimposed on the bias point. In this way it is possible to evaluate the effect of a small signal perturbation  $i_{c_{T5}}$  on the loop superimposed to the large signal  $I_{CT5}$  (i.e.  $I_{c_{T5}} = I_{CT5} + i_{c_{T5}}$ ).

If  $i_{c_{T5}} > i_{out}$  the error current  $i_e = i_{c_{T5}} - i_{out} > 0$  is sunk by the base of T4 and this effect will decrease the base current of T4 ( $i_{b_{T4}}$ ), its collector current ( $i_{c_{T4}}$ ) and its base-emitter voltage ( $v_{be_{T4}}$ ). Node F will pull-down by T4, so base-emitter voltage of T3, ( $v_{be_{T3}}$ ) and collector current of T0,  $i_{c_{T0}}$ , will decrease. In this way node E will pull-up by T6, decreasing  $v_{be_{T6}}$ ,  $i_{c_{T6}}$  and  $i_{c_{T5}}$ .

If  $i_{c_{T5}} < i_{out}$  the error current  $i_e = i_{c_{T5}} - i_{out} < 0$  is pushed by the base of T4 and this effect will increase the base current of T4 ( $i_{b_{T4}}$ ), its collector current ( $i_{c_{T4}}$ ) and its base-emitter voltage ( $v_{be_{T4}}$ ). Node F will pull-up by T4, so base-emitter voltage of T3 ( $v_{be_{T3}}$ ) and collector current of T0,  $i_{c_{T0}}$ , will increase. In this way node E will pull-down by T0, increasing  $v_{be_{T6}}$ ,  $i_{c_{T6}}$  and  $i_{c_{T5}}$ .

In order to improve the precision of the architecture, it is also possible to compensate the errors due to base currents. By analyzing the architecture, it is possible to observe that feedback loop force  $I_{CT5} = I_{OUT}$  and current mirror T6-T5 force  $I_{CT5} \simeq I_{CT6}$ , however  $I_{OUT} \neq I_{IN}$ :

$$\begin{cases} I_{IN} = & I_{CT6} + I_{B_{T6}} + I_{B_{T5}} = I_{CT6} + 2I_B \\ I_{OUT} = & I_{CT5} + I_{B_{T4}} = I_{CT5} + I_B \end{cases} \quad (2.31)$$

In order to get  $I_{OUT} = I_{IN}$ , it is necessary that:

$$I_{B_{T6}} + I_{B_{T5}} = I_{B_{T4}} \quad (2.32)$$

Base currents can be written as a function of collector current and current gain  $\beta$ , i.e.  $I_B = f(I_C, \beta) = \frac{I_C}{\beta}$ . If we assume that T5, T6 and T4 have the same current gain  $\beta$ , we can write eq. 2.32 using collector currents:

$$\frac{I_{C_{T6}}}{\beta} + \frac{I_{C_{T5}}}{\beta} = \frac{I_{C_{T4}}}{\beta} \quad (2.33)$$

Current mirror T6-T5 force  $I_{C_{T5}} \simeq I_{C_{T6}}$ , so to compensate the errors due to base currents it is necessary that the collector currents satisfy the following equation:

$$I_{C_{T4}} = 2I_{C_{T6}} \quad (2.34)$$

In order to reach the purpose it is possible to change the area of base-emitter junction of T0 and T3 setting  $A_{e_{T3}} = 2A_{e_{T0}}$ .

This architecture of PTAT stage suffers from current mismatch also after balancing base currents. Mirror current ratio between T2 and T1 differs from 1 because of Early Effect: collector's voltage of T1 is different from collector's voltage of T2:

$$\begin{cases} V_{C_{T1}} = & V_{BE_{T0}} - R_{PTAT} I_{PTAT} \\ V_{C_{T2}} = & V_{CC} - V_{EC_{T5}} \end{cases} \quad (2.35)$$

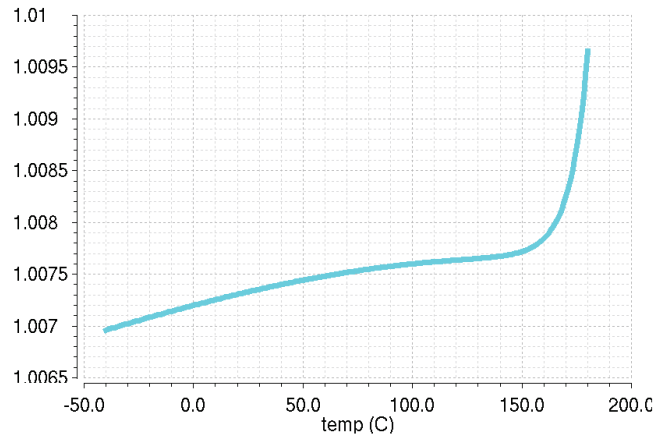


Figure 2.23: Mismatch collector currents T1-T2

It is possible to evaluate the precision of the PTAT voltage of this architecture: in fact, it is possible to compare the PTAT voltages generated by the third architecture and the PTAT voltage generated by the second architecture with an ideal PTAT voltage generated by Matlab. referring to fig. (2.20) “third Vptat” represents the PTAT voltage generated by the third architecture, “second Vptat” represents the PTAT voltage generated by the second architecture, while “ideal Vptat” represents the PTAT voltage generated by Matlab.



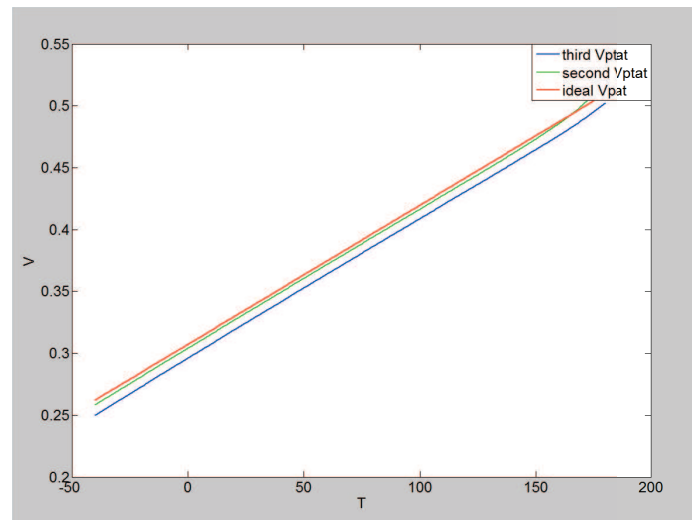


Figure 2.24: Comparison of different PTAT voltage

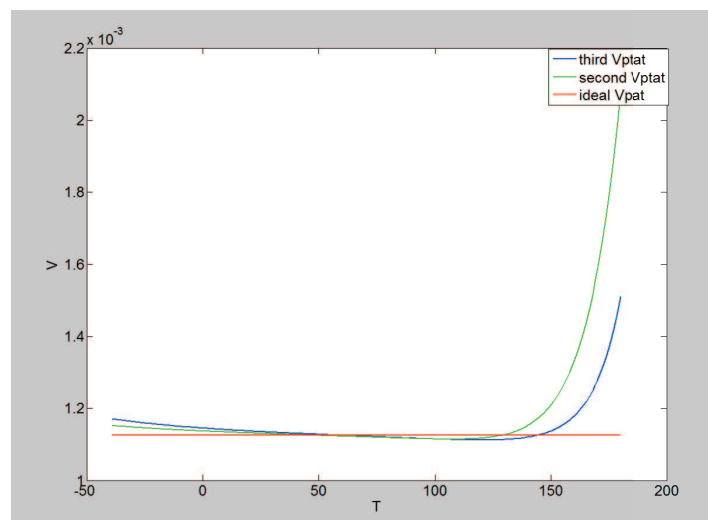


Figure 2.25: Comparison of the derivate of PTAT voltage

Fig. (2.24) and fig. (2.25) show that the PTAT voltage generated by the second architecture is always closer to the ideal PTAT voltage. Fig. (2.21) shows that real PTAT voltage rises up very quickly if  $T > 120^\circ\text{C}$ : this is due to the leakage.

In order to improve mirror ratio, a more complicated structure can be used fig. 2.26:

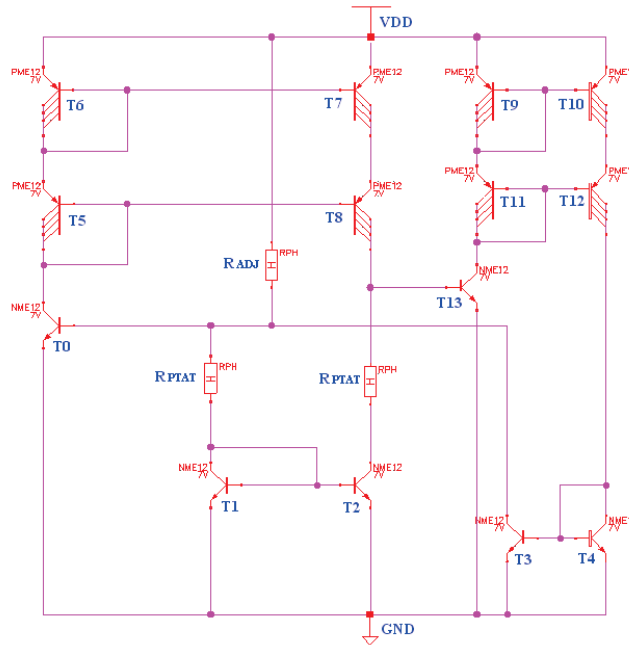


Figure 2.26: Third solution to avoid Early Effect

Using this schematic, collector voltage of T1 is very similar to collector voltage of T2:

$$\begin{cases} V_{C_{T0}} = V_{BE_{T0}} - R_{PTAT} I_{PTAT} \\ V_{C_{T2}} = V_{BE_{T13}} - R_{PTAT} I_{PTAT} \end{cases} \quad (2.36)$$

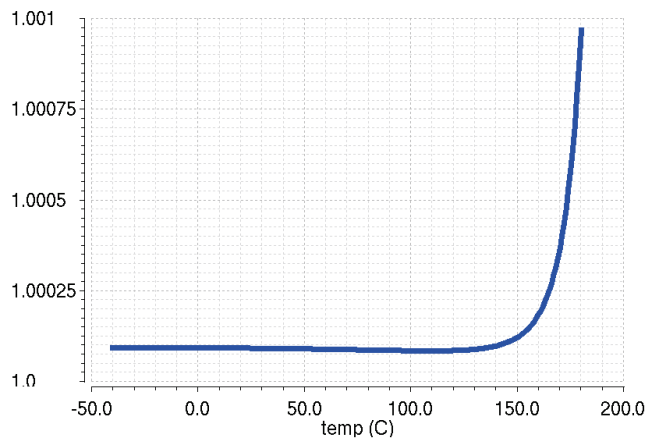


Figure 2.27: Mismatch collector currents T1-T2 will reduce

However this is not a good solution because headroom is the tradeoff:  $V_{CC_{min}} = 2V_{BE} + V_{CE}$  instead of  $V_{CC_{min}} = V_{BE} + V_{CE}$  (fig. 2.14 and fig. 2.17). Moreover current consumption increases and it is very difficult compensate all errors due to base currents.

### 2.3.5 Real PTAT stage with MOSFET

In the previous section a PTAT schematic realized with BJT has been identified. It has been illustrated that the base currents give an error and in order to improve the precision of PTAT stage their effect has been mitigated. Referring to 2.17, after balancing their effect, ideally it should be  $I_{C_{T9}} = I_{C_{T3}} + I_{C_{T2}}$ .

Instead, a perfect balancing of base currents is impossible and in practise  $I_{C_{T9}} \neq I_{C_{T3}} + I_{C_{T2}}$ , as shows the following figure:

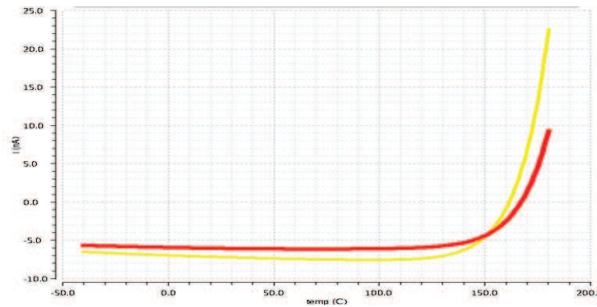


Figure 2.28: Mismatch base currents after optimization of feedback transistors

The base currents limit the precision of the PTAT stage: in order to improve precision of PTAT stage it is necessary to eliminate them. For this reason a new PTAT stage has to be designed: the current mirrors has to be realized with mosfet. Furthermore it is necessary to degenerate emitters' mosfet with a resistance in order to improve the matching.

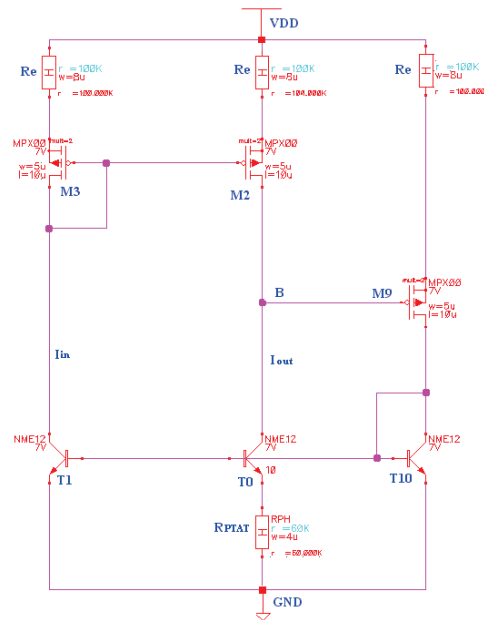


Figure 2.29: PTAT stage realized with mosfet

Using mosfet to implement current mirrors and to implement the feedback transistor, it isn't necessary to

compensate parasitic currents because mosfet M3, M2 and M9 don't inject currents.

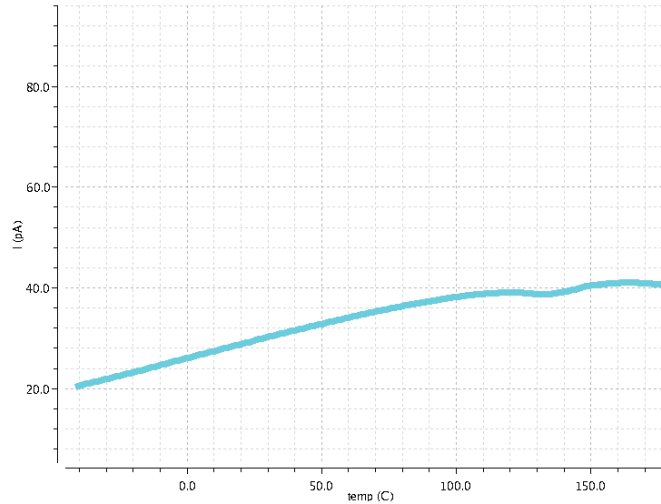


Figure 2.30: Mismatch collector currents  $I_{C_{M3}}$  and  $I_{C_{M2}}$

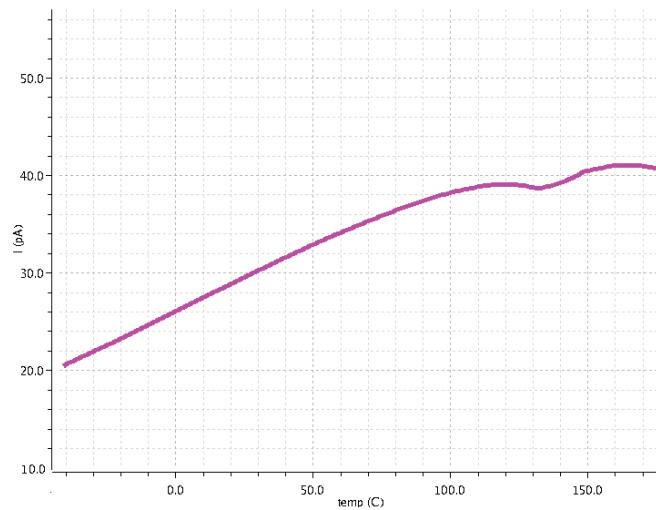


Figure 2.31: Mismatch drain currents  $I_{D_{T1}}$  and  $I_{D_{T0}}$

Currents mismatch are reduced and the precision of PTAT stage is improved by using mosfet. In the PTAT feedback loop there are already three BJT: their base currents can be considered as a disturb injected in the loop and so they are weakened by loop gain.

### 2.3.6 The source degeneration

Matching is the statistical study of the differences in the electrical parameters between identically designed components placed at a small distance in an identical environment and used with the same bias conditions. It is fundamental

to study the matching properties of transistors, because mismatch can seriously affect the performance of analog CMOS integrated circuits.

Mosfet that have identical dimensions and operating at equal current densities do not operate at identical gate-source voltages: the difference in gate-source voltage is called offset voltage.

The source degeneration can be used in order to improve the matching of transistors: this technique transfers part of matching sensitivity from transistors to resistors. The source degeneration improves the overall matching of the circuit and increase the output resistance of mosfet:

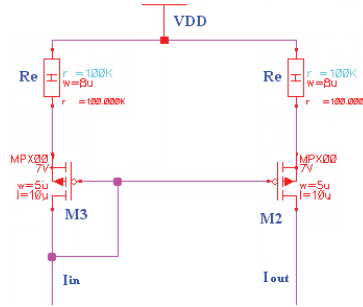


Figure 2.32: Current mirror with source degeneration

The current gain of the current mirror of fig. 2.32 if  $R_e = 0$ , i.e. without source degeneration is:

$$\frac{I_{out}}{I_{in}} = \frac{g_{m2}}{g_{m3}} \quad (2.37)$$

Instead if  $R_e \neq 0$  the current gain becomes:

$$\frac{I_{out}}{I_{in}} = \frac{g_{m2}}{g_{m3}} \frac{1 + g_{m3}R_e}{1 + g_{m2}R_e} \quad (2.38)$$

It is possible to evaluate the sensitivity functions of eq. 2.37 and of eq. 2.38. The sensitivity function  $S$  lets to correlate the variations of a physical quantity  $A$  with the variations of a parameter  $p$ :

$$S = \frac{\partial A}{\partial p} \frac{p}{A} \quad (2.39)$$

referring to fig. 2.32  $\frac{g_{m2}}{g_{m3}} \simeq 1$ . In order to analyze the effect of mismatch it is possible to write  $g_{m2} = kg_{m3}$  where  $k$  is a number close to 1. Eq. 2.37 and eq. 2.38 can be modified as follows:

$$\frac{I_{out}}{I_{in}} = k \quad (2.40)$$

$$\frac{I_{out}}{I_{in}} = k \frac{1 + g_{m3}R_e}{1 + kg_{m3}R_e} \quad (2.41)$$

It is possible to evaluate the sensitivity functions in order to correlate the variation of the current gain  $\frac{I_{out}}{I_{in}}$  with the variations of  $k$ . The sensitivity function  $S_1$  for eq. 2.37 is:

$$S_1 = \frac{\partial \frac{I_{out}}{I_{in}}}{\partial k} \frac{k}{\frac{I_{out}}{I_{in}}} = 1 \quad (2.42)$$

The sensitivity function  $S_2$  for eq. 2.38 is:

$$S_2 = \frac{\partial \frac{I_{out}}{I_{in}}}{\partial k} \frac{k}{\frac{I_{out}}{I_{in}}} = -\frac{1}{R_e g_{m1} + \frac{1}{k}} \quad (2.43)$$

In order to improve the matching it is necessary to reduce the sensitivity function: a great value of  $R_e$  reduces the dependence of  $S_2$  from the variations of  $k$ . It is possible to run a Monte Carlo Analysis (analyzing only mismatch), in order to evaluate the effect of  $R_e$  on current gain.

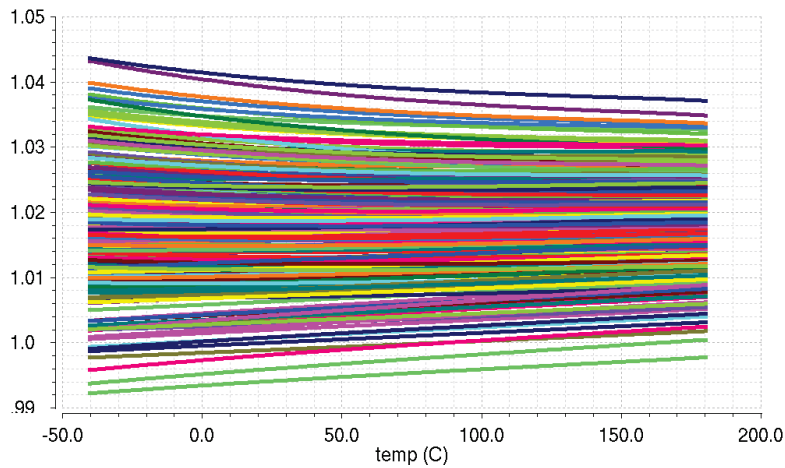


Figure 2.33: Current gain  $\frac{I_{out}}{I_{in}}$  of fig. 2.32 if  $R_e = 0$

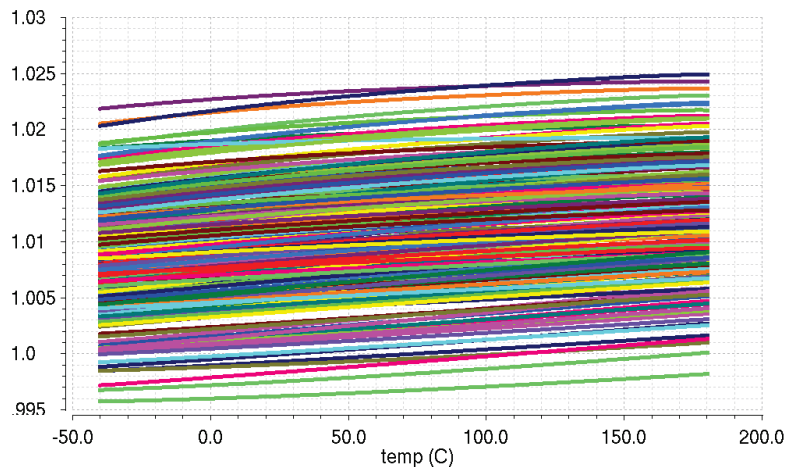


Figure 2.34: Current gain  $\frac{I_{out}}{I_{in}}$  of fig. 2.32 if  $R_e \neq 0$

It is possible to observe that the source-resistance reduces the “spread” (i.e. the amplitude dispersion of PTAT voltage). If source resistance is used in order to improve the matching, headroom is the tradeoff:  $R_e$  has to be chosen in order to minimize the drop voltage across it. At room temperature PTAT current is  $1\mu A$  and  $g_m$  of

mosfet is about  $10^{-5}S$ . In order to maximize the denominator of (2.43) and in order to minimize the drop voltage  $R_e I_{PTAT}$ , a good value for source resistance  $R_e$  is  $R_e = 100K\Omega$ .

### 2.3.7 Dominant pole compensation of PTAT stage

In order to guarantee the asymptotic stability of PTAT stage we must have a loop gain with a positive gain margin and a positive phase margin. Gain margin and phase margin are indexes of stability for a feedback system, though often only phase margin is used rather than both. If  $T$  is the gain loop of a feedback linear system, the phase margin is the difference between  $\angle T(\omega_{0dB})$  and  $-180^\circ$ ,  $\omega_{0dB}$  is the pulse at which  $|T|$  has unity gain, gain margin is the difference between unity gain and  $|T(\omega_{180^\circ})|$ , where  $\omega_{180^\circ}$  is the pulse at which the loop gain phase  $\angle T$  is  $-180^\circ$  fig.2.35.

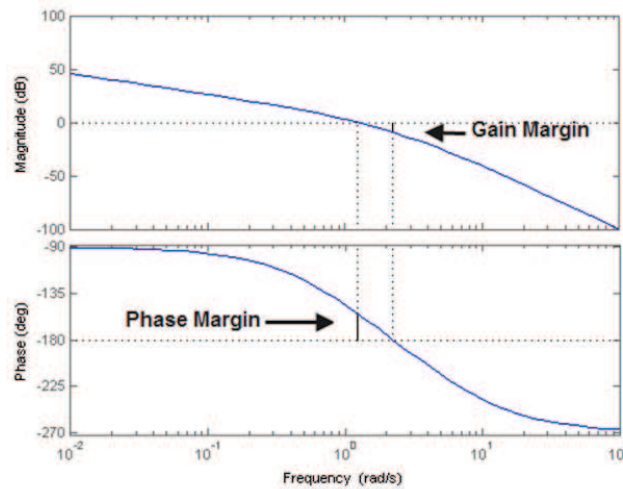


Figure 2.35: Gain margin and phase margin

Note that a target phase margin of  $60^\circ$  is desirable in a feedback amplifier design as a tradeoff between loop stability and settling time in the transient response. Typically the minimum acceptable phase is  $45^\circ$ .

Compensation is a means for controlling the location of poles. The most simple means to guarantee stability of feedback linear system is the dominant pole compensation: a pole is placed at an appropriate low frequency of the gain loop  $|T|$ . This pole has to reduce the gain loop to one ( $0dB$ ) at a frequency close to or just below the location of the next highest pole. The lowest frequency pole is called the dominant pole because it dominates the effect of all of the higher frequency poles.

Dominant pole compensation can be implemented for general purpose by adding an integrating capacitance: this capacitor creates a pole that is set at a frequency low enough to reduce the gain to one before the frequency of the first non dominant pole.

Though simple and effective, dominant pole compensation has two drawbacks:

- it reduces the bandwidth of the amplifier, thereby reducing available open loop gain at higher frequencies. This reduces the amount of feedback available at higher frequency;
- it reduces system's slew rate. This reduction results from the time the current takes for driving the compensated stage to charge the compensating capacitor. The result is the inability of the system to follow accurately rapidly changing signals.

The Miller compensation, in order to implement the dominant-pole compensation, results in the phenomenon of pole splitting. This results in the lowest frequency pole of the uncompensated amplifier “moving” to an even lower frequency to become the dominant pole, and the higher-frequency pole of the uncompensated amplifier “moving” to a higher frequency.

It is possible to analyze PTAT feedback loop in order to get the gain loop:

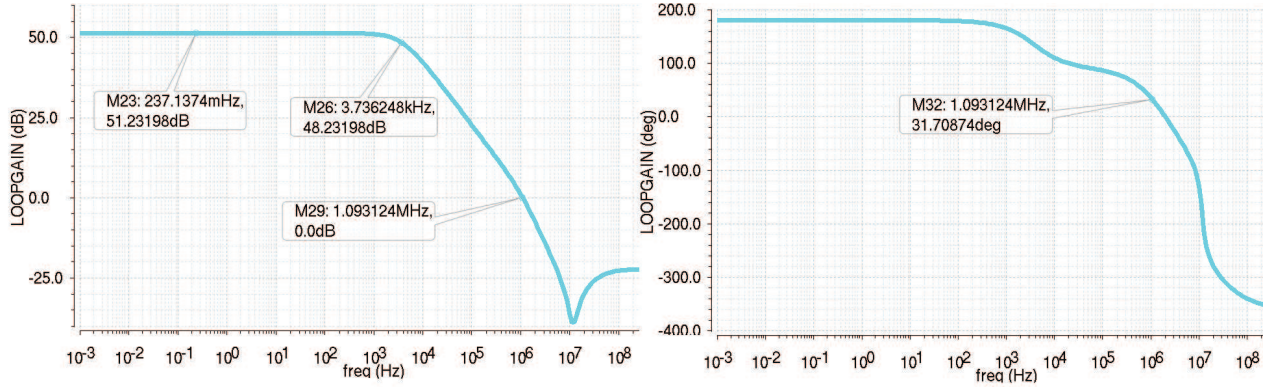


Figure 2.36: Gain loop of PTAT stage before compensation

PTAT loop has already a dominant pole gain loop and a positive phase margin about  $31^\circ$ . However this phase margin is too small, to guarantee the asymptotic stability. Now the PTAT stage has only “ideal” components: parasitic components as parasitic capacitors or parasitic resistors of the layout are not present. The parasitic capacitors may cause a further reduction of the phase margin so it is better to have a bigger phase margin ( $PM \sim 75^\circ$ ).

Analysing PTAT feedback loop fig. 2.29 it is possible to identify node B as a high impedance node:

$$R_{in} = r_{0M2} // r_{0T0} \quad (2.44)$$

where  $r_{0M2}$  is the drain resistance of  $M_2$  and  $r_{0T0}$  is the collector resistance of  $T_0$ . It is also possible to evaluate the capacitance of node B:

$$C_{in} = C_{ds2} + C_{db2} + C_{\mu0} + C_{gd2} + A_{v9}C_{gd9} \quad (2.45)$$

$A_{v9}$  is the voltage gain between gate and drain of M9  $A_{v9} = \frac{g_{m9}r_{09}R_{load}}{r_{09}(1+g_{m9}R_e)+R_e+R_{load}} \simeq \frac{g_{m9}R_{load}}{1+g_{m9}R_e} \simeq 0.106$   
 $A_{v9}$  is reduced by M9 emitter's degeneration and by M9 collector's resistance  $R_{load}$  is a low impedance.

Dominant pole's frequency is:

$$f_{dp} = \frac{1}{2\pi C_{in}R_{in}} \quad (2.46)$$

In order to improve phase margin it is necessary to split at a lowest frequency the dominant pole and so it is necessary to boost  $C_{in}$ .

It is useless place a capacitor between the gate and the drain of M9 because  $A_v < 1$ . It is better place a capacitor  $C_{adj}$  between the gate of M9 and ground. A resistor  $R_{adj}$  can also be added in series to  $C_{adj}$  in order to create a zero with negative real part to improve phase margin around  $1MHz$  to improve phase margin. Choosing  $C_{adj} = 1pF$  and  $R_{adj} = 160K\Omega$ , dominant pole is around  $f_{dp} = 1KHz$ . In this way the phase margin is  $85^\circ$ , the gain margin is  $14.854dB$  and the bandwidth of  $374KHz$ .



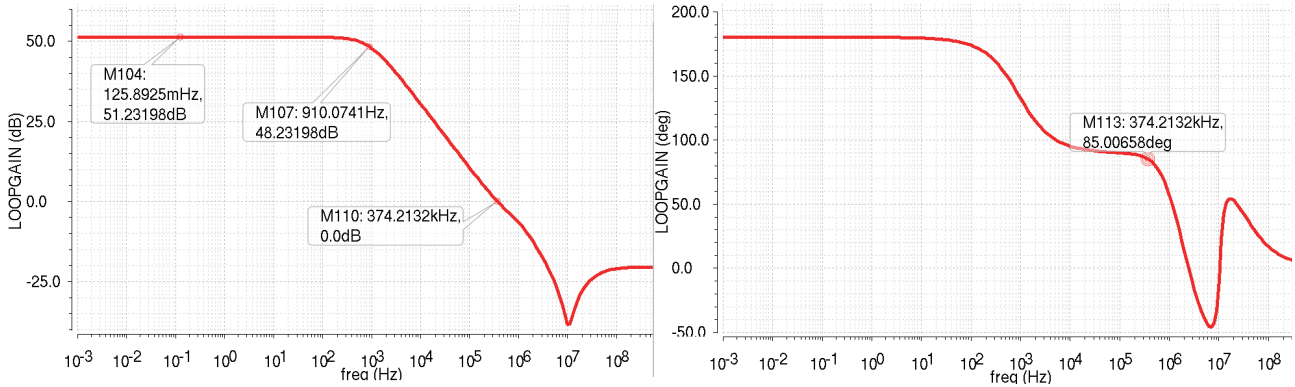


Figure 2.37: PTAT loop gain after dominant pole compensation

Bandgap voltage generator needs to be stable in a wide range of temperature: it is necessary for PTAT loop to be asymptotic stable both at low temperatures and high temperatures. PTAT gain margin and phase margin has to be evaluated at three different temperatures ( $-40^{\circ}\text{C}$ ,  $27^{\circ}\text{C}$ ,  $160^{\circ}\text{C}$ ) with a Montecarlo Analysis:

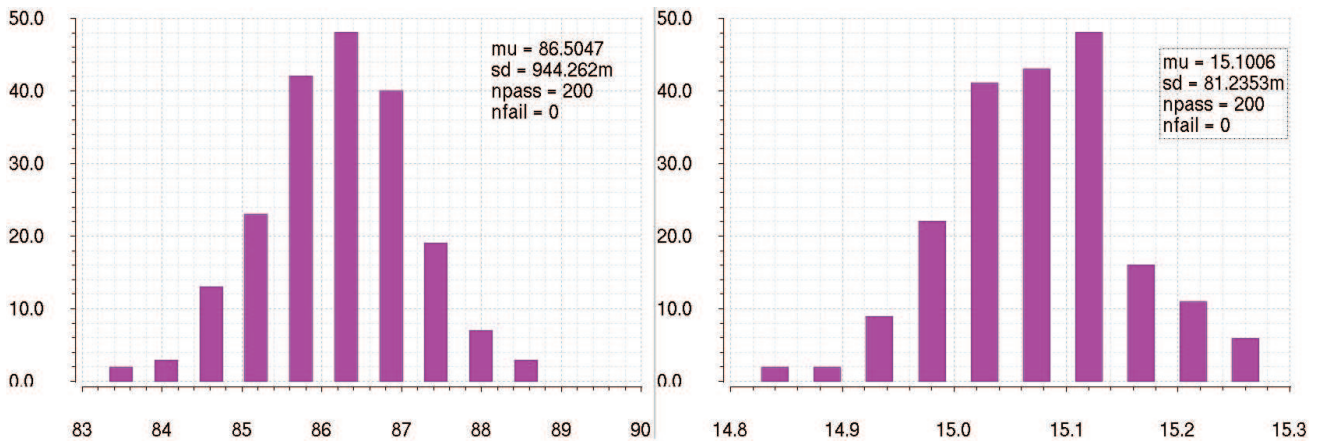


Figure 2.38: Phase margin and Gain margin of PTAT loop at  $-40^{\circ}\text{C}$

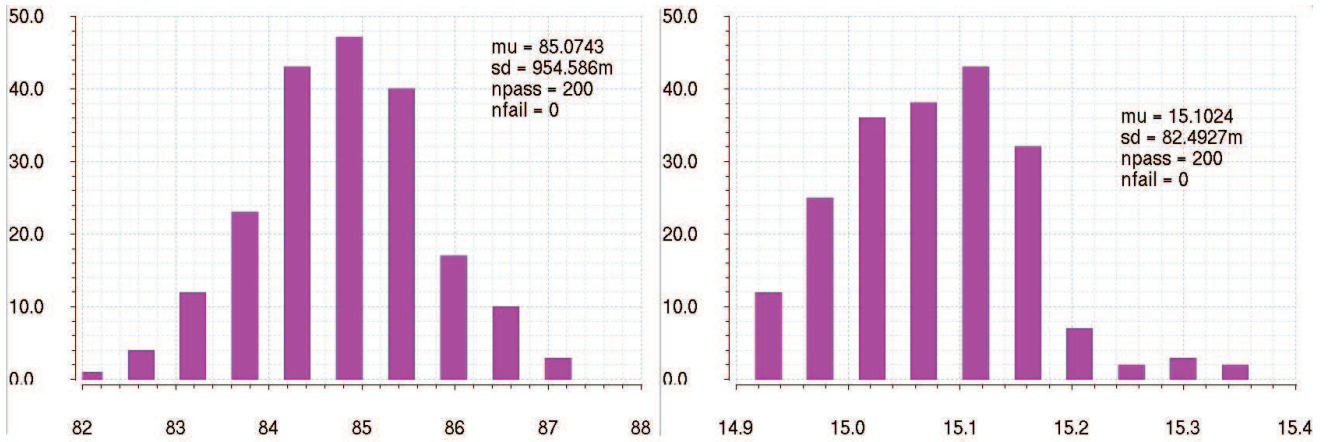


Figure 2.39: Phase margin and Gain margin of PTAT loop at 27°C

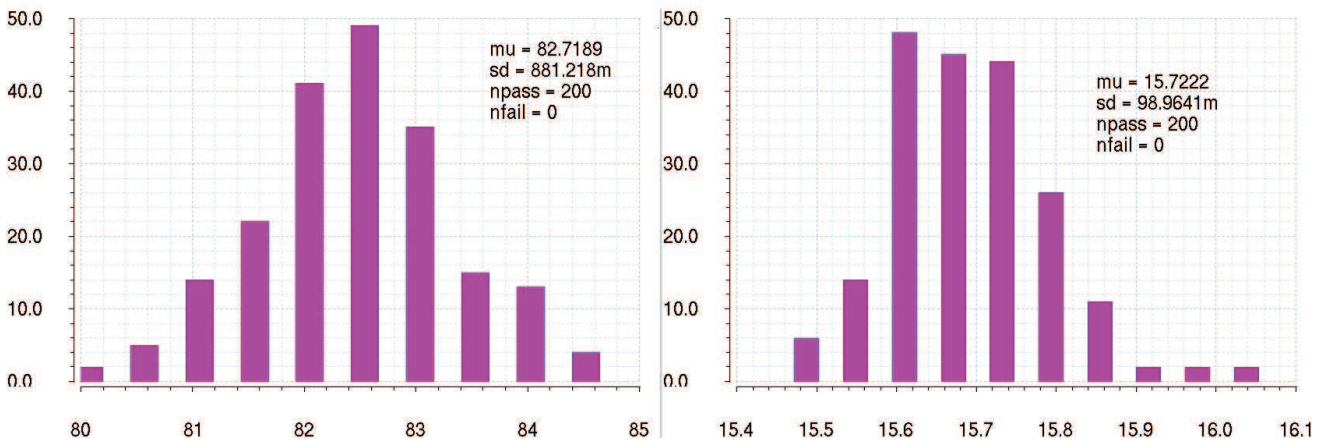


Figure 2.40: Phase margin and Gain margin of PTAT loop at 160°C

## 2.4 CTAT stage

### 2.4.1 Ideal schematic

The second brick to build of the first-order compensated bandgap voltage reference is the CTAT stage (“Complementary to absolute temperature”). A CTAT dependance of voltage is always used in conjunction with curvature correction schematic for first-order or second-order compensation. The core of CTAT stage is a  $V_{BE}$ : the CTAT component is generated by forcing a base-emitter voltage (diode) across a resistor and by mirroring the current flowing through the resistor elsewhere in the circuit. The bias current of CTAT stage should be PTAT in nature, in order to mitigate the nonlinear effects of the logarithmic component of  $V_{BE}$ . In order to implement the circuit of fig. 2.41 it is necessary to design a feedback loop for regulating the correct bias point:  $I_{PTAT} = I_{CT3}$ .

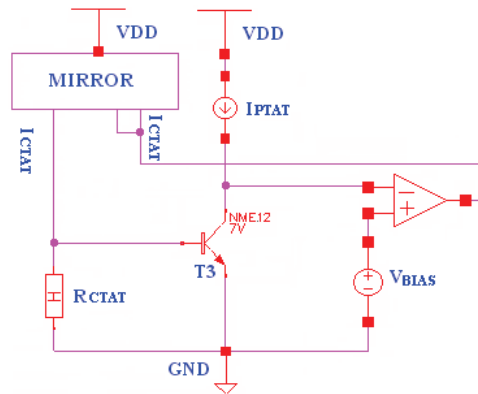


Figure 2.41: CTAT current generator

In order to force the correct bias point a feedback loop has to be implemented with mosfet fig. 2.42. In this way  $I_{PTAT} = I_{CT3}$ . The circuit starts to regulate when gate-source voltage of M5 will exceed the threshold voltage  $V_{THM5} = 1V$ .

It is possible to analyze the qualitative behaviour of PTAT loop by applying a small AC signal to the circuit. This small AC signal is superimposed on the bias point. In this way it is possible to evaluate the effect of a small signal perturbation  $i_{ptat}$  on the loop superimposed to the large signal  $I_{PTAT}$  (i.e.  $I_{ptat} = I_{PTAT} + i_{ptat}$ ).

If  $i_{ptat} > i_{CT3}$  the error current  $i_e = i_{ptat} - i_{CT3} > 0$  will pull-up the gate of M5 (n-mos). In this way the collector current of M5 ( $i_{CM5}$ ) will increase. Then this current will be mirrored by a current-mirror to the base of T3, so base current of T3 ( $i_{bT3}$ ), its  $v_{be}$  and its collector current  $i_{CT3}$  will increase, pulling-down the gate of M5.

If  $i_{ptat} < i_{CT3}$  the error current  $i_e = i_{ptat} - i_{CT3} < 0$  will pull-down the gate of M5 (n-mos). Gate voltage of M5 and its collector current will decrease.  $i_{CM5}$  will be mirrored by a current-mirror to the base of T3, so  $i_{bT3}$ ,  $v_{beT3}$  and  $i_{CT3}$  will decrease.

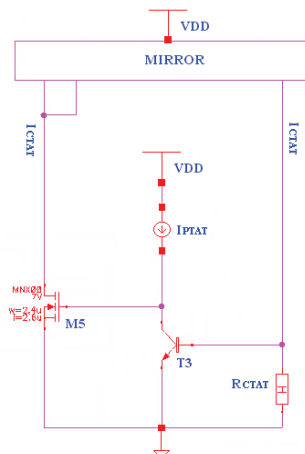


Figure 2.42: Ideal CTAT current generator with feedback loop

In order to evaluate the features of CTAT stage it is possible to implement fig.2.41 with an ideal CTAT stage using ideal current mirrors. The eq. 2.2 can be used in order to get the correct value of eq. 2.11  $R_{CTAT} = 468k\Omega$ .

A loop has been realized, in order to force a PTAT current through the base-emitter junction of T3. In the next section the loop will be stabilized with a dominant pole compensation like PTAT stage.

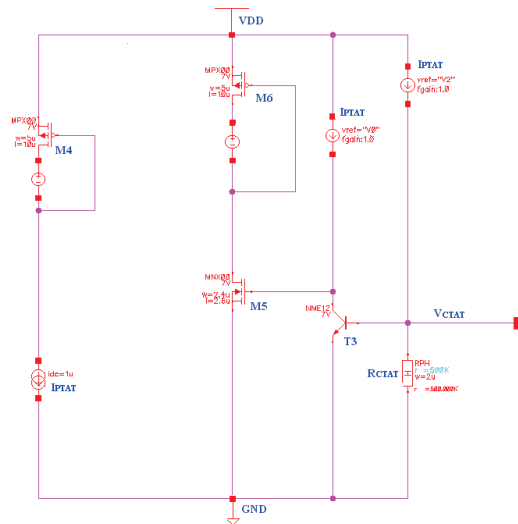


Figure 2.43: Ideal CTAT stage

### 2.4.2 Real schematic

Current mirrors realized with mosfet are better than current mirror realized with BJT, and so it is better to use mosfet to implement current mirror of CTAT stage. In this way it is possible to improve the precision of the stage. Moreover the emitter degeneration is useful to reduce the mismatch.

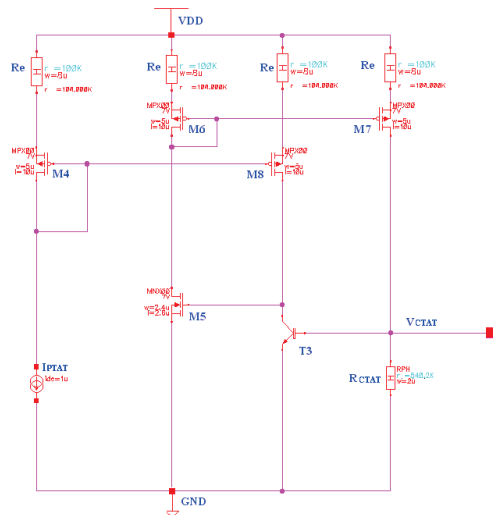


Figure 2.44: Real CTAT stage

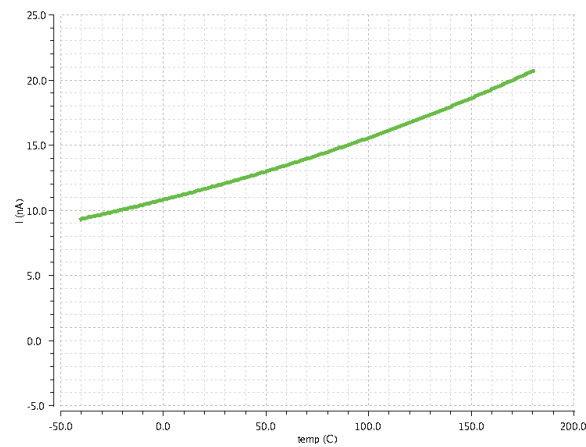


Figure 2.45: Current mismatch between M7 and M6

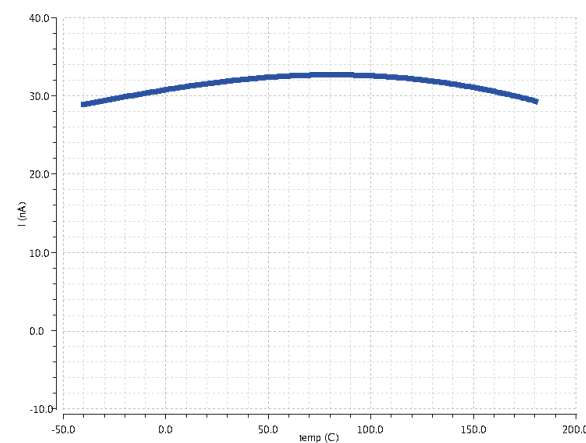


Figure 2.46: Current mismatch between M4 and M8

In real schematic transistor M4 doesn't exist, because PTAT current is directly mirrored from PTAT stage. It is however interesting to observe that there is a mismatch between the current of M4 and the current of M8. The mosfet are used to implement current mirrors in order to have an accurate mirror ratio in CTAT stage and in PTAT stage. In order to get bandgap output voltage it is necessary to sum PTAT current and CTAT current: in this way the currents have to be mirrored out the currents and it is also necessary to find a solution to avoid this loss of precision.

### 2.4.3 Dominant pole compensation of CTAT loop

As we have done for PTAT stage we will compensate CTAT stage to have a CTAT stage with a dominant pole loop gain. referring the circuit of fig. 2.47 we can evaluate the loop gain with a stability analysis of Cadence:

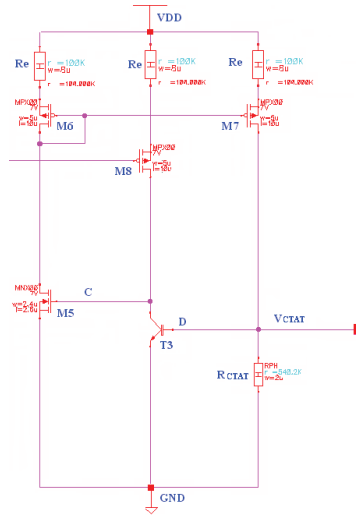


Figure 2.47: Real CTAT stage

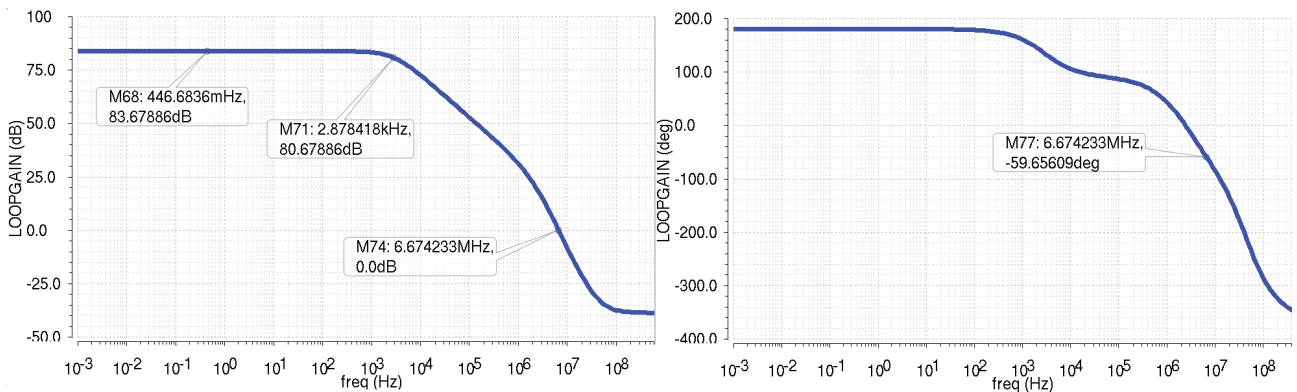


Figure 2.48: CTAT loop gain without dominant pole compensation

CTAT gain loop, without compensation, compared to PTAT loop has a negative phase margin and so CTAT feedback loop isn't asymptotic stable. The goal is getting a phase margin of about  $75^\circ$  and a positive gain margin in order to guarantee the asymptotic stability of the system at all temperatures also with layout parasitic elements. referring to fig. 2.47 we can identify two nodes:

- node C: this is a high impedance node. The impedance  $R_C$  of this node is given by parallel of collector resistance of T3,  $r_{o3}$ , and the series of drain resistance of M8,  $r_{o8}$  and its emitter resistance. While the capacitance of the node is given by  $C_C = C_{d8} + C_{\mu3} + C_{gs5} + C_{gd5}A_C$ .  $A_C = g_{m5}(R_e + \frac{1}{g_{m6}})$  represents Miller's gain between gate-drain of M5 but it isn't very big because the impedance of M6 is a low impedance.
- node D: this is a high capacitance node. The capacitance of this node is given by the capacitance between base-collector of T3  $C_{\mu3}$  multiplied by Miller's gain  $A_D = g_{m3}(R_e + r_{o8})$ . While the impedance of the node is given by the parallel between  $R_{CTAT}$  and the resistance  $r_{\pi3}$  (differential resistance between base-emitter of T3).

These nodes create the dominant pole and the first non dominant pole. The goal is getting a dominant pole loop gain and so it is necessary to find the lowest frequency pole and add a capacitor for getting a stable gain loop.

referring to fig. 2.48 a pole is at  $f = 2.88KHz$ .

It is possible to evaluate the frequency of the two poles:

$$f_{pC} = \frac{1}{2\pi R_C C_C} \simeq 31KHz \quad f_{pD} = \frac{1}{2\pi R_D C_D} \simeq 3KHz \quad (2.47)$$

Dominant pole is created by node D. In order to get a stable loop we can shift at lower frequency ( $\sim 100Hz$ )  $f_{pB}$  by adding a capacitance  $C_{comp} = 1pF$  between base and collector of T3 so Miller's Effect will boost its value. A nulling resistor  $R_{comp} = \frac{1}{g_{m3}} = 66K\Omega$  can also be added to eliminate the zero with real positive part.

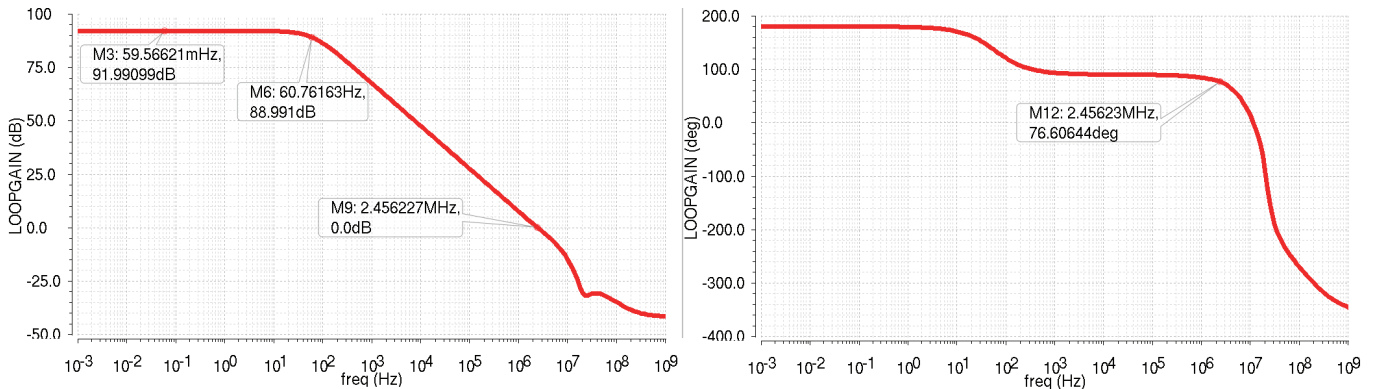


Figure 2.49: CTAT loop gain with dominant pole compensation

The pahse margin and gain margin has to be evaluated with a Montecarlo Analysis at three different temperatures:  $-40^{\circ}C$ ,  $27^{\circ}C$  and  $180^{\circ}C$ .

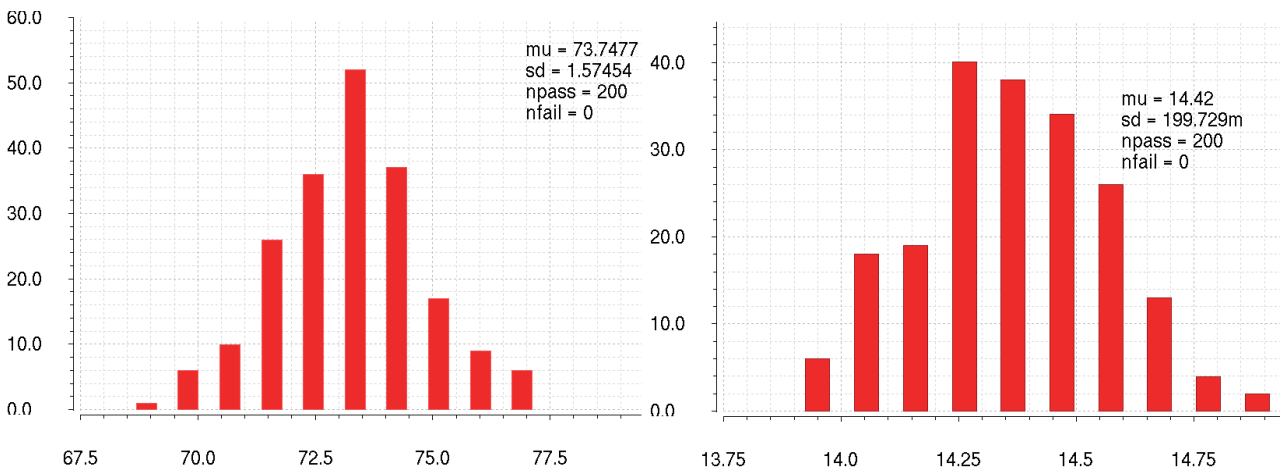


Figure 2.50: Phase margin and Gain margin of CTAT feedback loop at  $-40^{\circ}C$

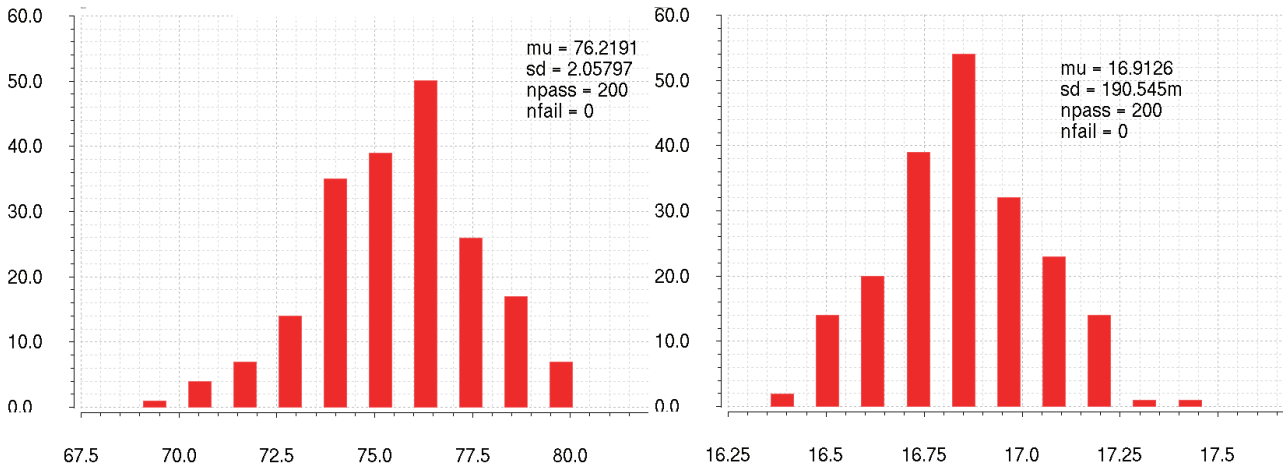


Figure 2.51: Phase margin and Gain margin of CTAT feedback loop at 27°C

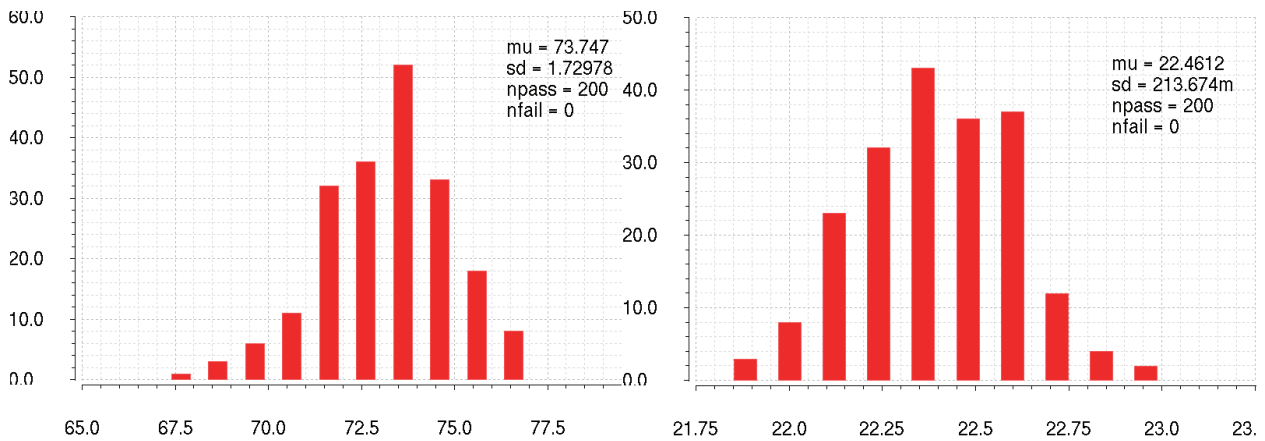


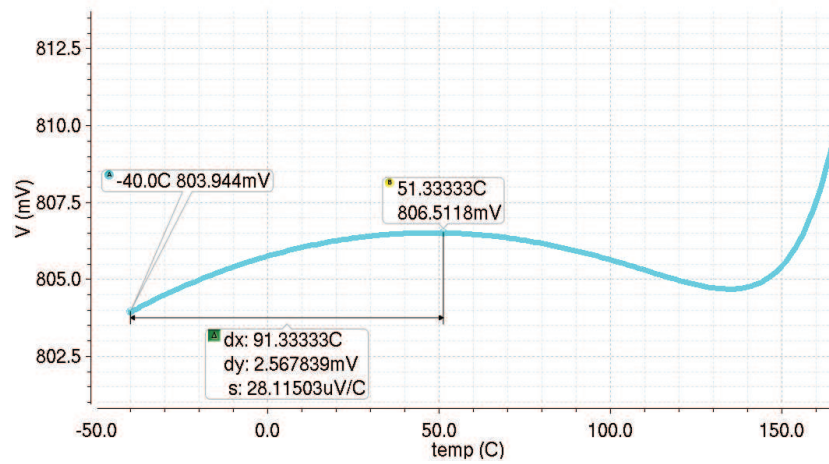
Figure 2.52: Phase margin and Gain margin of CTAT feedback loop at 160°C

## 2.5 First-order compensated bandgap output voltage

### 2.5.1 Evaluation of output voltage

It is possible to use a resistance to sum the accurate current generated in order to get the output voltage as reported in 2.9. The goal is to get  $V_{BG} = 800mV$  and so it is necessary to choose an output resistance to shift up the voltage until the correct value has been found. Choosing  $R_{OUT} = 340K\Omega$  bandgap output voltage becomes  $V_{BG} = 800mV$ .

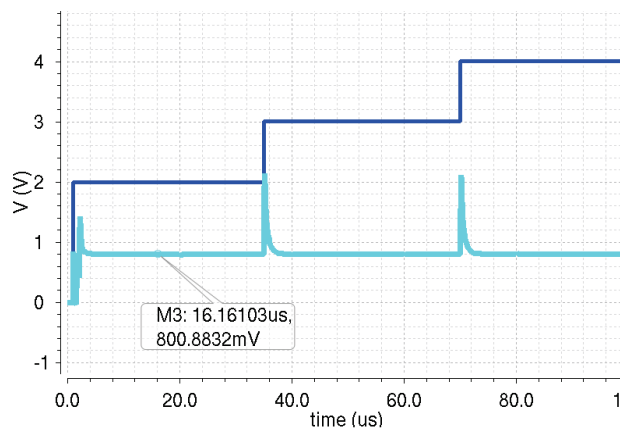


Figure 2.53:  $V_{PTAT}$ ,  $V_{CTAT}$  and  $V_{BG}$ 

## 2.5.2 Transient response with $V_{DD}$ steps

A transient analysis can be used in order to evaluate the step response of first-order bandgap reference. Transient analysis has been run at three different temperatures: low temperature ( $-40^{\circ}\text{C}$ ), room temperature ( $27^{\circ}\text{C}$ ) and high temperature ( $160^{\circ}\text{C}$ ).

In this way the transient response of the circuit has also been verified at the typical conditions. Power supply has been ramped up to the final value for this purpose ( $t_{rise} = 1\text{ns}$ ). The corresponding settling of  $V_{BG}$  is shown in fig. 2.54, 2.55 and 2.56.  $V_{BG}$  has been observed to settle down to its final value without any oscillations as  $V_{dd}$  ramps up in  $1\text{ns}$ .

Figure 2.54: Step response of the first-order compensated voltage bandgap generator at  $T = -40^{\circ}\text{C}$

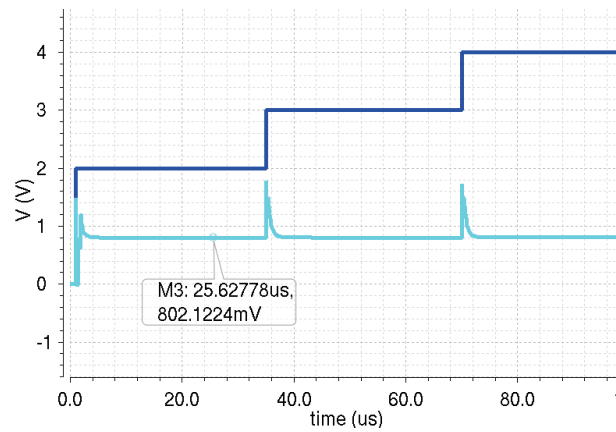


Figure 2.55: Step response of the first-order compensated voltage bandgap generator at  $T = 27^\circ\text{C}$

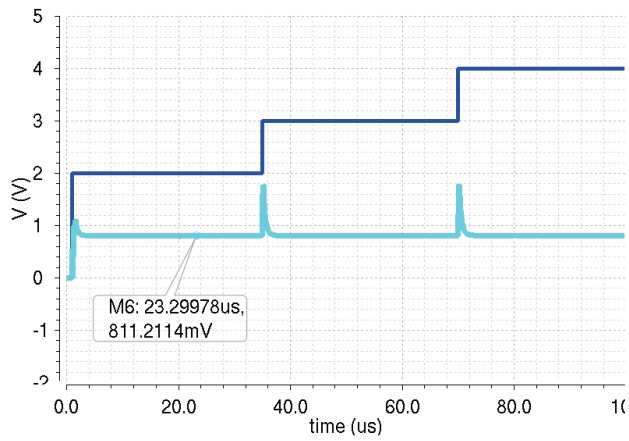


Figure 2.56: Step response of the first-order compensated voltage bandgap generator at  $T = 160^\circ\text{C}$



## Chapter 3

# DIODE LOOP SUPPLY1

### 3.1 Curvature correction

The first-order compensated bandgap, obtained in the previous chapter, has a residual error in the order of  $2.5mV$ . This error is due to the non-compensated non-linear part. The precision of the bandgap structure can be improved using also a curvature correction of the non-linear part. In addition to canceling the first-order terms, curvature corrected bandgap references attempt to approximately cancel the non-linear component of base emitter voltage (diode). The curvature-correcting component of a high-order bandgap reference can also be effectively generated through the use of different temperature-dependent currents and a diode voltage loop. The classical method for doing such compensation is through the addition of a non-linear PTAT term to the output voltage relation of the first-order bandgap reference. The non-linear PTAT term has logarithmic behaviour: the idea is to eliminate the negative temperature dependence of the logarithmic term in  $V_{BE}$  with a positive non-linear term.

Base emitter voltage can be written as in [12]:

$$V_{BE} = V_{G0} - BT - Cf(T) \quad (3.1)$$

where  $V_{G0}$  (extrapolated diode voltage at  $0^\circ K$ ),  $B$  and  $C$  are temperature-independent constants, while  $T$  is the absolute temperature. First-order references sum a positive linear temperature-dependent voltage to a base-emitter voltage to eliminate the effects of temperature linear dependence. The goal of a high-order bandgap references is to sum a temperature-dependent voltage exhibiting both a positive linear and a positive non-linear temperature dependence:

$$V_{BG} = V_{BE}(T) + V_x(T) = V_{BE}(T) + DT + Ef_2(T) \simeq V_{G0} \quad (3.2)$$

### 3.2 Ideal Diode Loop

A particular technique used to get  $Ef_2(T)$  is to generate a non linear correction current. The circuit realization is illustrated in 3.1

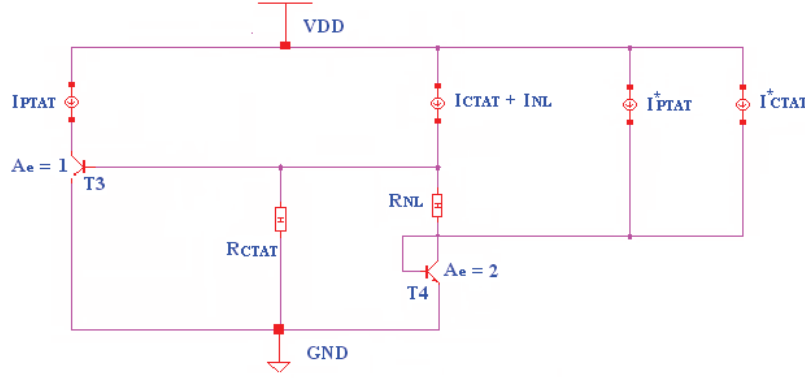


Figure 3.1: Diode-loop curvature ideal topology

Nonlinear current component  $I_{NL}$  is defined by different temperature-dependent currents and by a transistor loop comprised of  $T_3$ ,  $T_4$  and  $R_{NL}$ :

$$I_{NL} = \frac{V_t}{R_{NL}} \ln \left( \frac{I_{C3} A_{e4}}{A_{e3} I_{C4}} \right) = \frac{V_t}{R_{NL}} \ln \left( \frac{2I_{PTAT}}{I_{NL} + I_{CONST}} \right) \quad (3.3)$$

where  $I_{CONST}$  is a current whose temperature dependance is dominated also by the temperature coefficient of the resistors used in the circuit. Output voltage  $V_{BG}$  can be written as:

$$V_{BG} = R_{LOAD}(I_{PTAT} + I_{CTAT} + I_{NL}) \quad (3.4)$$

Eq. 3.4 can be written by using the correct expressions for  $I_{PTAT}$ ,  $I_{CTAT}$  and  $I_{NL}$ :

$$V_{BG} = \frac{R_{LOAD}}{R_{PTAT}} \left( \frac{k_B}{q} T \ln Q \right) + \frac{R_{LOAD}}{R_{CTAT}} \left( V_{G0} - \frac{V_{G0} - V_{BE}(T_R)}{T_R} T - (\eta - 1) \frac{k_B T}{q} \ln \frac{T}{T_R} \right) + \frac{R_{LOAD}}{R_{NL}} \left( V_t \ln \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right) \quad (3.5)$$

In order to get an accurate high-order compensation, at  $T = T_R = 338^\circ\text{K}$  (reference temperature), the linear part of  $V_{BG}$  has to be constant, i.e. it is necessary to satisfied the equation:

$$\frac{\partial V_{BG_{linear}}}{\partial T} = \frac{R_{LOAD}}{R_{PTAT}} \frac{k_B}{q} \ln Q - \frac{R_{LOAD}}{R_{CTAT}} \frac{V_{G0} - V_{BE}(T_R)}{T_R} = 0 \quad (3.6)$$

In this way, it is possible to find the correct value for  $R_{CTAT}$ :

$$R_{CTAT} = R_{PTAT} \frac{V_{G0} - V_{BE}(T_R)}{T_R \frac{k_B}{q} \ln Q} \quad (3.7)$$

Moreover, in order to get a perfect cancellation, it is necessary that:

$$\frac{\eta - 1}{R_{CTAT}} \left( \frac{k_B T}{q} \right) \ln \frac{T}{T_R} = \frac{V_t}{R_{NL}} \ln \frac{2I_{PTAT}}{I_{NL} + I_{const}} \quad (3.8)$$

Therefore it is necessary that the arguments of the two logarithms are the same and then exploit the property of logarithm  $\ln(1) = 0$  in order to check if the compensation has been done correctly. At  $T = T_R$  the following equation has to be satisfied:

$$\left[ \ln \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right]_{T=T_R} = \left[ \ln \frac{T}{T_R} \right]_{T=T_R} = 0 \quad (3.9)$$

Eq. 3.9 is satisfied only if at  $T = T_R$  the following equation is satisfied:

$$2I_{PTAT} = I_{NL} + I_{const} \quad (3.10)$$

Eq. 3.10 can be simply checked by observing the voltage across  $R_{NL}$ . In fact, when  $2I_{PTAT} = I_{NL} + I_{PTAT} + I_{CTAT}$ , the voltage drop across  $R_{NL}$  is zero, (i.e.  $V_{NL} = 0$ ).

Moreover the arguments of the two logarithms of eq. 3.8 are the same, if their derivative is the same:

$$\frac{\partial \left( \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right)}{\partial T} = \frac{1}{T_R} = \frac{1}{338^\circ\text{K}} = 2.95 \times 10^{-3} / ^\circ\text{K} \quad (3.11)$$

$R_{NL}$  can be simply calculated as:

$$R_{NL} = \frac{R_{CTAT}}{\eta - 1} \quad (3.12)$$

if equations 3.9, 3.10 and 3.11 are satisfied.

In order to study step by step the non-linear compensation, it is useful an ideal schematic with ideal current mirror, realized with CCCS:

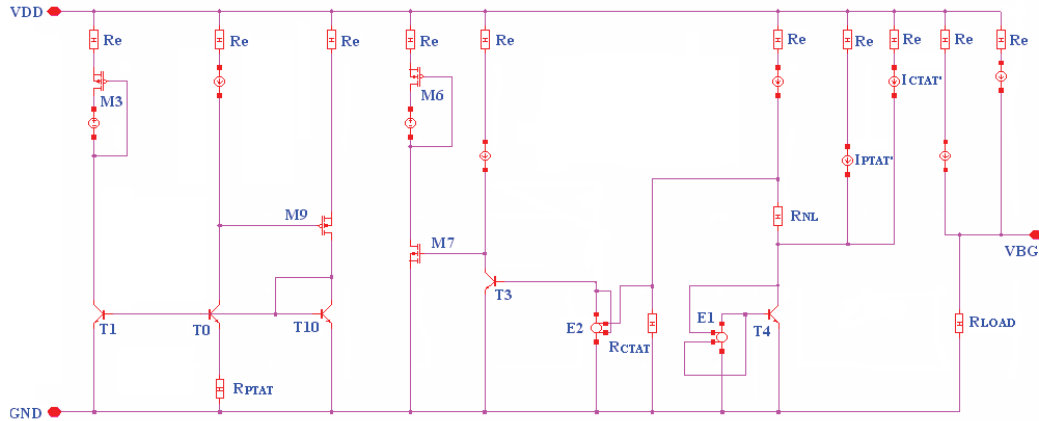


Figure 3.2: Ideal Diode Loop

where  $R_{PTAT} = 60\text{K}\Omega$  and  $Q = \ln \frac{A_{eT0}}{A_{eT1}} = \ln 10 = 2.30$ . In PTAT current has the value of  $1\mu\text{A}$  at room temperature.

The analysis of first-order has shown that base currents may cause errors and loss of precision: in order to avoid these negative effects, it is possible to use ideal “operational amplifier” (E1 and E2 realized with E-gain Cadence block) to deliver the base currents of T4 and T3.

referring to eq. 3.9, it is fundamental that  $\left[ \ln \frac{T}{T_R} = \ln \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right]_{T=T_R} = 0$  to get an accurate compensation.  $I_{const}$  is given by the sum of  $I_{PTAT}$  and  $I_{CTAT}$ , but to satisfied the equation it is necessary to change the mirror ratio with a tuning process, getting the currents  $I'_{CTAT}$  and  $I'_{PTAT}$ . The equation has a recursive nature, so it is impossible to solve it analytically. The accuracy of our tuning process can be evaluated by the observation of drop voltage across  $R_{NL}$ . If the gain of current mirrors is chosen correctly, voltage drop across  $R_{NL}$  (i.e.  $V_{NL}$ ) will be zero at  $T = T_R = 338^\circ\text{K}$ . Moreover if  $R_{NL}$  changes, output voltage “rotates” around  $T = T_R = 338^\circ\text{K}$ .

Using an ideal circuit, it is possible to find the following values in order to satisfy all the equations:

$$R_{PTAT} = 60\text{K}\Omega \quad R_{CTAT} = 468\text{K}\Omega \quad R_{NL} = 134\text{K}\Omega \quad \frac{I'_{CTAT}}{I_{CTAT}} = 1 \quad \frac{I'_{PTAT}}{I_{PTAT}} = 0.65 \quad (3.13)$$

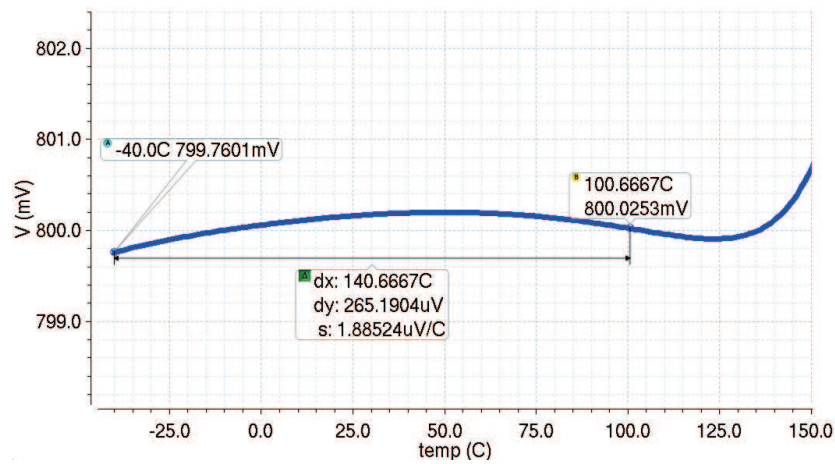


Figure 3.3: Output voltage

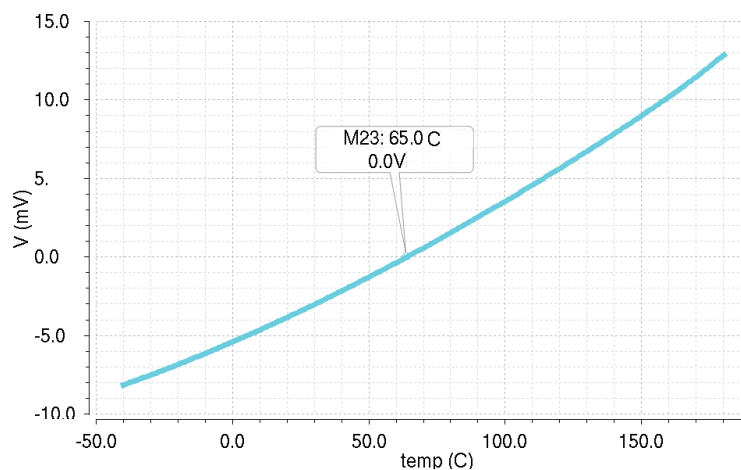


Figure 3.4: Non linear voltage  $V_{NL}$  across  $R_{NL}$

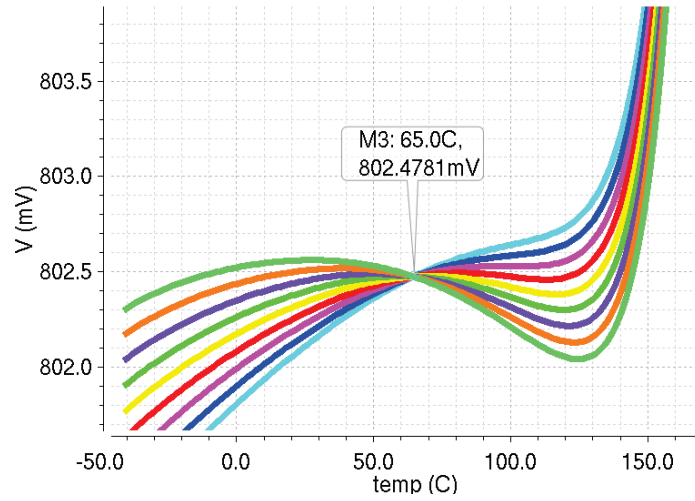


Figure 3.5: Changing  $R_{NL}$  output voltage “rotates” around  $T = T_R$

### 3.3 Evaluation of current mismatch

#### 3.3.1 Output currents mirrored

The PTAT current and the CTAT current are generated with great accuracy by using mosfet in order to eliminate the errors due to base currents of BJT. However in order to get the badgap output voltage, it is necessary to mirror PTAT current and CTAT current with great precision and sum them through an output resistance. It is possible to evaluate the mirror ratio using real current mirrors. Referring to fig. 3.6

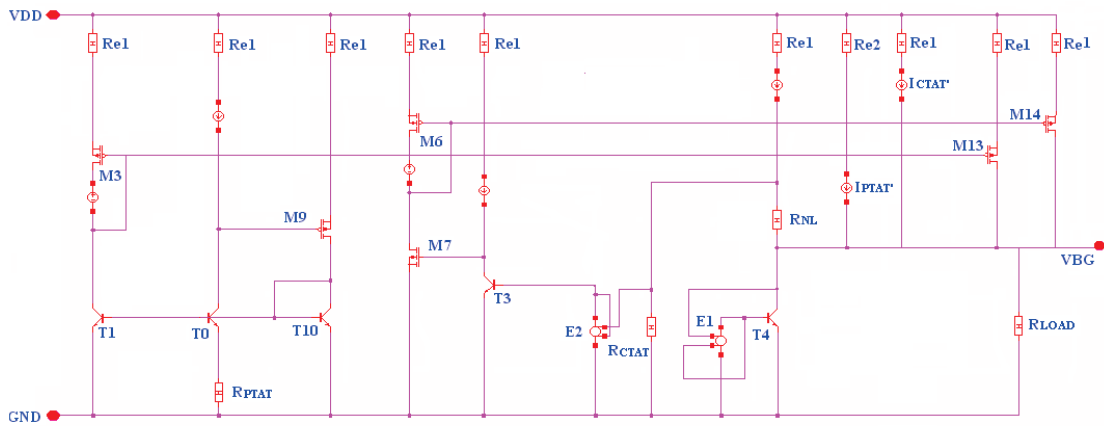


Figure 3.6: Evaluation of real mirror ratio in  $I_{OUT} = I_{PTAT} + I_{CTAT} + I_{NL}$  generation

The ratio  $\frac{I_{CM14}}{I_{CM6}}$  and the ratio  $\frac{I_{CM13}}{I_{CM3}}$  fig. 3.21 and fig. 3.8 can be evaluated by Cadence:



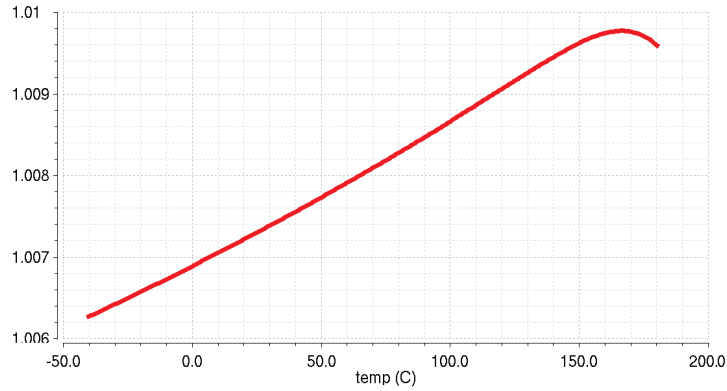


Figure 3.7: Mismatch between CTAT collector current of M14 and CTAT collector current of M6 (ideal  $\frac{I_{C_{M14}}}{I_{C_{M6}}} = 1$ )

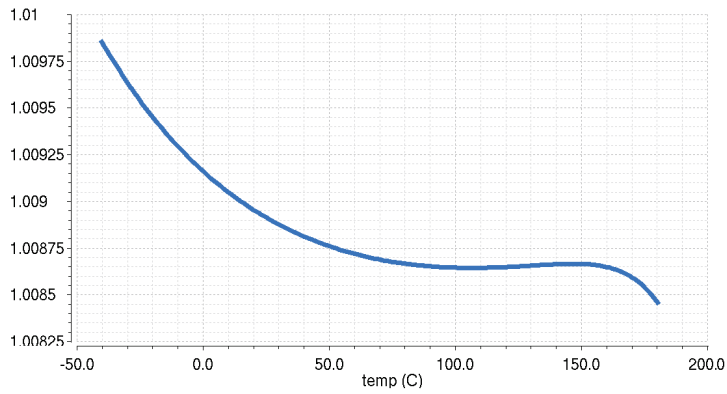


Figure 3.8: Mismatch between PTAT collector current of M13 and PTAT collector current of M3 (ideal  $\frac{I_{C_{M13}}}{I_{C_{M3}}} = 1$ )

Collector current of M14 and collector current of M6 have a mismatch of  $13 \div 14$  nA (room temperature), while collector current of M9 and collector current of M3 have a mismatch of  $15 \div 16$  nA (room temperature). The error in mirrored currents is about  $1.3\% \div 1.6\%$ .

### 3.3.2 Currents mirrored to generate $I_{const}$

Voltage drop across  $R_{NL}$  is zero at  $T = T_R$ . However if equations 3.13 are not satisfied, the value of  $T_R$  changes. In other words, if  $I_{PTAT}^*$  and  $I_{CTAT}^*$  are bad mirrored the temperature  $T_R'$ , at which drop voltage across  $R_{NL}$  is zero, is different from  $T_R = 338^\circ\text{K}$ .

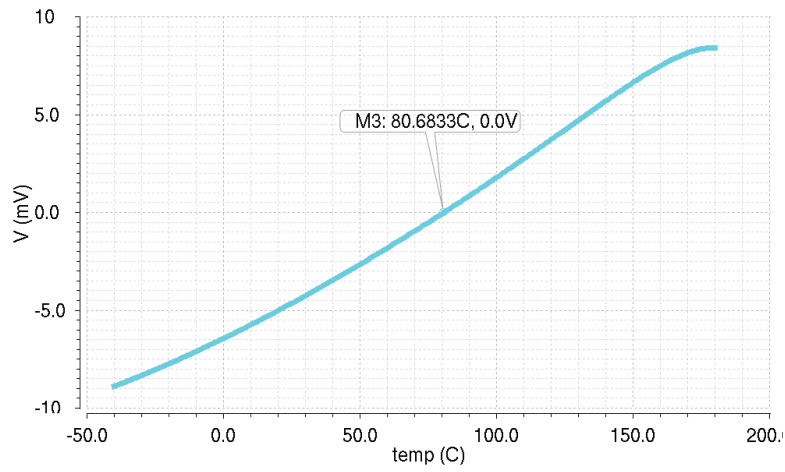


Figure 3.9: Wrong compensation temperature

It is fundamental that the currents are mirrored with the correct ratio. It is interesting to evaluate how the PTAT current and the CTAT current are mirrored by real current mirrors and if the relations about current eq. 3.13 are really satisfied. Referring to fig. 3.10

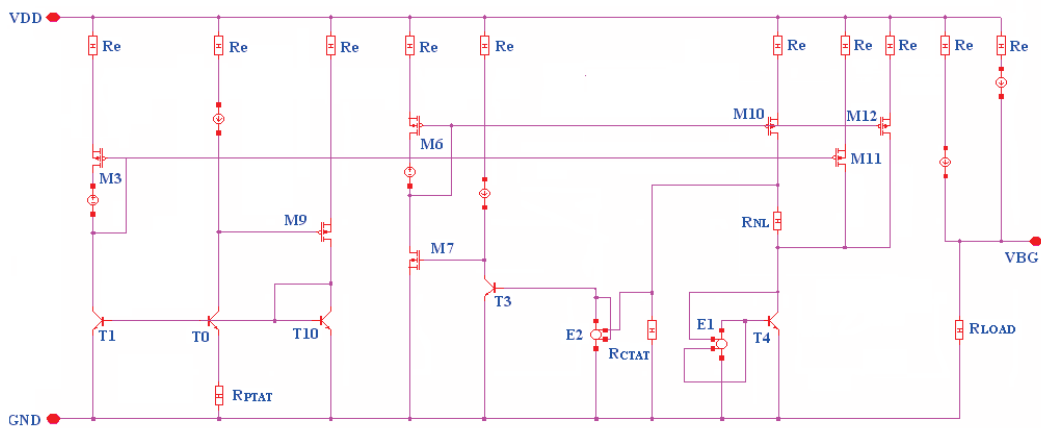


Figure 3.10: Evaluation of real mirror ratio in  $I_{const}$  generation

voltage-mode output is characterized by the sum of temperature-dependent

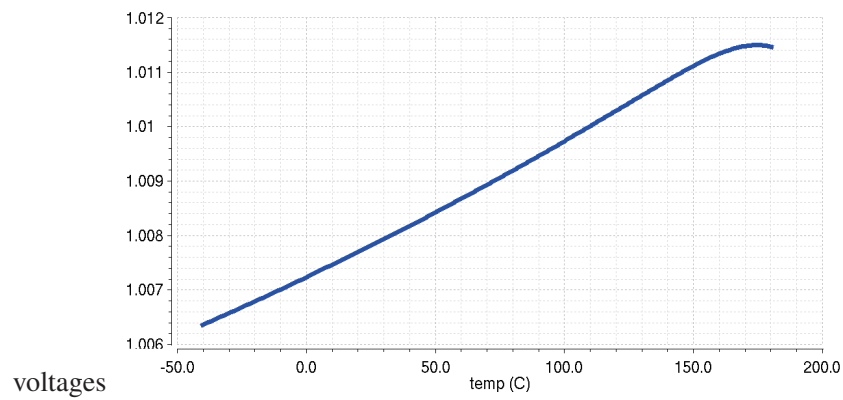


Figure 3.11: Mismatch between collector current of M10 and collector current of M6 (ideal  $\frac{I_{C_{M10}}}{I_{C_{M6}}} = 1$ )

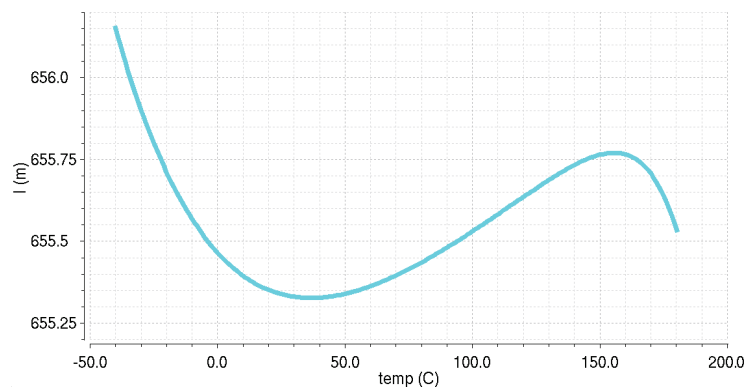


Figure 3.12: Mismatch between collector of M11 and collector current of M3 (ideal  $\frac{I_{C_{M11}}}{I_{C_{M3}}} = 0.65$ )

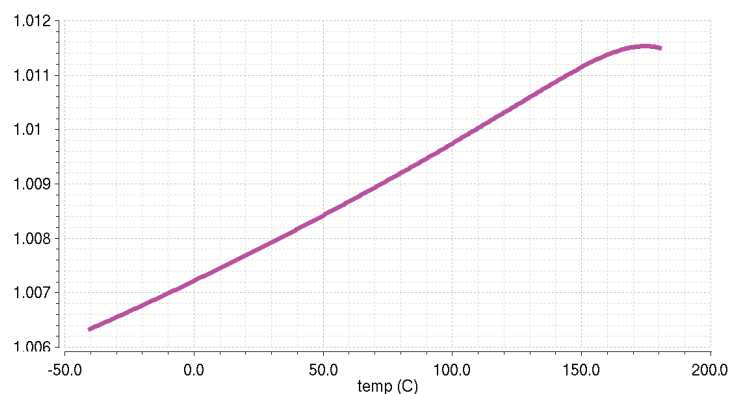


Figure 3.13: Mismatch between collector of M12 and collector current of M6 (ideal  $\frac{I_{C_{M12}}}{I_{C_{M6}}} = 1$ )

Collector current of M10 and collector current of M6 have a mismatch of  $6 \div 12$  nA (room temperature), i.e. the current is mirrored by M10-M6 with an error of  $0.6\% \div 1.2\%$ .

Collector collector current of M11 and collector current of M3 have a mismatch of  $5 \div 6$  nA (room temperature), i.e. the current is mirrored by M11-M3 with an error of  $0.77\% \div 0.92\%$ .

Collector collector current of M12 and collector current of M6 have a mismatch of  $6 \div 12$  nA (room temperature), i.e. the current is mirrored by M12-M6 with an error of  $0.6\% \div 1.2\%$ .

### 3.4 Improving current mirrors

Current mismatches are introduced by mosfet current mirrors. When the mosfet work in saturation region their current can be described by the simplified equation:

$$I_{DS} = \frac{1}{2}k\frac{W}{L}(V_{GS} - V_{TH})^2 \quad (3.14)$$

It is possible to describe the current gain of a current mirror, realized with two mosfet M1 and M2, as:

However when the mosfet work in saturation region their current isn't constant, as predicted by eq. 3.14, but slightly depends on the drain - source voltage too:

$$I_{DS} = \frac{1}{2}k\frac{W}{L}(V_{GS} - V_{TH})^2(1 + \lambda V_{DS}) \quad (3.15)$$

where  $\lambda$  is a technological parameter.

This effect is known as channel length modulation. This happens because if  $V_{DS}$  increases, the reverse voltage of Drain-Substrate junction increases too. Increasing the reverse voltage, space charge region increases and moves pinch-off point toward source. Consequently the channel decreases in size and there is a reduction of its resistance. However the drop voltage through drain and source remains constant and the current in the channel increases.

The inverse ratio of the slope of  $I_{DS}$ , i.e.  $\left(\frac{\partial I_{DS}}{\partial V_{DS}}\right)^{-1} = \left(\frac{1}{\frac{1}{2}k\frac{W}{L}(V_{GS} - V_{TH})^2\lambda}\right)$ , is a resistance:

$$r_0 = \left(\frac{\partial I_{DS}}{\partial V_{DS}}\right)^{-1} = \left(\frac{1}{\frac{1}{2}k\frac{W}{L}(V_{GS} - V_{TH})^2\lambda}\right) \quad (3.16)$$

$r_0$  represents the output resistance of transistor. Ideally if  $\lambda = 0$   $r_0 = \infty$  and  $I_{DS}$  doesn't depend on  $V_{DS}$ . However  $r_0$  always has a finite value. For reducing the dependence of  $I_{DS}$  from  $V_{DS}$  it is necessary to improve the output resistance of current mirror.

It is possible to calculate  $\lambda$ , by evaluating  $I_{DS} = f(V_{GS}, V_{DS}, \lambda)$  with a parametric analysis:

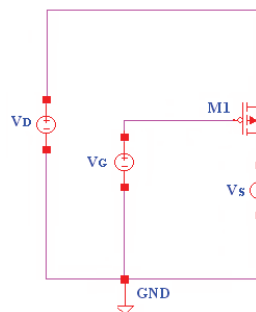


Figure 3.14: Circuit used to evaluate  $I_{DS} = f(V_{GS}, V_{DS}, \lambda)$

In this technology  $\lambda$  is not a constant parameter and it depends also by bias point ( $-0.169V^{-1} \leq \lambda \leq -0.131V^{-1}$ ).

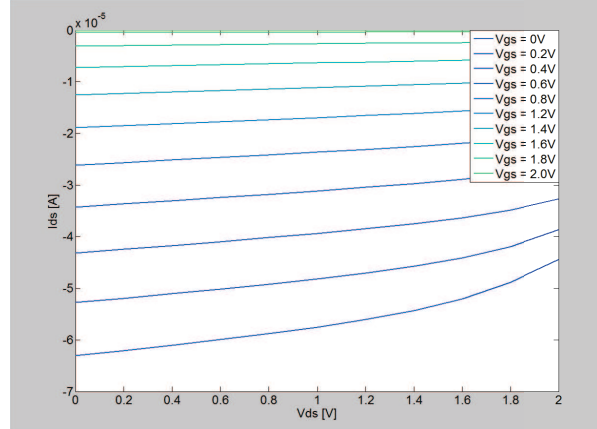


Figure 3.15:  $I_{DS} = f(V_{GS}, V_{DS}, \lambda)$

For a given gate-source voltage eq. 3.15 can be written as:

$$I_{DS} = mV_{DS} + q \quad (3.17)$$

where  $m = \frac{1}{2}k\frac{W}{L}(V_{GS} - V_{TH})^2\lambda$  and  $q = \frac{1}{2}k\frac{W}{L}(V_{GS} - V_{TH})^2$ . Using a matlab script it is possible to find the coefficients ( $m$  and  $q$ ) of a polynomial  $P(X)$  of degree  $N = 1$  (i.e. a trendline) that fits  $I_{DS}$  best in a least-squares sense.

$\lambda$  can be evaluated as:

$$\lambda = \frac{m}{q} \quad (3.18)$$

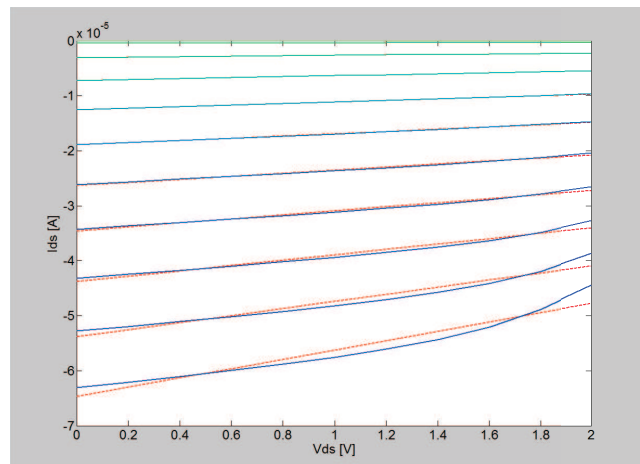


Figure 3.16: Rectification of  $I_{DS} = f(V_{GS}, V_{DS}, \lambda)$  with a trendline

In order to reduce the dependence of  $I_{DS}$  from  $V_{DS}$  it is necessary to improve  $r_0$ : for this purpose it is possible to use cascode current mirrors. The cascode is a two-stage amplifier composed of a transconductance amplifier (common source) followed by a current buffer (common gate).

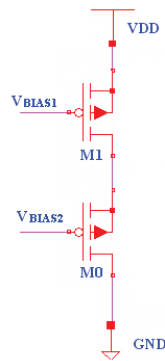


Figure 3.17: Cascode

Compared to a single amplifier stage, this combination has a high output impedance:

$$Z_{out} = r_{0M1} + r_{0M0} + g_{mM0}r_{0M0}r_{0M1} \simeq g_{mM0}r_{0M0}r_{0M1} \quad (3.19)$$

Cascode is a good solution to improve the precision of the mirrors. Figures 3.21, 3.22, 3.23, 3.24, 3.25, compare the mismatch between collector currents of mosfet before using cascode and after using cascode. The current mismatch without using cascode is represented with a green line. However using cascode current mirrors headroom is the tradeoff: minimum voltage supply increases. Moreover it is necessary to find a solution to generate the bias voltage  $V_A$  fig. 3.19.

In order to generate voltage  $V_A$  it is necessary to pay attention to the drain voltage of M18 and M19: the drain voltage of these transistors is our bandgap output voltage. Mosfet have to be in saturation region to work correctly, i.e. for each of them  $|V_{DS}| \geq |V_{GS} - V_{TH}|$  and  $V_S > V_{BG} = 800mV$ . It is possible to use a simple a simple first-order voltage reference in order to generate the correct bias for cascode:

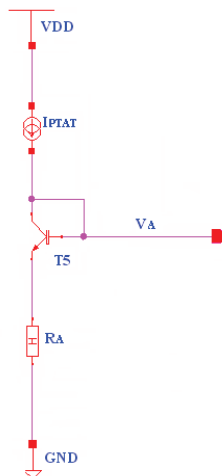


Figure 3.18: Bias for cascode

Referring to fig. 3.19, it is possible to write:

$$R_A I_{PTAT_{max}} + V_{BE_{T5_{min}}} - V_{GD_{M8}} = R_A I_{PTAT_{max}} + V_{BE_{T5_{min}}} - (V_{GS_{M8}} - V_{DS_{M8}}) \geq 800mV \quad (3.20)$$

Eq. 3.20 can be written by adding  $\pm V_{TH}$ :

$$R_A I_{PTAT_{max}} + V_{BE_{T5_{min}}} - V_{GS_{M8}} + V_{DS_{M8}} + V_{TH} - V_{TH} = R_A I_{PTAT_{max}} + V_{BE_{T5_{min}}} - V_{OV_{M8}} + V_{DS} - V_{TH} \geq 800mV \quad (3.21)$$

where  $V_{OV_{M8}} = V_{GS_{M8}} - V_{TH}$

In order to be in saturation region,  $V_{GS}$  and  $V_{DS}$  have to satisfied the following equation:

$$V_{DS_{M8}} \geq V_{GS_{M8}} - V_{TH} \quad (3.22)$$

Choosing  $V_{DS_{M8}} = V_{GS_{M8}} - V_{TH}$ , i.e.  $V_{DS_{M8}} = V_{OV_{M8}}$  eq. 3.21 yields a negative value for  $R_A$ :

$$R_A \geq \frac{800mV - V_{BE_{T5_{min}}} + V_{TH}}{I_{PTAT_{max}}} \quad (3.23)$$

It is possible to use only a BJT to generate  $V_A$ . However it is better to add a resistance because:

- $V_A$  becomes the output voltage of a first-order bandgap reference and it becomes flatter than the simple  $V_{BE_{T5}}$ ;
- $V_A = V_{BE_{T5}} + R_A I_{PTAT}$  lets to improve the mirror ratio of M10-M6, M11-M3, M12-M6, M13-M6, M14-M3 because they will have very similar drain voltages.

The chosen value is  $R_A = 1M\Omega$ . This value lets to improve the mirror ratio of the current mirrors, however head-room is the tradeoff:

$$V_{DD_{min}} = R_e I_{PTAT_{min}} + V_{DS} + R_A I_{PTAT_{min}} + V_{BE_{max}} \simeq 2V + V_{DS} \quad (3.24)$$

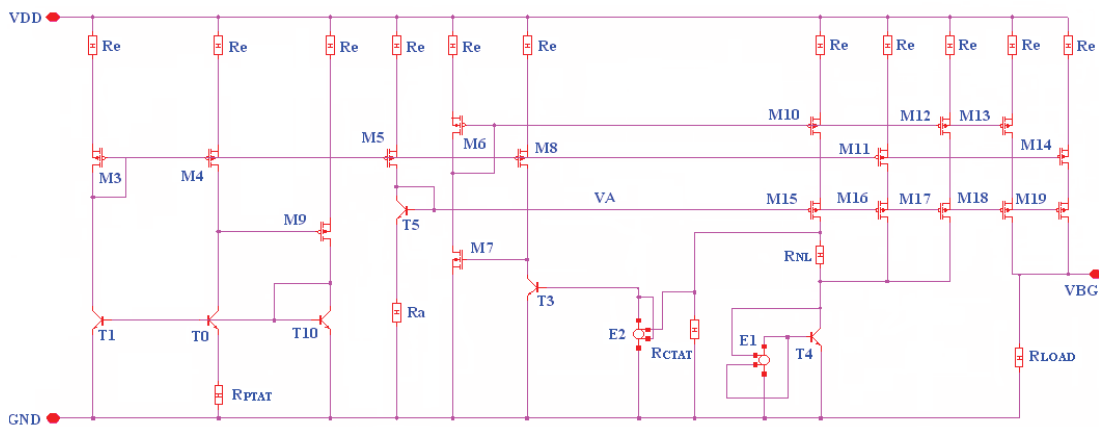


Figure 3.19: Nonlinear-order without base currents

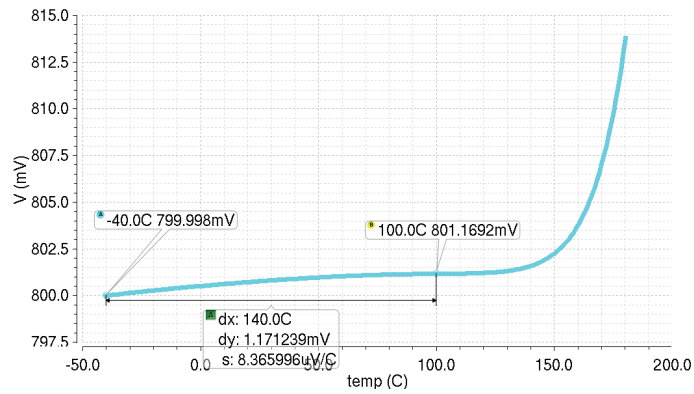


Figure 3.20: Output bandgap voltage

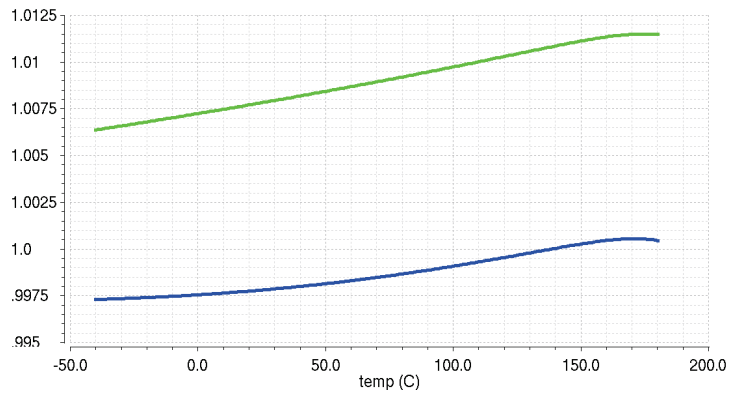


Figure 3.21: Mismatch between collector current of M10 and collector current of M6 (ideal  $\frac{I_{CM10}}{I_{CM6}} = 1$ )

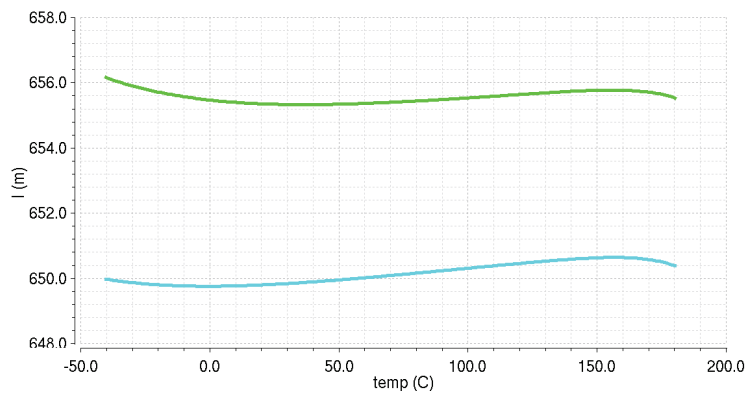


Figure 3.22: Mismatch between collector current of M11 and collector current of M3 (ideal  $\frac{I_{CM11}}{I_{CM3}} = 0.66$ )



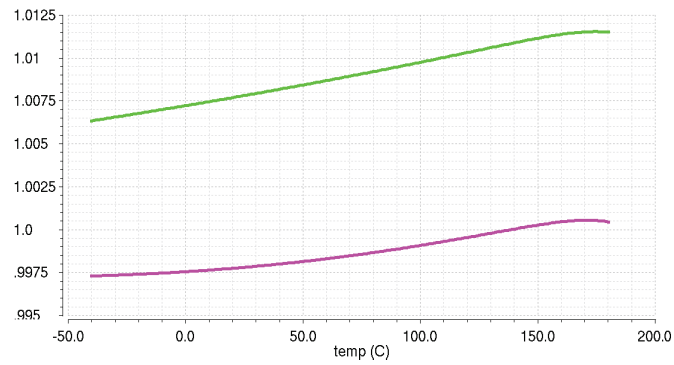


Figure 3.23: Mismatch between collector current of M12 and collector current of M6 (ideal  $\frac{I_{C_{M12}}}{I_{C_{M6}}} = 1$ )

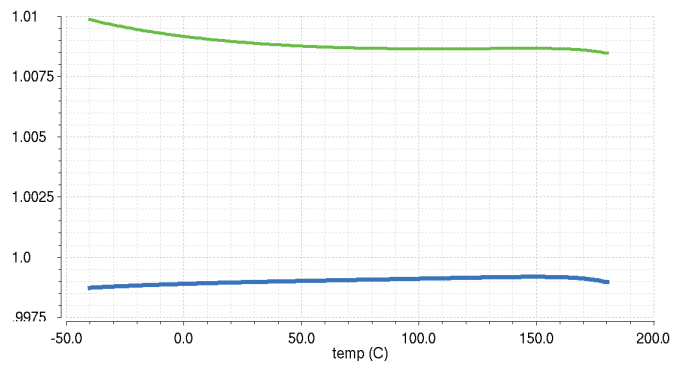


Figure 3.24: Mismatch between the PTAT collector current of M14 and the PTAT collector current of M6 (ideal  $\frac{I_{C_{M14}}}{I_{C_{M6}}} = 1$ )

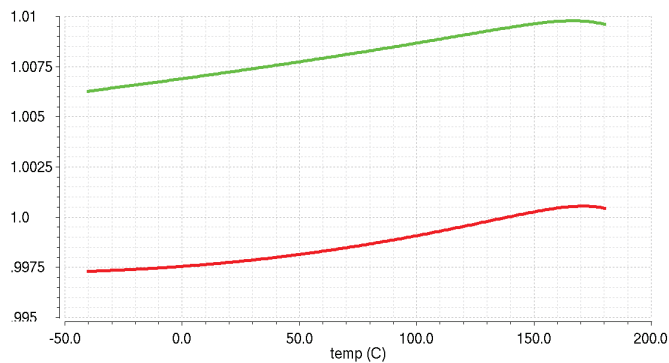


Figure 3.25: Mismatch between the CTAT collector current of M13 and the CTAT collector current of M3 (ideal  $\frac{I_{C_{M13}}}{I_{C_{M3}}} = 1$ )

### 3.5 Effects of base currents

Base currents are considered a parasitic element and a particular circuit has been used to avoid its effects. Is it possible to remove the ideal amplifier (realized with Cadence E-grain block) and evaluate the loss of precision due to the presence of these parasitic currents.

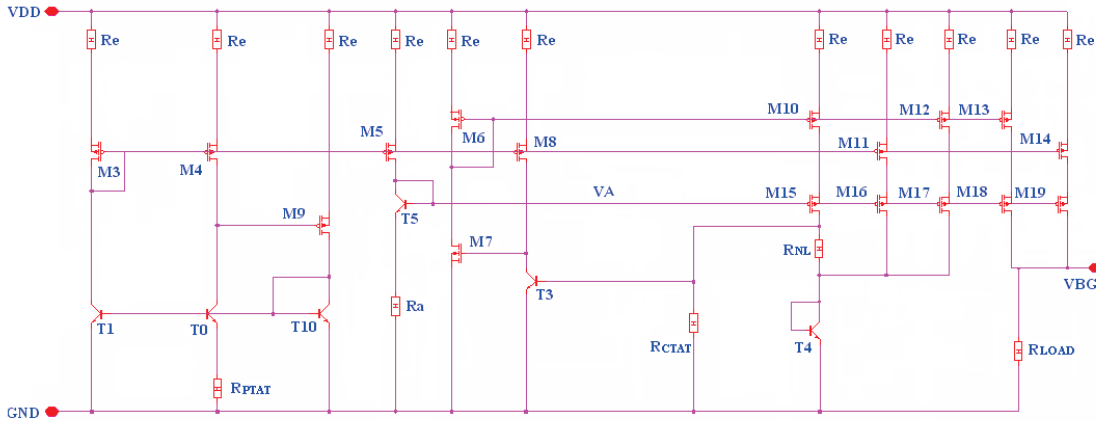


Figure 3.26: Real second-order compensated bandgap voltage reference

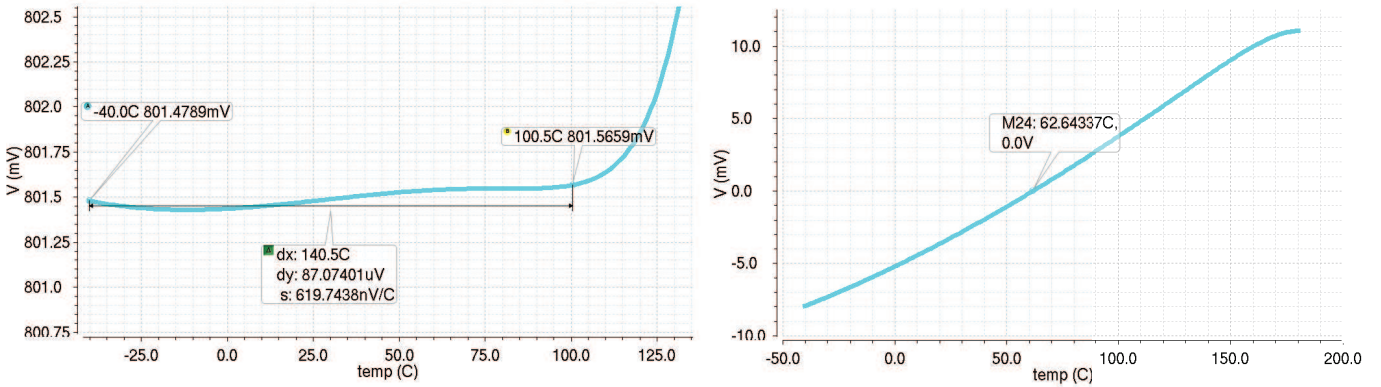


Figure 3.27: Output voltage Non linear voltage  $V_{NL}$  across  $R_{NL}$

By analyzing KVL T1-T0- $R_{PTAT}$ :

$$I_{PTAT} = I_{E_{T0}} = \frac{1}{R_{PTAT}} \frac{k_B}{q} T \ln Q \quad (3.25)$$

The collector current of T0 and T1 can be written as:

$$I_{C_{T0}} = I_{C_{T1}} = I_{E_{T0}} - I_{B_{T0}} = I_{PTAT} - I_{B_{T0}} \quad (3.26)$$

The feedback loop of PTAT stage impose that  $I_{C_{T0}} = I_{D_{M4}} = I_{PTAT} - I_{B_{T0}}$ : the error due to base current is mirrored to the CTAT stage, to the nonlinear stage and to the output stage.

The collector current of T3 is:

$$I_{C_{T3}} = I_{PTAT} - I_{B_{T0}} \quad (3.27)$$

instead of  $I_{C_{T3}} = I_{PTAT}$ .

The drain current of M11 and M16 is:

$$I_{D_{M11}} = I_{D_{M16}} = 0.65(I_{PTAT} - I_{B_{T0}}) \quad (3.28)$$

The drain current of M15 is:

$$I_{D_{M15}} = I_{NL} + I_{CTAT} + I_{B_{T3}} \quad (3.29)$$

$I_{D_{M15}}$  is mirrored at output stage of bandgap reference.

The collector current of T4 is:

$$I_{C_{T4}} = I_{NL} + I_{PTAT}^* + I_{CTAT}^* - I_{B_{T4}} = I_{NL} + 0.65(I_{PTAT} - I_{B_{T0}}) + I_{CTAT} + I_{B_{T3}} - I_{B_{T4}} = I_{NL} + I_{CONST} - 0.65I_{B_{T0}} + I_{B_{T3}} - I_{B_{T4}} \quad (3.30)$$

If  $\delta_{I_B} = I_{B_{T3}} - I_{B_{T4}}$  the collector current of T4 becomes:

$$I_{C_{T4}} = I_{NL} + I_{CONST} - 0.65I_{B_{T0}} + \delta_{I_B} \quad (3.31)$$

The base current of T4 changes the expression of  $I_{NL}$ . Referring to eq. 3.3, the new expression of  $I_{NL}$  becomes:

$$I_{NL} = \frac{V_t}{R_{NL}} \ln \left( \frac{I_{C_{T3}}}{J_S A_{e3}} \frac{J_S A_{e4}}{I_{C_{T4}}} \right) = \frac{V_t}{R_{NL}} \ln \left( \frac{2(I_{PTAT} - I_{B_{T0}})}{I_{NL} + I_{CONST} - 0.65I_{B_{T0}} + \delta_{I_B}} \right) \quad (3.32)$$

In this way the bandgap output voltage  $V_{BG} = R_{LOAD}(I_{PTAT} + I_{CTAT} + I_{NL})$  can be calculated considering also the base currents of T3 and T4:

$$V_{BG} = \frac{R_{LOAD}}{R_{PTAT}} \left( \frac{k_B}{q} T \ln Q \right) - R_{LOAD} I_{B_{T0}} + \frac{R_{LOAD}}{R_{CTAT}} (V_{BE}) + \frac{R_{LOAD}}{R_{NL}} \ln \left( \frac{2(I_{PTAT} - I_{B_{T0}})}{I_{NL} + I_{CONST} - 0.65I_{B_{T0}} + \delta_{I_B}} \right) + R_{LOAD} I_{B_{T3}} \quad (3.33)$$

If  $\Delta I_B = I_{B_{T3}} - I_{B_{T0}}$  it is possible to write  $V_{BG}$  as:

$$V_{BG} = \frac{R_{LOAD}}{R_{PTAT}} \left( \frac{k_B}{q} T \ln Q \right) + \frac{R_{LOAD}}{R_{CTAT}} (V_{BE}) + \frac{R_{LOAD}}{R_{NL}} \ln \left( \frac{2(I_{PTAT} - I_{B_{T0}})}{I_{NL} + I_{CONST} - 0.65I_{B_{T0}} + \delta_{I_B}} \right) + R_{LOAD} \Delta I_B \quad (3.34)$$

In this way it is possible to reduce the effect of base currents of T3 and T4.

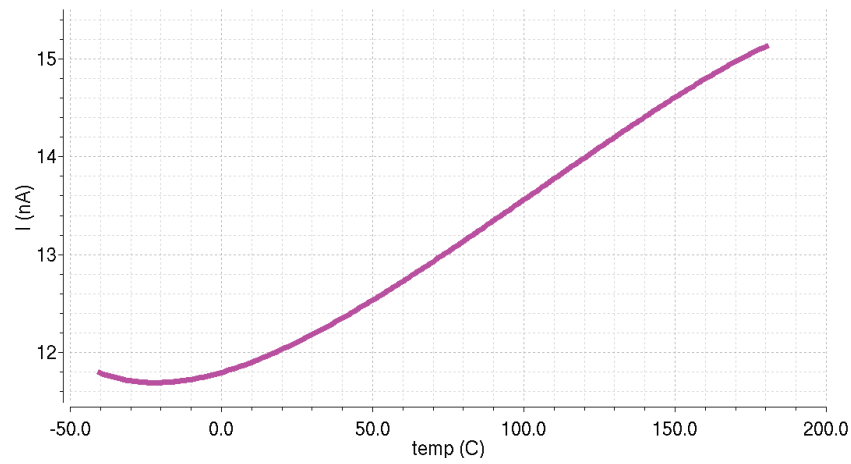


Figure 3.28: Base current of T0

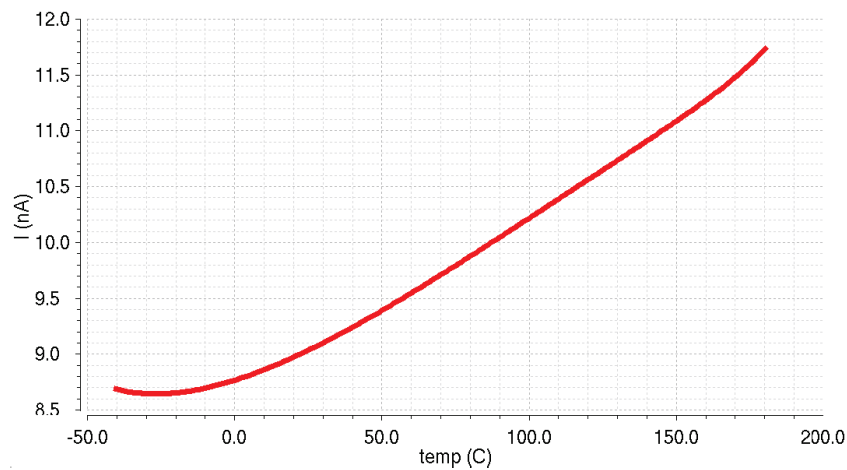


Figure 3.29: Base current of T3

### 3.6 Stabilization of non-linear order bandgap reference: dominant pole compensation

It is fundamental that the precise nonlinear-order voltage reference is stable, like the first-order voltage reference of the previous chapter. It is possible to guarantee the asymptotic stability with a dominant pole gain loop with a phase margin of  $75^\circ$  for PTAT stage and CTAT stage at all the temperatures.

PTAT stage of first-order bandgap reference and PTAT stage of diode loop are the same: the capacitor and the resistance used to stabilize first-order PTAT loop can be used also to stabilize the non-linear PTAT loop. If  $C_{adj} = 1pF$  and  $R_{adj} = 160k\Omega$  are chosen, there is a dominant pole around  $f_{dp} = 1kHz$ , a phase margin of  $85^\circ$  a gain margin of  $14.854dB$  and a bandwidth of  $374kHz$ .

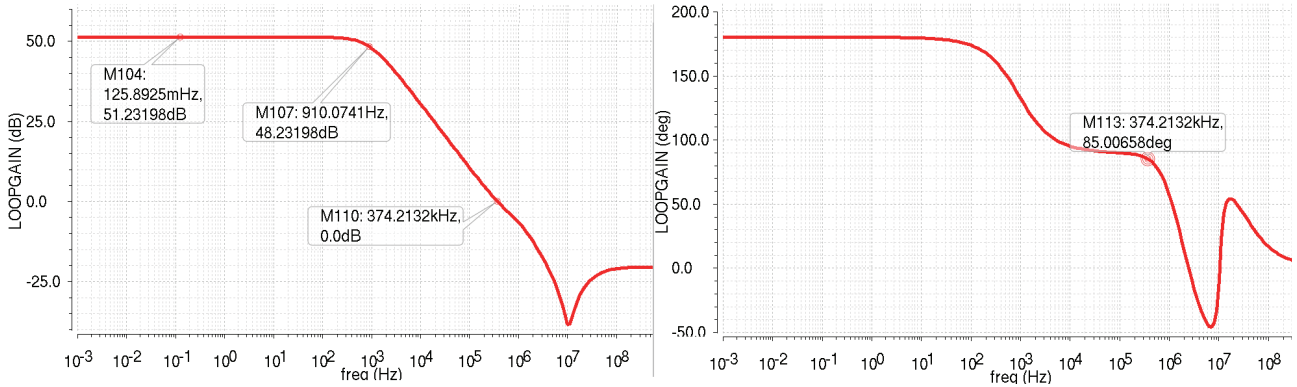


Figure 3.30: PTAT loop gain after dominant pole compensation

Using these values for  $C_{adj}$  and  $R_{adj}$  PTAT loop is stable at all temperatures.

CTAT stage of Diode Loop is a bit different from CTAT stage of first-order, because of the non-linear stage. It is possible to analyze CTAT stage with a Cadence stb analysis:

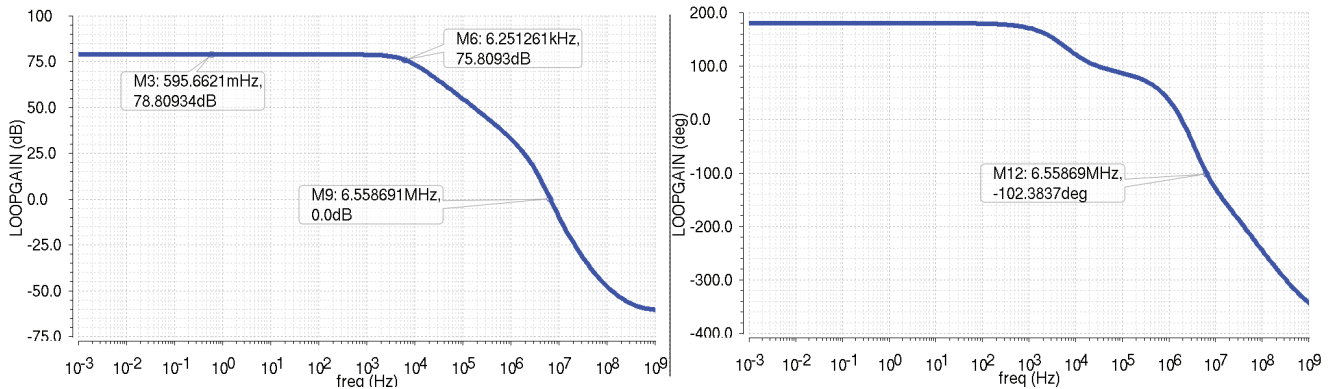


Figure 3.31: CTAT loop gain before dominant pole compensation

CTAT gain loop, without compensation, has a negative phase margin and so CTAT feedback loop is not asymptotically stable. A phase margin of  $50^\circ$  is obtained by using the same value to compensate the CTAT stage of the first-order bandgap voltage reference ( $C_{comp} = 1 pF$  and  $R_{comp} = 66 K\Omega$ ).

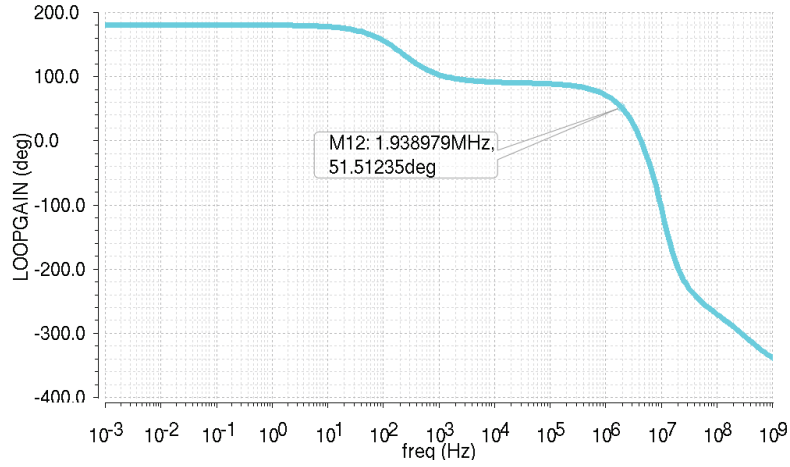


Figure 3.32: Phase margin of CTAT loop using  $C_{comp} = 1pF$  and  $R_{comp} = 66K\Omega$

A phase margin of  $\sim 75^\circ$  and a positive gain margin guarantee the asymptotic stability of the system at all temperatures also with layout parasitic elements. Referring to fig.3.26, it is possible to identify the frequency of dominant pole ( $f_{dp}$ ) and the frequency of the first non dominant pole ( $f_{ndp}$ ):

$$f_{dp} = \frac{1}{2\pi R_{dp} C_{cdp}} \simeq 6.25KHz \quad f_{ndp} = \frac{1}{2\pi R_{ndp} C_{ndp}} \simeq 1.724MHz \quad (3.35)$$

where:

$$R_{dp} = R_{CTAT} // r_{\pi 2} // \left( \frac{1}{g_{m4}} + R_{NL} \right) \quad C_{dp} = C_{\mu 3} g_{m3} [(R_e + r_{o8}) // r_{oT3}]$$

$$R_{ndp} = (r_{o8} + R_e) // r_{o3} \quad C_{ndp} = C_{d8} + C_{\mu 3} + C_{gs7} + C_{gd7} g_{m5} (R_e + \frac{1}{g_{m6}})$$

In order to get a higher phase margin, it is necessary to split dominant pole at lowest frequency. Using a capacitor  $C_{comp} = 6pF$  between base and collector of T3 (so Miller's Effect will boost its value) and a nulling resistor  $R_{comp} = \frac{1}{g_{m3}} = 27.624K\Omega$  (to eliminate the zero with real positive part) loop gain has  $79^\circ$  of phase margin and  $19.23dB$  of gain margin .

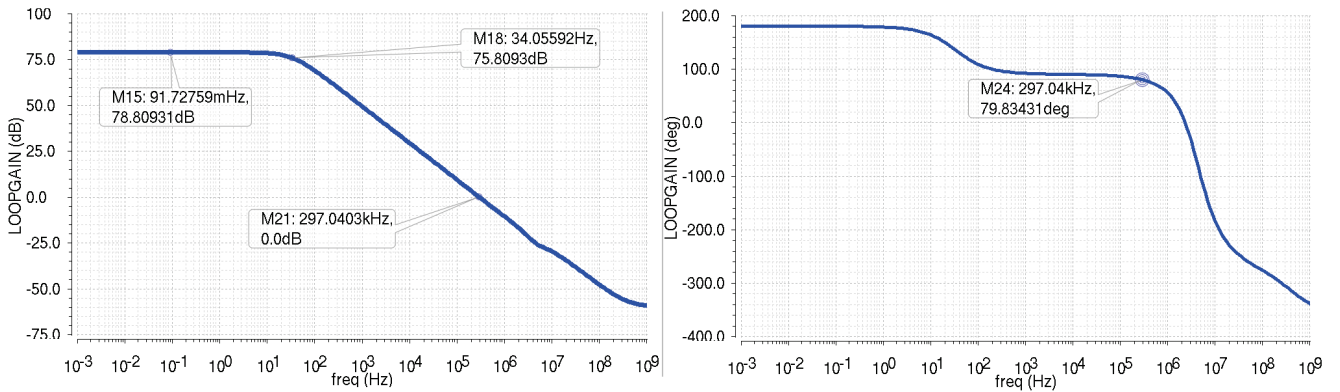


Figure 3.33: CTAT loop gain after dominant pole compensation

Using a Montecarlo Analysis, it is possible to evaluate gain margin and phase margin of CTAT loop at three different temperatures, to verify the asymptotic stability of the loop.

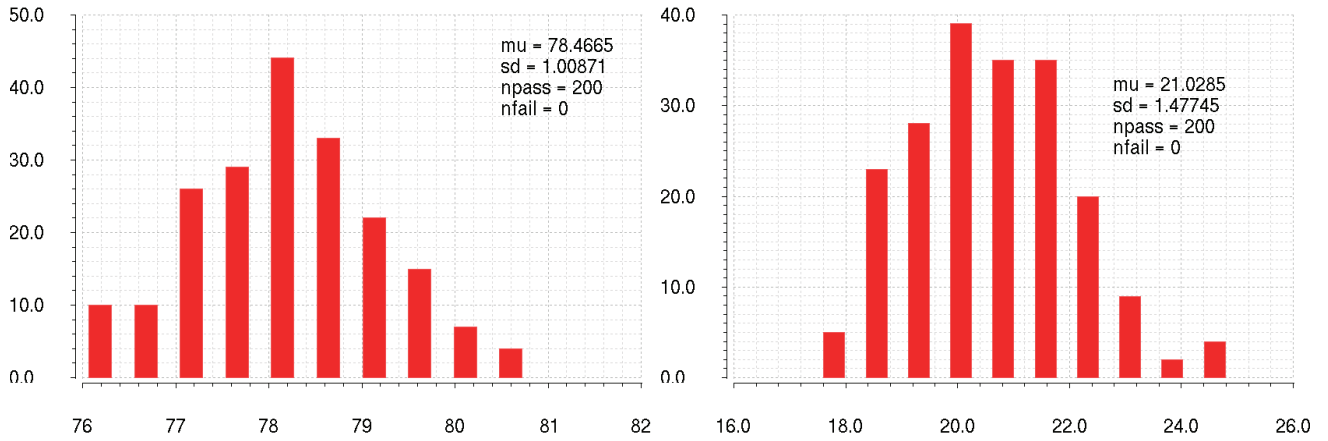


Figure 3.34: Phase margin and Gain margin of CTAT feedback loop at  $-40^{\circ}\text{C}$

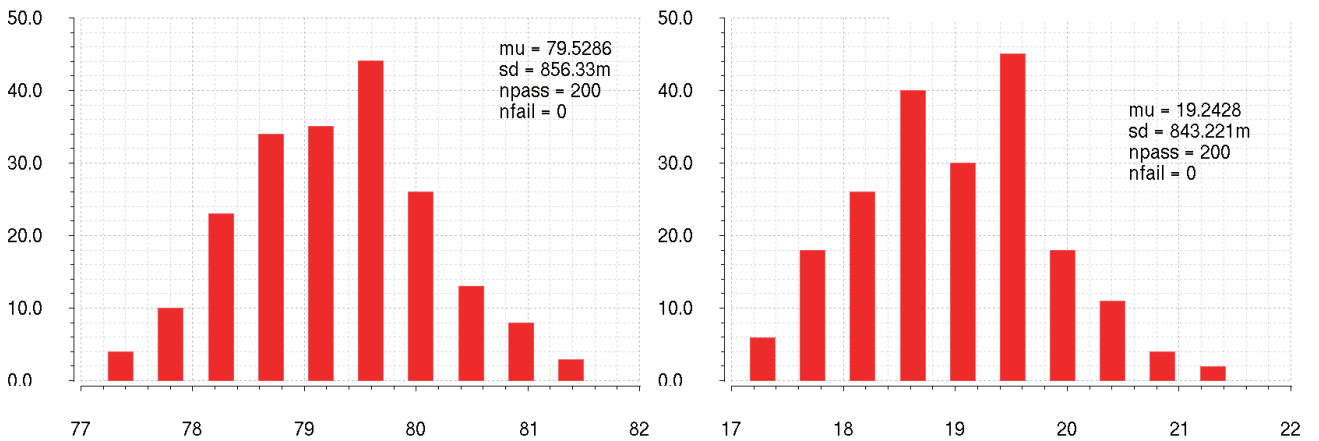


Figure 3.35: Phase margin and Gain margin of CTAT feedback loop at  $27^{\circ}\text{C}$

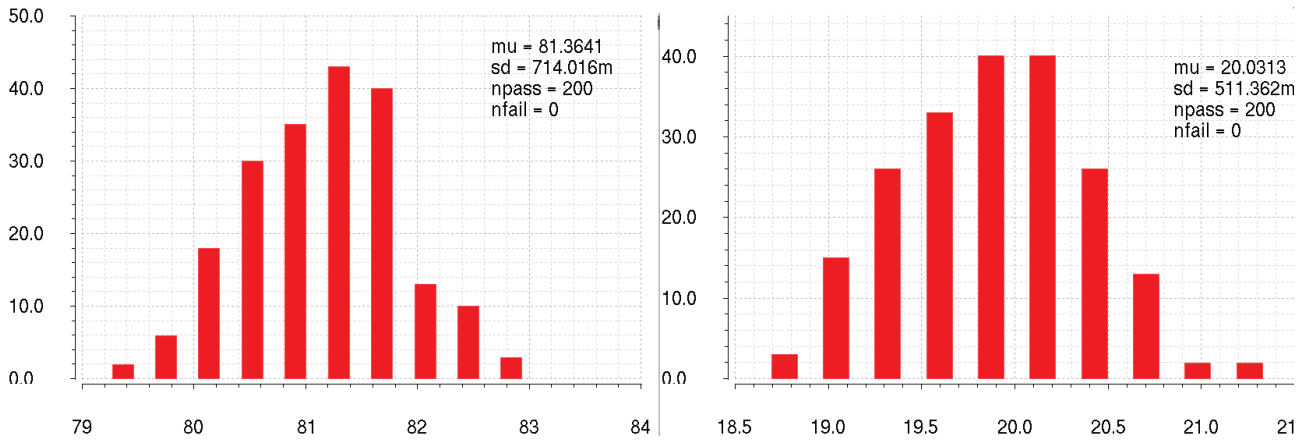


Figure 3.36: Phase margin and Gain margin of CTAT feedback loop at 160°C

### 3.6.1 Transient response with $V_{DD}$ steps

A transient analysis can be used in order to evaluate the step response of nonlinear-order bandgap reference. Transient analysis has been run at three different temperatures: low temperature ( $-40^{\circ}\text{C}$ ), room temperature ( $27^{\circ}\text{C}$ ) and high temperature ( $160^{\circ}\text{C}$ ).

In this way the transient response of the circuit has also been verified at the typical conditions. Power supply has been ramped up to the final value for this purpose ( $t_{rise} = 1\text{ns}$ ). The corresponding settling of  $V_{BG}$  is shown in fig. 3.37, 2.55 and 3.39.  $V_{BG}$  has been observed to settle down to its final value without any oscillations as  $V_{dd}$  ramps up in  $1\text{ns}$ .

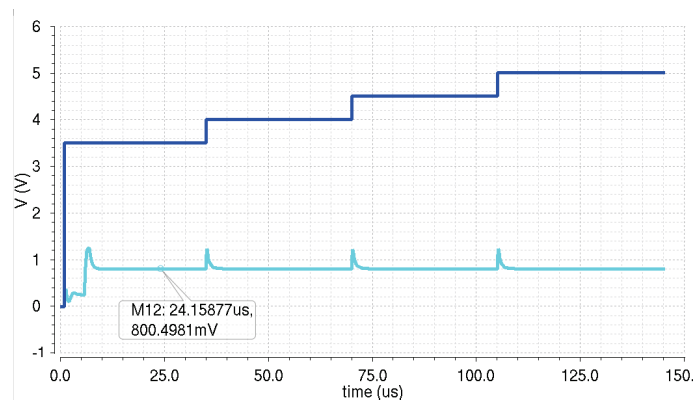


Figure 3.37: Step response of the first-order compensated voltage bandgap generator at  $T = -40^{\circ}\text{C}$



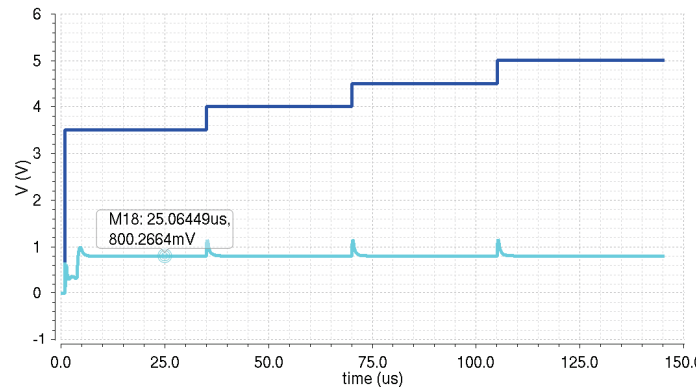


Figure 3.38: Step response of the first-order compensated voltage bandgap generator at  $T = 27^\circ\text{C}$

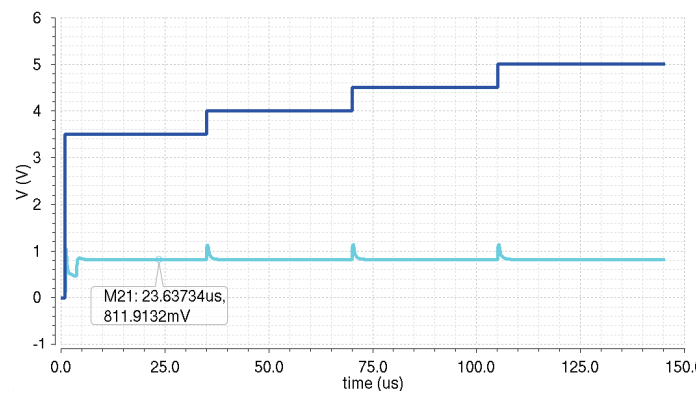


Figure 3.39: Step response of the first-order compensated voltage bandgap generator at  $T = 160^\circ\text{C}$

## Chapter 4

# TRIMMING NETWORKS

### 4.1 Analysis of Process-Induced Errors

Integrated references suffer from practical nonidealities associated with the fabrication process in any technology. These parasitic effects can degrade the accuracy of the most well-designed bandgap reference; they manifest themselves in the form of current mirror mismatches, resistor tolerance, resistor temperature coefficient (TC), resistor mismatches, Early Voltage, transistor mismatches, package shifts and input voltages offset.

Some of these errors are systematic and effectively predicted by the simulator, like TC and Early Voltage. However device mismatches, tolerance and package effects are random in nature and they will therefore vary from chip to chip and from wafer to wafer. Consequently, the robustness of a particular design depends on its susceptibility to process and package effects.

It is possible to analyze the effects of process and mismatch and to their impact on the accuracy of PTAT current and base-emitter voltage generation. In the following analysis of the effects of process and mismatch,  $I_{PTAT}$  and  $V_{BE}$  will be the physical quantities generate by the ideal circuit (i.e. without mismatch effects or errors), while  $I'_{PTAT}$  and  $V'_{BE}$ , as the same quantities generate by real circuit (i.e. considering one or more effects of mismatch and errors).  $I_{PTAT}$  and  $V_{BE}$  are the two fundamental bricks of all bandgap voltage reference: if process and mismatch change them, also output voltage, compensation temperature  $T_R$ , current consumption will change.

#### 4.1.1 Current mirror mismatch

referring to fig. 3.26, a mismatch in the current flowing through the collectors of transistors T1 and T0 affects the base-emitter voltage as well as the PTAT current. It is possible to quantify the mismatch of the mirror:

$$I_{C_{T1}} = I_{C_{T0}} (1 + \delta_M) \quad (4.1)$$

where  $\delta_M$  is the percent current mismatch between the collector currents in transistor T1 and T0. This mismatch involves the base-emitter voltage as well as the PTAT voltage term.

Effect of current mirror mismatch on PTAT stage:

Using eq. 4.1, it is possible to analyze eq. 2.5 the equation of PTAT current:

$$I'_{PTAT} = \frac{V_t}{R_{PTAT}} \ln \left[ \frac{I_{C_{T0}} (1 + \delta_M) Q}{I_{C_{T0}}} \right] = \frac{V_t}{R_{PTAT}} \ln Q \left[ 1 + \frac{\ln(1 + \delta_M)}{\ln Q} \right] = I_{PTAT} \left[ 1 + \frac{\ln(1 + \delta_M)}{\ln Q} \right] = I_{PTAT} \left[ 1 + \frac{\delta_M}{\ln Q} \right] \quad (4.2)$$

The final approximation is true only if  $\delta_M$  and  $\frac{\delta_M}{\ln Q}$  are much less than 1.

Effect of current mirror mismatch on CTAT stage:

The effects of this mismatch affects also base-emitter voltage of T3. Using eq. 2.4:

$$V_{BE_{T3}} = V_t \ln \left( \frac{I_{PTAT}}{J_s A_e} \right) \quad V'_{BE_{T3}} = V_t \ln \left( \frac{I'_{PTAT}}{J_s A_e} \right) \quad (4.3)$$

The mismatch between  $V_{BE_{T3}}$  and  $V'_{BE_{T3}}$  can be quantify as  $V_t \frac{\delta_M}{\ln Q}$  by the following equation:

$$V'_{BE_{T3}} = V_{BE_{T3}} + (V'_{BE_{T3}} - V_{BE_{T3}}) = V_{BE_{T3}} + V_t \ln \left( \frac{I'_{PTAT}}{I_{PTAT}} \right) = V_{BE_{T3}} + V_t \ln \left( 1 + \frac{\ln(1 + \delta_M)}{\ln Q} \right) \simeq V_{BE_{T3}} + V_t \frac{\delta_M}{\ln Q} \quad (4.4)$$

Effect of current mirror mismatch on Output stage:

The effect of current mirror mismatch on bandgap output voltage can be evaluated by the equation:

$$V'_{BG} = R_{LOAD} \left( I'_{PTAT} + \frac{V'_{BE_{T3}}}{R_{CTAT}} \right) = R_{LOAD} \left[ I_{PTAT} \left( 1 + \frac{\delta_M}{\ln Q} \right) + \frac{1}{R_{CTAT}} \left( V_{BE_{T3}} + V_t \frac{\delta_M}{\ln Q} \right) \right] = V_{BG} + \frac{\delta_M}{\ln Q} R_{LOAD} \left( \frac{V_t}{R_{CTAT}} + I_{PTAT} \right) \quad (4.5)$$

#### 4.1.2 Transistor mismatch

The area ratio between T1 and T0 has also some errors. These errors ( $\delta_{NPN}$ ) manifest themselves by yielding an effective gain-ratio of  $(1 + \delta_{NPN}) Q$ . The effective current gain between T1 and T0 is  $(1 + \delta_{NPN}) Q$ .

Effect of transistor mismatch on PTAT stage:

$$\Delta V_{BE} = V_{BE_{T0}} - V_{BE_{T1}} = V_t \ln \frac{I_{CT0} J_s A_{eT0}}{I_{CT1} J_s A_{eT1}} = V_t \ln \frac{A_{eT0}}{A_{eT1}} = V_t \ln [(1 + \delta_{NPN}) Q] \simeq V_t (\delta_{NPN} + \ln Q) = \Delta V_{BE} + V_t \delta_{NPN} \quad (4.6)$$

$$I'_{PTAT} = \frac{\Delta V_{BE}}{R_{PTAT}} = \frac{\Delta V_{BE} + V_t \delta_{NPN}}{R_{PTAT}} = I_{PTAT} + \frac{V_t \delta_{NPN}}{R_{PTAT}} \quad (4.7)$$

Effect of transistor mismatch on CTAT stage:

The mismatch between the two transistors changes PTAT current and this effect cause a mismatch also in base-emitter voltage of T3:

$$V'_{BE_{T3}} = V_{BE_{T3}} + (V'_{BE_{T3}} - V_{BE_{T3}}) = V_{BE_{T3}} + V_t \ln \left( \frac{I'_{PTAT}}{I_{PTAT}} \right) = V_{BE_{T3}} + V_t \ln \left( \frac{I_{PTAT} + \frac{V_t \delta_{NPN}}{R_{PTAT}}}{I_{PTAT}} \right) \simeq V_{BE_{T3}} + V_t \frac{\delta_{NPN}}{\ln Q} \quad (4.8)$$

Effect of transistor mismatch on Output stage:

The effect of current mirror mismatch on bandgap output voltage can be evaluated by the equation:

$$V'_{BG} = R_{LOAD} \left( I'_{PTAT} + \frac{V'_{BE_{T3}}}{R_{CTAT}} \right) = V_{BG} + V_t \delta_{NPN} R_{LOAD} \left( \frac{1}{R_{PTAT}} + \frac{1}{R_{CTAT} \ln Q} \right) \quad (4.9)$$

### 4.1.3 Early Voltage

Early voltage affects the reference by altering PTAT current. Consequently, the base-emitter voltage relationship is also affected as well as PTAT term of the reference.

Effect of Early Voltage on PTAT stage:

The effects on the PTAT current are observed by taking into account the Early Voltage on collector current relationship:

$$I_C = J_s A_e \left( 1 + \frac{V_{CE}}{V_A} \right) \exp \left( \frac{V_{BE}}{V_t} \right) \quad (4.10)$$

or on the base-emitter voltage relationship:

$$V_{BE} = V_t \ln \frac{I_C}{J_s A_e \left( 1 + \frac{V_{CE}}{V_A} \right)} \quad (4.11)$$

$V_A$  is the Early Voltage. Now it is possible to evaluate again the PTAT current, using also Early Voltage:

$$I'_{PTAT} = \frac{V_t}{R_{PTAT}} \ln \left[ \frac{Q \left( 1 + \frac{V_{CE_{T0}}}{V_A} \right)}{\left( 1 + \frac{V_{CE_{T1}}}{V_A} \right)} \right] = I_{PTAT} \left[ 1 + \frac{1}{\ln Q} \ln \left( \frac{1 + \frac{V_{CE_{T0}}}{V_A}}{1 + \frac{V_{CE_{T1}}}{V_A}} \right) \right] \simeq I_{PTAT} \left( 1 + \frac{V_{CE_{T0}} - V_{CE_{T1}}}{V_A \ln Q} \right) \quad (4.12)$$

$V_{CE_{T0}}$  and  $V_{CE_{T1}}$  are the collector-emitter voltages of the transistors T0 and T1. The final approximation is true only if  $\frac{V_{CE_{T0}}}{V_A}$  and  $\frac{V_{CE_{T1}}}{V_A}$  are much less than 1.

Effect of Early Voltage on CTAT stage:

The effect on base-emitter voltage can be described by the following equation:

$$V'_{BE_{T3}} = V_{BE_{T3}} + \left( V'_{BE_{T3}} - V_{BE_{T3}} \right) = V_{BE_{T3}} + V_t \ln \left( \frac{I'_{PTAT}}{I_{PTAT}} \right) \simeq V_{BE_{T3}} + V_t \ln \left( 1 + \frac{V_{CE_{T0}} - V_{CE_{T1}}}{V_A \ln Q} \right) \simeq V_{BE_{T3}} + V_t \left( \frac{V_{CE_{T0}} - V_{CE_{T1}}}{V_A \ln Q} \right) \quad (4.13)$$

Effect of Early Voltage on Output stage:

The effect of current mirror mismatch on bandgap output voltage can be evaluated by the equation:

$$V'_{BG} = R_{LOAD} \left( I'_{PTAT} + \frac{V'_{BE_{T3}}}{R_{CTAT}} \right) = V_{BG} + R_{LOAD} \left( I_{PTAT} + \frac{V_t}{R_{CTAT}} \right) \left( \frac{V_{CE_{T0}} - V_{CE_{T1}}}{V_A \ln Q} \right) \quad (4.14)$$

### 4.1.4 Resistor process dependency

Process may cause a resistor inaccuracy ( $\delta_R$ ) in order of 10%-20%. This inaccuracy affects directly all the current reference and all the current consumption of the circuit, since the current is derived from a voltage across a resistance.

Effect of Resistor mismatch on PTAT stage:

The effect resistor mismatch on PTAT current generation:

$$I'_{PTAT} = \frac{V_t}{R_{PTAT}(1 + \delta_R)} \ln Q = \frac{R_{PTAT}}{R_{PTAT}(1 + \delta_R)} \frac{V_t}{R_{PTAT}} \ln Q = \frac{1}{(1 + \delta_R)} I_{PTAT} = I_{PTAT} \left( 1 - \frac{\delta_R}{1 + \delta_R} \right) \quad (4.15)$$

Effect of Resistor mismatch on CTAT stage:

This mismatch affects also base-emitter voltage:

$$V'_{BE_{T3}} = V_{BE_{T3}} + (V'_{BE_{T3}} - V_{BE_{T3}}) = V_{BE_{T3}} + V_t \ln \left( \frac{I'_{PTAT}}{I_{PTAT}} \right) = V_{BE_{T3}} + V_t \ln \left( \frac{1}{(1 + \delta_R)} \right) \simeq V_{BE_{T3}} - V_t \delta_R \quad (4.16)$$

Effect of Resistor mismatch on Output stage:

The effect of current mirror mismatch on bandgap output voltage can be evaluated by the equation:

$$V'_{BG} = R_{LOAD} \left( I'_{PTAT} + \frac{V'_{BE_{T3}}}{R_{CTAT}} \right) = R_{LOAD} (1 + \delta_R) \left[ I_{PTAT} \left( 1 - \frac{\delta_R}{1 + \delta_R} \right) + \frac{V_{BE_{T3}} - \delta_R V_t}{R_{CTAT} (1 + \delta_R)} \right] = V_{BG} - \frac{R_{LOAD} V_t}{R_{CTAT}} \delta_R$$

#### 4.1.5 Resistors temperature coefficient

The temperature coefficient (TC) of the resistors also affects the temperature-drift performance of the reference. Resistors  $R_{PTAT}$ ,  $R_{CTAT}$  and  $R_{LOAD}$  should be made of the same material, so they will have the same TC.

PTAT term and CTAT term of the reference relation are unaffected:

$$V_{BG} = \frac{R_{LOAD}}{R_{PTAT}} V_{PTAT} + \frac{R_{LOAD}}{R_{CTAT}} V_{CTAT} \quad (4.17)$$

The parasitic effects due to the TC of resistor  $R_{PTAT}$ , however, affect the reference. The behaviour of a realistic resistor can be described by the following equation:

$$R(T) = R(T_R) \left[ 1 + A(T - T_R) + B(T - T_R)^2 \right] \quad (4.18)$$

where A and B are the linear and quadratic temperature coefficients,  $R(T)$  is resistance's value at a temperature T,  $R(T_R)$  is the resistance at room temperature.

#### 4.1.6 Package-shift effects

The process of packaging a reference circuit induces variations on the output voltage. The mechanical stresses superimposed may alter the characteristics of the p-n junctions, which are the fundamental brick of all the reference circuit. Ceramic packages, instead of plastic packages, don't cause significant stresses.

#### 4.1.7 Threshold Voltage Mismatch

Referring to fig. 3.26, a mismatch in threshold voltage  $\Delta V_{TH}$  and current factor  $\Delta \beta$  ( $\beta = C_{ox} \mu \frac{W}{L}$ ) differences are the dominant sources of mismatching between MOS transistors. These random differences have a normal distribution with zero mean and their deviation depends on device area.

The threshold voltage of a mosfet may be expressed as:

$$V_{TH} = \Phi_{MS} + 2\Phi_B + \frac{1}{C} (Q_B - Q_F + qD_I) \quad (4.19)$$

where  $\Phi_{MS}$  is the gate-semiconductor work function difference,  $\Phi_B$  is Fermi potential in the bulk,  $C$  is the gate oxide capacitance per unit area,  $Q_B$  is the depletion charge density,  $Q_F$  is the fixed oxide charge density and  $D_I$  is the threshold adjust implant ions. The standard deviation of  $V_{TH}$  may be determined if it is possible to evaluate the

standard deviations of the various terms of 4.19: these terms are shaped as independent gaussian random variables with zero means and their standard deviation depends on the device area:

$$\frac{\sigma_{Q_B}^2}{Q_B^2} = \frac{1}{4LWW_dN_A} \quad \frac{\sigma_{Q_f}^2}{Q_f} = \frac{q}{LW} \quad \frac{\sigma_{D_I}^2}{D_I^2} = \frac{1}{LW} \quad \frac{\sigma_C^2}{C^2} = \frac{A_{ox}}{LW} \quad (4.20)$$

where  $L$  and  $W$  are the dimentions of the channel of mosfet,  $W_d$  is the depletion layer width,  $N_A$  is the substrate doping and  $A_{ox}$  is the variance in oxide thickness and permittivity.

The variance of  $V_{TH}$  may be written [Kadaba] as follows:

$$\sigma_{V_{TH}}^2 = \frac{1}{C^2} \left( \sigma_{Q_B}^2 + \sigma_{Q_f}^2 + q^2 \sigma_{D_I}^2 \right) + \frac{\sigma_C^2}{C^2} \left( \frac{Q_B^2}{C^2} + \frac{Q_f^2}{C^2} + \frac{qD_I^2}{C^2} \right) \quad (4.21)$$

#### 4.1.8 Current Factor Mismatch

The conductance constant is given by:

$$K = \mu C \frac{W}{L} \quad (4.22)$$

where  $\mu$  is the channel mobility. It is possible to express the variance of  $K$  in terms of the variances of  $\mu$ ,  $C$ ,  $W$  and  $L$ , which are shaped as independent gaussian random variables with zero means and their standard deviation depends on the device area. It is necessary to know the factor that affect the mobility, to determine its variance. At room temperature and a moderate gate bias the electron mobility is mainly governed by scattering due to interface charge centers and phonons.  $\mu$  can be expressed by an empirical relationship:

$$\mu = \frac{\mu_0}{1 + \alpha N_f} \quad (4.23)$$

where  $\alpha$  and  $\mu_0$  are two empirical constants, while  $N_f$  is the total dopant concentration. The variance of  $\mu$  can be expressed as:

$$\frac{\sigma_\mu^2}{\mu^2} = \left( \frac{A_\mu}{LW} \right)^2 = \frac{\alpha^2}{(1 + \alpha Q_f)^4} \frac{N_f}{LW} \quad (4.24)$$

The variance of  $K$  can be expressed as follows:

#### 4.1.9 Solutions implemented against mismatch effects

In previous chapter two simple solutions have been shown to contrast the effects of mismatch:

- Cascode current mirrors: these structure let increase the output resistance of the current mirror. In this way it is possible to mitigate Early Effects, i.e. the dependance of the collector current  $I_C$  by voltage drop between drain and source  $V_{DS}$ .
- Degeneration of mosfet emitters: using an emitter resitance (emitter degeneration) it is possible to improve matching of transistors.

#### 4.1.10 Monte Carlo Analysis

The effects of process and mismatch can be evaluated running a Monte Carlo Analysis in Cadence. This kind of analysis refers to “statistic blocks”, where statistical distributions and correlations of netlist parameters are specified. For each iteration of Monte Carlo Analysis (200 runs), new pseudo-random values are generated for the specified netlist parameters and the list of analysis are then executed. This analysis becomes therefore a tool that allows to examine and predict the effects of statistical variations of parameters. The statistical blocks allow to specify batch-to-batch (process) and per-instance (mismatch) variations for netlist parameters. These statistically-varying netlist parameters can be referenced by models or instances in the main netlist and may represent IC manufacturing process variation, or component variations for board-level designs for example.

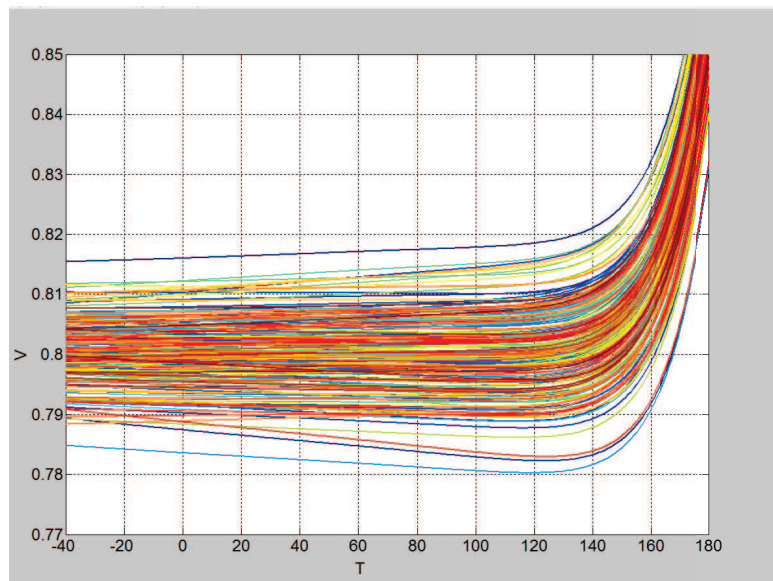


Figure 4.1: Monte Carlo Analysis of bandgap reference output voltage

Fig. 4.1 represents the “beam” of output voltage after Monte Carlo Analysis. It is possible to evaluate the output voltage at room temperature  $27^{\circ}\text{C}$ , i.e. the distance between nominal reference ( $800\text{mV}$ ) and the values at room temperature of all the runs.

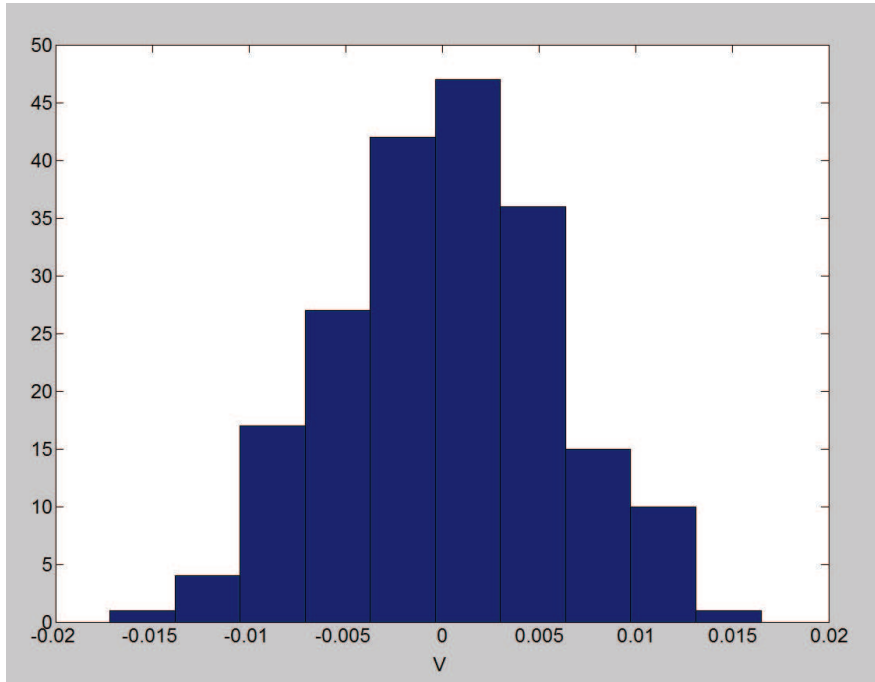


Figure 4.2: Amplitude spread's distribution of bandgap output voltage  $V_{BG}(T = 27^\circ\text{C}) - 800\text{mV}$   $\mu_D = 2.8383 \times 10^{-4}$   $\sigma_D = 5.731 \times 10^{-3}$

Process and mismatch cause, compared to nominal run, an inaccuracy of  $\pm 17\text{mV}$ . So the expected bandgap output voltage is  $(800 \pm 17)\text{mV}$ .

The bandgap output voltage is contaminated by leakage for  $T > 100^\circ\text{C}$ . In fact bandgap output voltage for  $T > 100^\circ\text{C}$  rises-up very quickly and this behaviour is due to leakage currents.

In order to avoid this contamination and to evaluate the “real” slope, it is necessary to consider bandgap output voltage only for  $T \leq 100^\circ\text{C}$ . In this way the slope  $\epsilon_s$  is evaluated as the difference  $\epsilon_s = V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$ . In other words the slope of bandgap output voltage (ideally equals to zero) is evaluated between  $T_{min} = -40^\circ\text{C}$  and  $T_{max} = 100^\circ\text{C}$ .

The difference  $V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$  can be calculated for each run of Monte Carlo Analysis in order to find a distribution of bandgap output voltage slope:



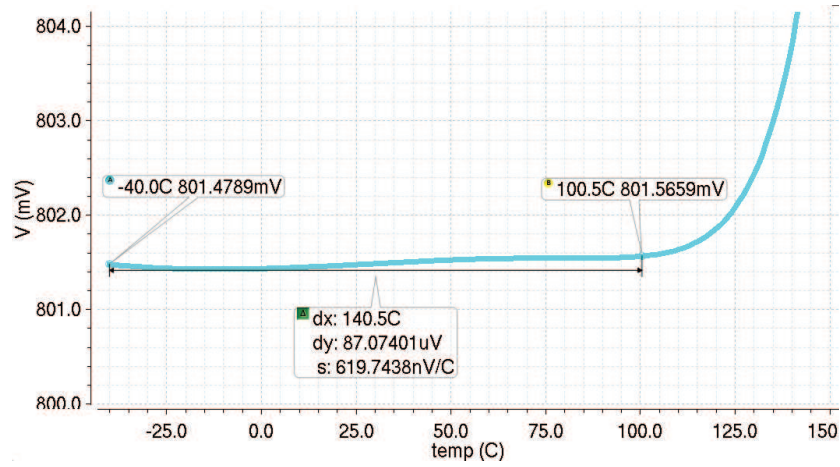


Figure 4.3: Solpe of bandgap output voltage (nominal run)

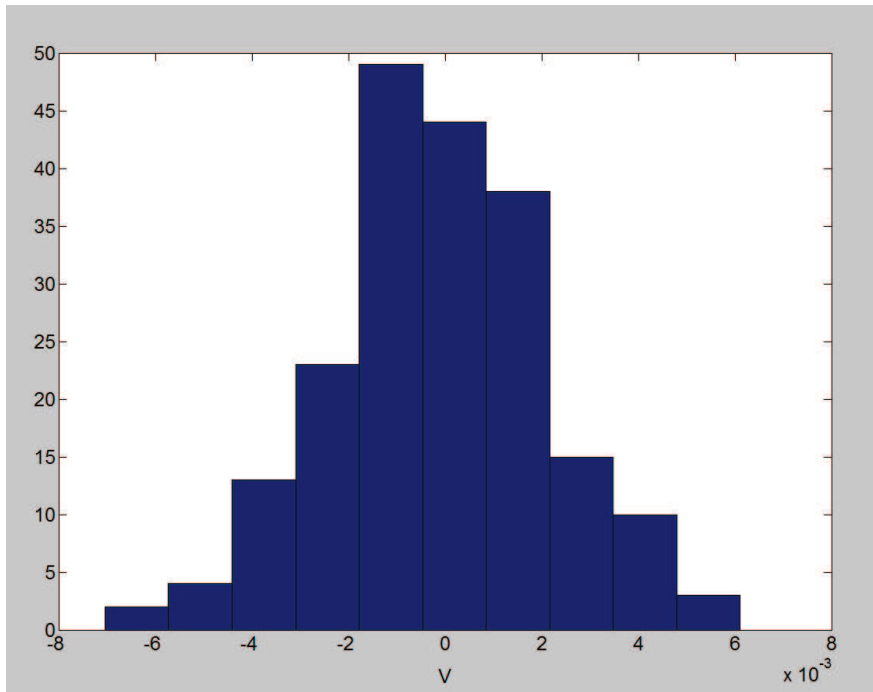


Figure 4.4: Slope distribution of bandgap output voltage  $V_{BG}(T = 100^{\circ}\text{C}) - V_{BG}(T = -40^{\circ}\text{C})$   $\bar{\mu}_S = 1.311 \times 10^{-4}\text{V}$   $\sigma_S = 2.192 \times 10^{-3}$

Process and mismatch, cause an inaccuracy also in the slope of bandgap output voltage: the worst runs have a slope of  $\epsilon_{s_1} = V_{BG_1}(T = 100^{\circ}\text{C}) - V_{BG_1}(T = -40^{\circ}\text{C}) = -7\text{mV}$  and  $\epsilon_{s_2} = V_{BG_2}(T = 100^{\circ}\text{C}) - V_{BG_2}(T = -40^{\circ}\text{C}) = 6\text{mV}$

### 4.1.11 Corner Analysis

The effects of process can be evaluated running a Corner Analysis in Cadence. Usually process variation is treated as a percentage variation in the performance calculation. Process parameters can be impurity concentration densities, oxide thickness and diffusion depths. These are caused by non uniform conditions during depositions and/or during diffusions of the impurities. This introduces variations in the resistances and in the transistor parameters such as threshold voltage. Process corners are used to investigate the parameter variations. The limited resolution of the photolithographic process may cause variations in the dimensions of the devices ( $W/L$ ). Process variations are due to variations in the manufacture conditions such as temperature, pressure and dopant concentration: there are variations in the process parameter over the whole chip. This causes, for example, the propagation delay to be different from production lot to production lot, because a smaller transistor is faster and therefore the propagation delay is smaller. Process parameters such as threshold voltage, mobility, etc. are different functions of temperature at different process corners.

Three fundamental corners exist: typical, fast and slow. Fast and slow corners vary different parameters, as for example carrier mobilities that are higher and lower than normal respectively. Cadence can mix the three corners for each batch of device in order to get corners combinations.

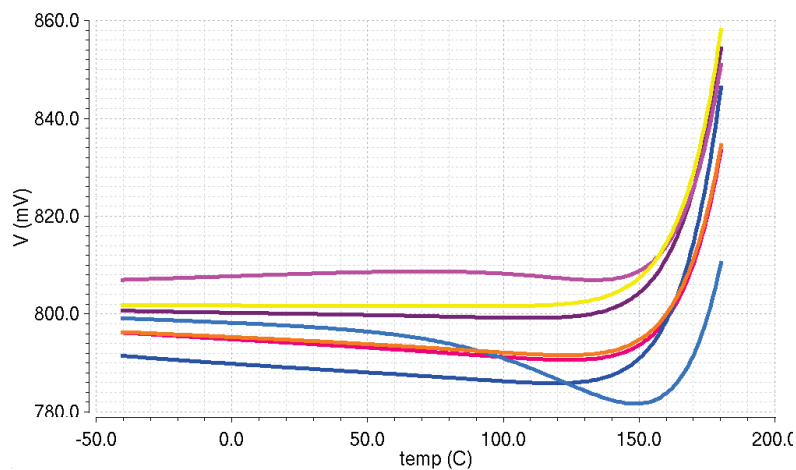


Figure 4.5: Corner Analysis of bandgap output voltage

It is possible to evaluate the output voltage at room temperature  $27^{\circ}\text{C}$ , i.e. the distance between nominal reference ( $800\text{mV}$ ) and the value at room temperatures of all the runs:

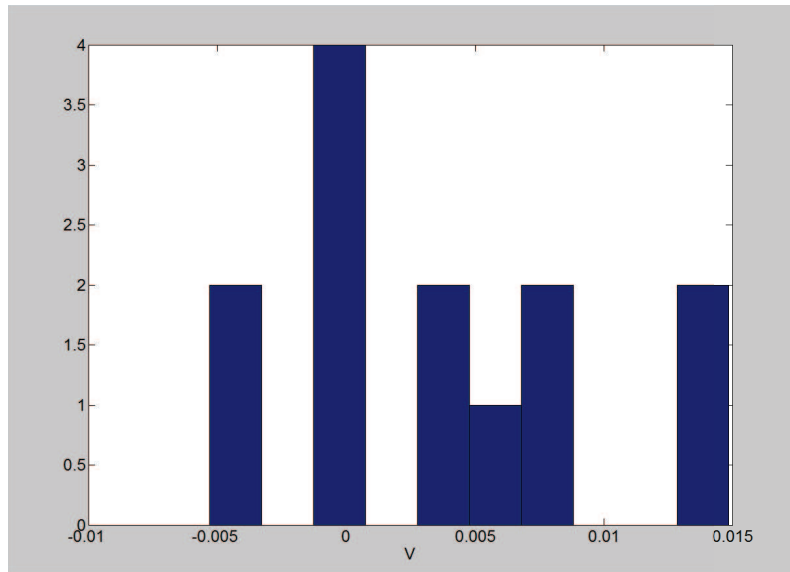


Figure 4.6: Amplitude spread distribution of bandgap output voltage  $V_{BG}(T = 27^\circ\text{C}) - 800\text{mV}$

Bandgap output voltage's nominal value at  $T = 27^\circ\text{C}$  is  $V_{BG} = 800\text{mV}$ , however process may change this value. Corner Analysis shows that the worst case is  $V_{BG} = (800 + 15)\text{mV}$ .

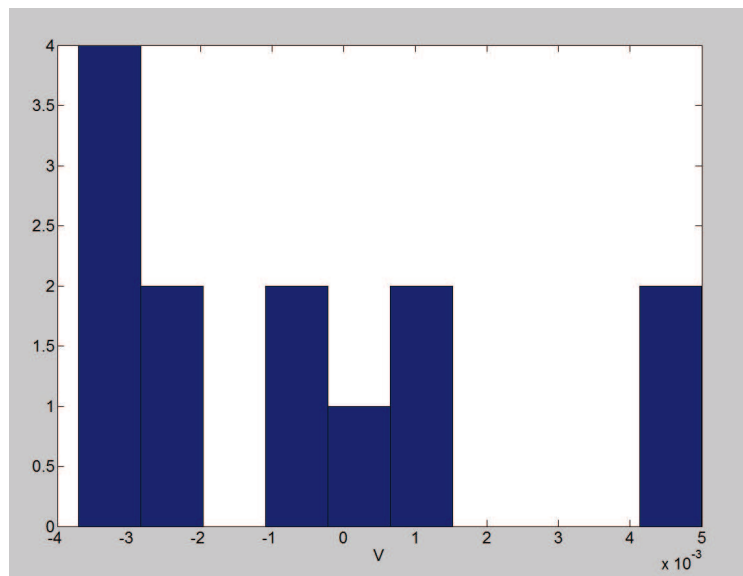


Figure 4.7: Slopes distribution of bandgap output voltage  $V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$

Nominal bandgap output voltage is flat: its slope between  $T_{min} = -40^\circ\text{C}$  and  $T_{max} = 100^\circ\text{C}$  is  $V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C}) = 695.298\mu\text{V}$ . However Corner Analysis shows that process may curve bandgap output voltage: the worst case is  $V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C}) = 5\text{mV}$ .

Corner Analysis may predict the negative effects of process, however Corner Analysis has not a statistical distribution, differently from Monte Carlo Analysis. A Corner Analysis can be run in order to evaluate the worst

case of process, while Monte Carlo Analysis can be run in order to evaluate the statistical distribution of parameters due to process and mismatch.

## 4.2 Trimming Techniques

As it has been shown previously, all precision references are subject to random parasitic effects and therefore require a post-fabrication adjustments (trim). Trimming is similar to fine-tuning, or adjusting, the value of a component. Trimming is typically use to achieve improvement of offset voltage of an op-amp, of the absolute value of a reference voltage or a specific delay time in a cascade of logic gates. Trimming techniques aim to make variable or programmable a component of the circuit: typical trimming networks are comprised of resistors that can be either short-circuited or open-circuited at room temperature. Trimming parameters are regulated by trimming words, i.e. binary words. Trimming throughout the temperature range is possible, but the required time is usually unacceptable for most commercial products. The associated cost and the space in silicon limit the implementation of trimming. There are two fundamental type of trimming: static trimming and dynamic or functional trimming. Static trimming, is a means of adjusting the resistor value without power being applied, while dynamic or functional trimming consists of adjusting a resistor to a specified value while the circuit is under power.

Over the years, trimming techniques have been developed to improve the accuracy of integrated circuits. There are three basic types of trimming techniques: Zener zap, fusible links and laser trim.

### 4.2.1 Zener zap

This trimming techniques provides the possibility of short-circuiting two nodes, like two resistor terminal. Zener diode sets and resistor sets are interconnected in a configuration that allow to cut elements out of a circuit. A current is forced to flow into the cathode of a small Zener diode to short-circuit both terminals of the diode (breakdown of p-n junction). Typically this current requires a large voltage drop across the cathode and anode terminals of the diode. The large reverse current through the diode dissipates enough localized power to permanently destroy the p-n junction. The metal is molten by current flowing resulting in a short-circuit between the anode and the cathode of the device. Zener zapping is highly reliable and stable over time. The number steps are limited because only incremental and finite resistor elements are either short-circuited or open-circuited.

### 4.2.2 Fusible links

Link fuse trimming is the process of selecting a desired resistance from a series of increasing resistors shorted by thin jumper wires. Fuses, unlike Zener-zapping diodes, are normally short-circuited devices capable of being open-circuited once trimmed. The fuses are typically fabricated with aluminium or polysilicon. Connected to each end of a fuse are two probe pads. Through these probe pads, a current, in the order of mA, is applied to selected fuses and so it opens the fuse.

Probe pads are extremely large when compared to the size of the resistor on an IC. Fuses are physically destroyed when a significant amount of current is forced to flow through them, thereby creating an open circuit. Fusible links are therefore less intrusive than Zener-zapping diodes, which usually require higher voltage to program. Precautionary steps are required to avoid metal regrowth resulting from “on-chip electromigration”.

### 4.2.3 Laser-trimmable resistors

Laser trim can also be used to cut metal links already short-circuiting resistor segments. There are two widely used laser systems:

- neodymium-doped yttrium aluminum garnet (YAG) crystal laser;
- carbon dioxide (CO<sub>2</sub>) laser;

In laser trimming a beam is focused, through a series of mirrors and lenses, on the plane of the resistor. The laser beam then hits the resistor material, energy is absorbed, and the material heats up and vaporizes.

As a result, the resistance of a single resistor is effectively modified. However there aren't digital bits defining the range and accuracy of the trim. The equivalent resistance is adjusted to a wide range of values since the laser redefines the physical shape of the device. The tradeoff of laser trimming is time and money. The process requires expensive equipment and considerable time to align and to trim each resistor. Laser trimming is the most effective and the most power-intrusive method of trimming a resistor. However Zener-zapping diodes and fusible links are the most economical.

### 4.3 Trimming options

In order to improve the precision of bandgap voltage reference and to limit the effects of process and mismatch, it is possible to trim some parameters, at room temperature  $T_{trim} = 27^\circ\text{C}$ , using zener-zap. The goals are to reduce the inaccuracy of the slope and the "spread" of the bandgap output voltage. Referring to fig. 4.8, there are four trimming options.

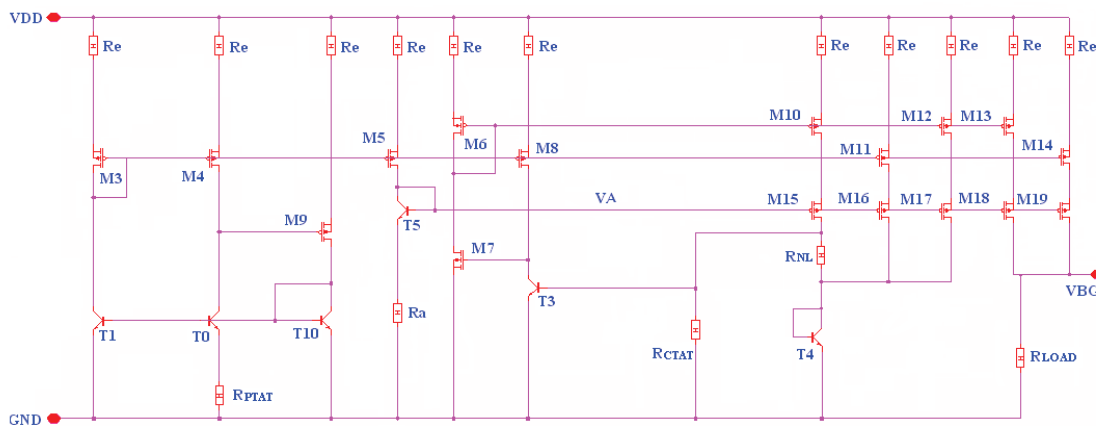


Figure 4.8: Complete Diode Loop

#### 4.3.1 Description of the Zener zap mechanism

In fig. 4.9 a cross section of a device used to implement the Zener zap is drawn. The cathode is heavily doped N-material. The anode of the Zener device is created by a moderately doped P-diffusion.

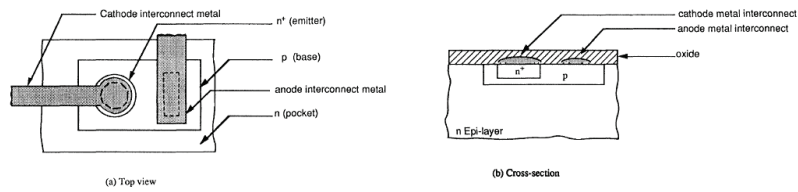


Figure 4.9: Bipolar Zener Device

The equation governing the breakdown voltage of the junction is given by:

$$BV = \frac{\epsilon (N_A + N_D) E_{crit}^2}{2qN_A N_D} \quad (4.25)$$

where  $\epsilon$  is the permittivity of silicon,  $N_A$  and  $N_D$  are the acceptor and donor doping densities of P-materials and N-materials respectively,  $q$  is the charge of an electron and  $E_{crit}$  is the maximum field that can be impressed across the depletion region of the P-N junction without leading avalanche breakdown [Analysis and Design Integrated circuits].

If the diode is  $N_D$  doped, eq. 4.25 may be simplified to:

$$BV \simeq \frac{\epsilon E_{crit}^2}{2qN_A} \quad (4.26)$$

Eq. 4.26 indicates that the P-material doping will determine the junction breakdown and that a higher doping will result in a lower breakdown voltage. The power dissipated by the junction at breakdown is given by:

$$P = BV \cdot I \quad (4.27)$$

where  $I$  is the current conducted across the junction during breakdown. If there aren't limits imposed upon  $I$ , the junction will be very rapidly destroyed. However, if  $I$  is limited, the power  $P$  will cause localized heating around the area where the current is concentrated. If  $I$  is applied for a fixed time, sufficient heating can occur to cause migration of atoms of the metal interconnect from the cathode terminal to the anode terminal along the path of breakdown current. This migration manifest itself as a trace of metal (aluminium) embedded in the silicon along the path of  $I$  (near the surface).

The mechanism by which the transport of metal atoms takes place with conducting electrons is known as electromigration. It occurs in metal lines at high current densities and elevated temperatures and consists of the movement of metal atoms toward the positive terminal of the conductor. The amount of current required to cause fusing will vary with the process and the device size. However it is possible to design structures for a given process that will fuse at low current.

The fusing phenomenon could be divided into two steps:

- the first initial breakdown and heating phase in which the metal interconnected atoms are mobilized and begin to flow across the junction;
- the second phase involves the carrying of a sufficient number of metal atoms to create a low resistance path through the silicon.

The resistance of a fused Zener diode is in the order of  $10\Omega$ .

### 4.3.2 First trimming option

Resistance of PTAT stage ( $R_{PTAT}$ ) could be trimmed in order to improve slope's accuracy of bandgap output voltage. If this resistance changes, the bias of CTAT stage and the bias of M11 changes. This is not a good solution, in fact a variation in PTAT current will impact on the bias CTAT stage. Moreover, compensation temperature  $T_R = 65^\circ\text{C}$  of the circuit depends on the current of M11-M16 ( $I_{PTAT}^*$ ) and on the current of M13-M18 ( $I_{CTAT}^*$ ). In Chapter 4 a tuning process has been used "to centre" the correct compensation temperature: as a result, a particular current ratio between  $I_{PTAT}^*/I_{PTAT}$  and  $I_{CTAT}^*/I_{CTAT}$  has been found.

Process and mismatch will change the parameters of the circuit and their variation will change also the compensation temperature.

At  $T = T_R$  voltage drop across  $R_{NL}$  has to be zero; so it is possible to evaluate the statistical distribution of  $T_R$ , using a Monte Carlo analysis, by observing the temperature of zero-voltage across  $R_{NL}$ .

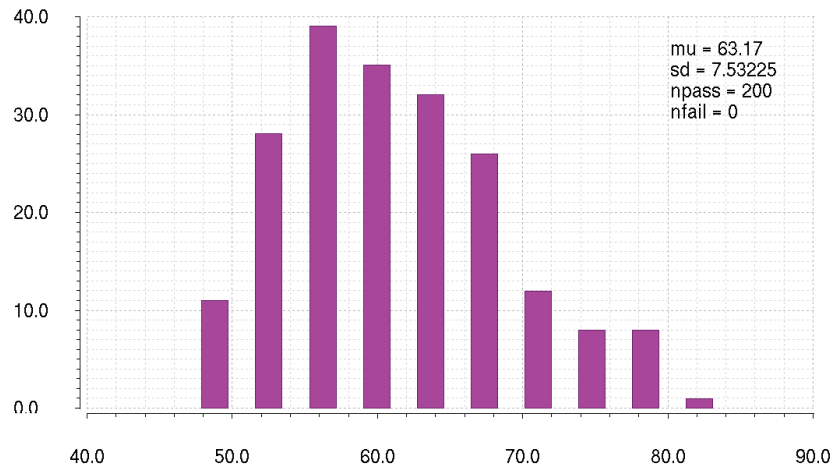


Figure 4.10: Statistical distribution of  $T = T_R$

However, it is hard to predict how a variation of  $I_{PTAT}$  could influence  $T = T_R$ . The equations we have found in Chapter 4 have a recursive nature and it is not possible to find an analytical relation that could describe a variation of  $T_R$  as function of  $I_{PTAT}$ .

### 4.3.3 Second trimming option

It is possible to trim  $I_{PTAT}^*$  or  $I_{CTAT}^*$  in order to reduce the slope inaccuracy of bandgap voltage reference. Trimming one of these currents, the precision of current gains  $I_{PTAT}^*/I_{PTAT}$  or  $I_{CTAT}^*/I_{CTAT}$  improves; in this way it is possible to avoid the effects of process and mismatch. Moreover this trimming option reduces also the inaccuracy of temperature compensation  $T_R$ , getting a very flat bandgap output voltage. The currents through M11-M16 ( $I_{PTAT}^*$ ) or the currents through M13-M18 ( $I_{CTAT}^*$ ) can be adjusted in different levels. The current flowing through mosfet can be regulated with a matrix of switches:

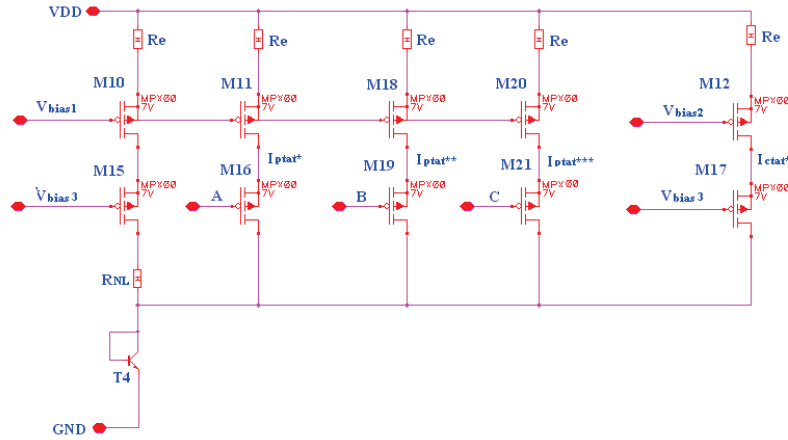


Figure 4.11: Switches matrix used to trim the current which generates  $I_{PTAT}^*$

This is not a good solution, because it is hard to predict how  $I_{PTAT}^*$  and  $I_{CTAT}^*$  will influence bandgap output voltage, because, due to the recursive nature of equations, it is impossible to find an analytical relation that could describe a variation of  $T_R$  as function of  $I_{PTAT}^*$  or  $I_{CTAT}^*$ . Moreover it is necessary to design a bias for the matrix of switches.

### 4.3.4 Third trimming option

In order to reduce the inaccuracy of the bandgap output voltages slope, trimming  $R_{NL}$  is a good solution. If the value of  $R_{NL}$  changes, the bandgap output voltage “rotates” around the compensation temperature.

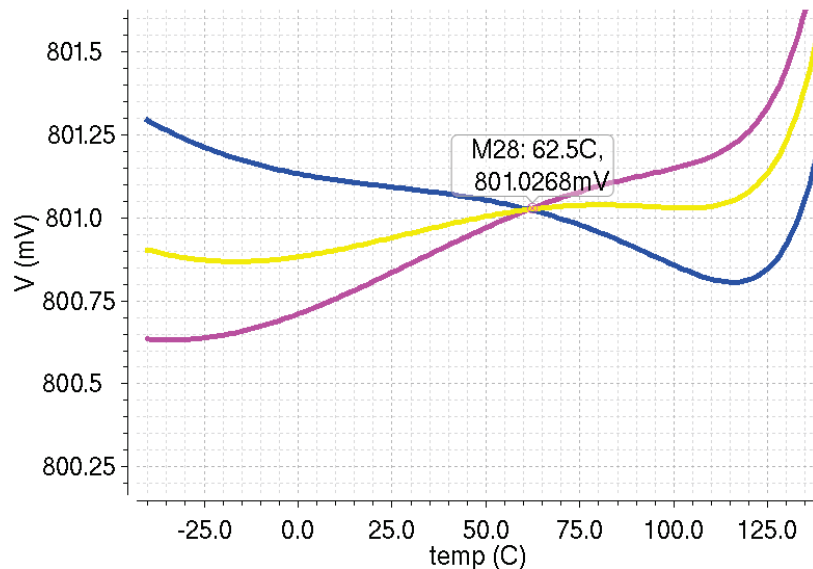


Figure 4.12: Changing  $R_{NL}$ , the bandgap output voltage will rotate around  $T = T_R$ . The yellow line is the nominal output voltage ( $R_{NL} = 134K\Omega$  and  $R_{LOAD} = 344K\Omega$ )



If  $R_{NL}^* < 134K\Omega$ , the bandgap output voltage has a PTAT behaviour (like purple line  $R_{NL} = 132k\Omega$  in fig. 4.12). If  $R_{NL}^* > 134K\Omega$ , bandgap output voltage has with a CTAT behaviour (like blue line  $R_{NL} = 137k\Omega$  in fig. 4.12).

A variation of  $R_{NL}$  will not modify the position of compensation temperature  $T = T_R$  and will not change the current gain  $I_{PTAT}^*/I_{PTAT}$  or the current gain  $I_{CTAT}^*/I_{CTAT}$ . It is possible to verify this by running a Monte Analysis.

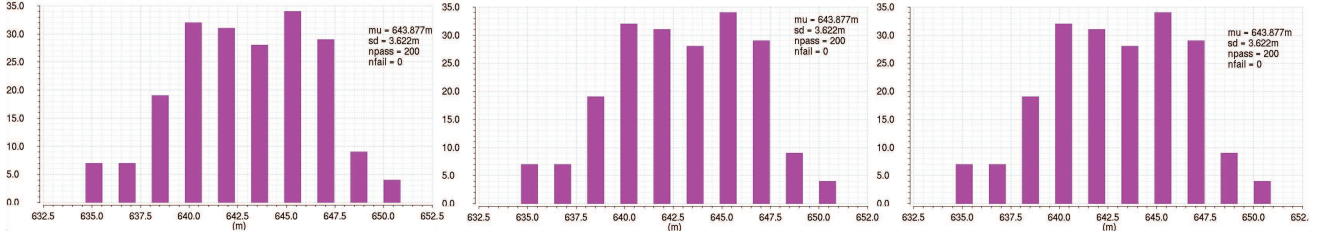


Figure 4.13: Distributions of  $I_{PTAT}^*/I_{PTAT}$  ratio at  $T = 27^\circ\text{C}$   $R_{NL1} = 130k\Omega$   $R_{NL2} = 134k\Omega$   $R_{NL3} = 138k\Omega$

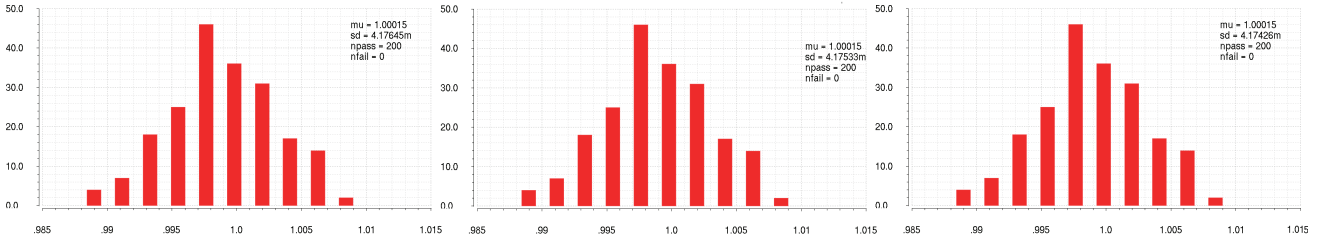


Figure 4.14: Distributions of  $I_{CTAT}^*/I_{CTAT}$  ratio at  $T = 27^\circ\text{C}$   $R_{NL1} = 130k\Omega$   $R_{NL2} = 134k\Omega$   $R_{NL3} = 138k\Omega$

It is possible to observe that the statistical distribution of  $I_{PTAT}^*$  and  $I_{CTAT}^*$  is always the same, although  $R_{NL}$  changes.

### 4.3.5 Fourth trimming option

Trimming  $R_{LOAD}$  is the only solution to reduce the “spread” of bandgap output voltage (i.e. the distance of the bandgap output voltage from its nominal value). Bandgap output voltage is obtained by the sum of three currents across an output resistance  $R_{LOAD}$ :

$$V_{BG} = R_{LOAD} (I_{PTAT} + I_{CTAT} + I_{NL}) \quad (4.28)$$

Eq. 4.28 is a linear equation, so by changing  $R_{LOAD}$  it is possible to “shift” up and down bandgap output voltage.

## 4.4 Trimming Resistances

Two resistors-trimming networks has been implemented in the bandgap voltage reference, in order to limit negative effects of process and mismatch. The resistors trim are:

- $R_{NL}$ : trimming this resistance, it is possible to change the slope of the bandgap output voltage in order to get a flatter output voltage;

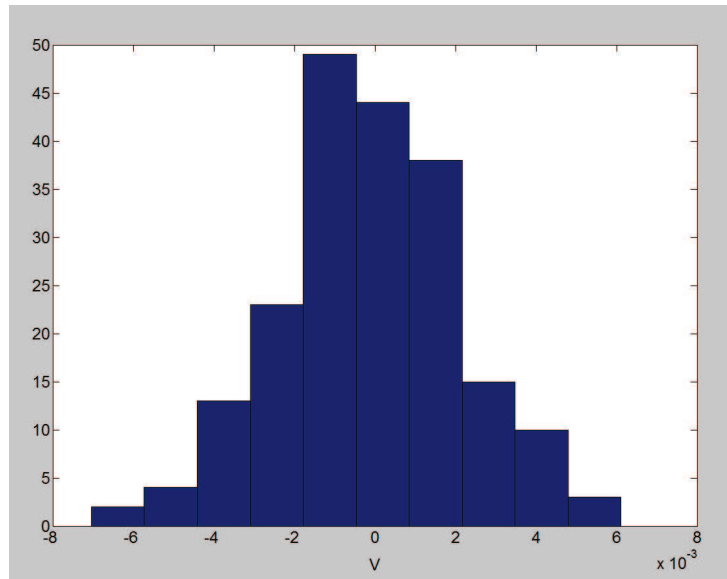


Figure 4.15: Initial spread of bandgap output voltage's slope  $V_{BG}(T = 100^{\circ}\text{C}) - V_{BG}(T = -40^{\circ}\text{C})$   $\bar{\mu}_S = 1.311 \times 10^{-4}\text{V}$   $\sigma_S = 2.192 \times 10^{-3}$

- $R_{LOAD}$ : trimming this resistance it is possible to shift up and down bandgap output voltage in order to reduce the distribution its initial spread.

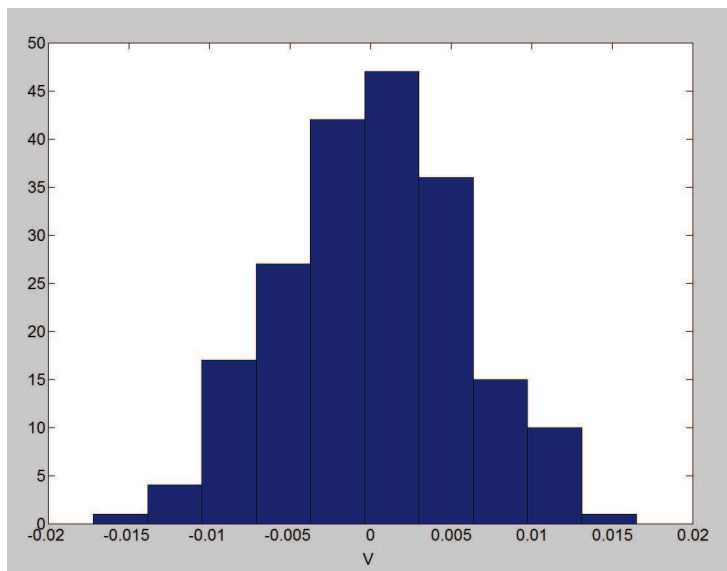


Figure 4.16: Output voltage's distribution  $V_{BG}(T = 27^{\circ}\text{C}) - 800\text{mV}$   $\bar{\mu}_D = 2.8383 \times 10^{-4}$   $\sigma_D = 5.731 \times 10^{-3}$

In order to trim a resistance, a wide range of trim is required; it is possible to cascade trim elements and resistors as shown in fig. 4.17

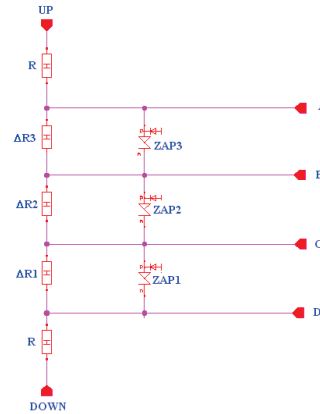


Figure 4.17: Cascaded trimming elements

For  $N$  trims in a resistor string,  $N+1$  bond pads will be required. If individual pads are provided as shown in fig. 4.17 to independently access any of the trim elements, then all possible combinations of trim could be achieved.

For  $N$  trimming elements, there are  $2^N$  trim combinations. It follows that if the trims are binary weighted ( $\Delta R_1$  is the lowest-order trim) all integer combinations of the lowest-order trim may be achieved up to  $(2^N - 1) \Delta R_1$ . The trim range  $R_T$  could be determined as the maximum resistance change when all the resistors are opened:

$$R_T = \Delta R_1 (2^N - 1) \quad (4.29)$$

The situation can be summarized as follows: a trim string, fig. 4.17, where the lowest order trimming step is  $\Delta R_1$ . Then  $\Delta R_2 = 2\Delta R_1$  and  $\Delta R_3 = 2\Delta R_2$ . The resolution of trimming is  $\Delta R_1$  and the trim range is  $(2^N - 1) \Delta R_1$ .

If  $\Delta V = f(\Delta R_1, \Delta R_2, \Delta R_3)$  is a linear function, i.e.  $\Delta V = I(\Delta R_1 + \Delta R_2 + \Delta R_3)$  the minimum voltage resolution step is:

$$step = \frac{\Delta V}{2^N - 1} \quad (4.30)$$

#### 4.4.1 Trimming $R_{NL}$

Minimum amount of silicon space used versus number of trimming bits is a tradeoff. More trimming bits require more area on silicon. Two trimming networks are designed for  $R_{NL}$ :

- first network:  $N = 3$  trimming bits (i.e.  $2^3 = 8$  trimming words) are used. referring to fig. 4.15  $\Delta V \simeq 13mV$ . Using this network, the step is  $\frac{\Delta V}{2^3 - 1} = \frac{13mV}{2^3 - 1} = 1.86mV$  and the goal is to centre the statistical distribution between  $\pm \frac{1.86mV}{2} = \pm 0.926mV$ .
- second network:  $N = 2$  trimming bits (i.e.  $2^2 = 4$  trimming words) are used. referring to fig. 4.15  $\Delta V \simeq 13mV$ . Using this network, the step is  $\frac{\Delta V}{2^2 - 1} = \frac{13mV}{2^2 - 1} = 4.33mV$  and the goal is to centre the statistical distribution between  $\pm \frac{4.33mV}{2} = \pm 2.17mV$ .

It is necessary to find an analytical relation to evaluate  $\Delta R_1$ . Now only two bounds are known:

- upper bound: the smallest resistance has to give a PTAT behaviour to bandgap output voltage:  $V_{BG}(T = 100^\circ C) - V_{BG}(T = -40^\circ C) \simeq 7mV$ .

- lower bound: the biggest resistance has to give a CTAT behaviour to bandgap output voltage:  $V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C}) \simeq -6\text{mV}$ .

Using the equations described in the previous Chapter, it is possible to observe that if  $R_{NL}$  changes, also  $I_{NL}$  changes. In fact, the equations have a recursive nature  $V_{BG} = R_{LOAD}(I_{PTAT} + I_{CTAT} + I_{NL})$ :

$$V_{BG} = \frac{R_{LOAD}}{R_{PTAT}} \left( \frac{k_B}{q} \ln Q \right) T + \frac{R_{LOAD}}{R_{CTAT}} \left[ V_{G0} - \frac{V_{G0} - V_{BE}(T_R)}{T_R} T - (\eta - 1) \frac{k_B T}{q} \ln \frac{T}{T_R} \right] + \frac{R_{LOAD}}{R_{NL}} V_i \left( \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right) \quad (4.31)$$

It is impossible to find analytically  $\Delta R_1$  using this equation, however the equation can be made easier:  $I_{NL}$  is ignored. The goal is to write the new value of  $R_{NL}^*$  as a function of the “desidered slope” of bandgap output voltage  $\Delta V$  and  $T$ :

$$R_{NL}^* = f(\Delta V, T) \quad (4.32)$$

Eq. 4.31 can be written for two different temperatures  $T_1 = -40^\circ\text{C}$  and  $T_2 = 100^\circ\text{C}$ , using for  $R_{NL}$  and  $R_{LOAD}$  the values calculated in previous Chapter:  $R_{NL} = 134\text{K}\Omega$  and  $R_{LOAD} = 344\text{K}\Omega$ .

$$\begin{cases} V_{BG_1} = \alpha_1 + \frac{V_{i1}}{R_{NL}} \ln \left( \frac{2I_{PTAT_1}}{I_{const_1}} \right) - \frac{\eta-1}{R_{CTAT}} \frac{k_B}{q} \ln \left( \frac{T_1}{T_R} \right) \\ V_{BG_2} = \alpha_2 + \frac{V_{i2}}{R_{NL}} \ln \left( \frac{2I_{PTAT_2}}{I_{const_2}} \right) - \frac{\eta-1}{R_{CTAT}} \frac{k_B}{q} \ln \left( \frac{T_2}{T_R} \right) \end{cases} \quad (4.33)$$

where  $\alpha_1 = \frac{R_{LOAD}}{R_{PTAT}} \left( \frac{k_B}{q} \ln Q \right) T_1 + \frac{R_{LOAD}}{R_{CTAT}} \left[ V_{G0} - \frac{V_{G0} - V_{BE}(T_R)}{T_R} T_1 \right]$        $\alpha_2 = \frac{R_{LOAD}}{R_{PTAT}} \left( \frac{k_B}{q} \ln Q \right) T_2 + \frac{R_{LOAD}}{R_{CTAT}} \left[ V_{G0} - \frac{V_{G0} - V_{BE}(T_R)}{T_R} T_2 \right]$

Now eq. 4.31 can be written for the same temperatures  $T_1 = -40^\circ\text{C}$  and  $T_2 = 100^\circ\text{C}$ , the same  $R_{LOAD} = 344\text{K}\Omega$ , but an unknown value:  $R_{NL}^*$ .

$$\begin{cases} V_{BG_1}^* = \alpha_1 + \frac{V_{i1}}{R_{NL}^*} \ln \left( \frac{2I_{PTAT_1}}{I_{const_1}} \right) - \frac{\eta-1}{R_{CTAT}} \frac{k_B}{q} \ln \left( \frac{T_1}{T_R} \right) \\ V_{BG_2}^* = \alpha_2 + \frac{V_{i2}}{R_{NL}^*} \ln \left( \frac{2I_{PTAT_2}}{I_{const_2}} \right) - \frac{\eta-1}{R_{CTAT}} \frac{k_B}{q} \ln \left( \frac{T_2}{T_R} \right) \end{cases} \quad (4.34)$$

Now the nominal slope  $\Delta V_{BG} = V_{BG_2} - V_{BG_1}$  can be use as “reference” to evaluate the unknown slope  $\Delta V_{BG}^* = V_{BG_2}^* - V_{BG_1}^*$ . In other words the following system of equation has to be solved:

$$\begin{cases} \Delta V_{BG} = V_{BG_2} - V_{BG_1} \\ \Delta V_{BG}^* = V_{BG_2}^* - V_{BG_1}^* \\ \Delta V = \Delta V_{BG} - \Delta V_{BG}^* \end{cases} \quad (4.35)$$

$\Delta V$  is the desidered slope. The system of equations 4.35 yields  $R_{NL}^* = \left\{ \frac{1}{R_{NL}} - \frac{\Delta V}{\left[ V_{i1} \ln \left( \frac{2I_{PTAT_1}}{I_{const_1}} \right) - V_{i2} \ln \left( \frac{2I_{PTAT_2}}{I_{const_2}} \right) \right] R_{LOAD}} \right\}^{-1}$

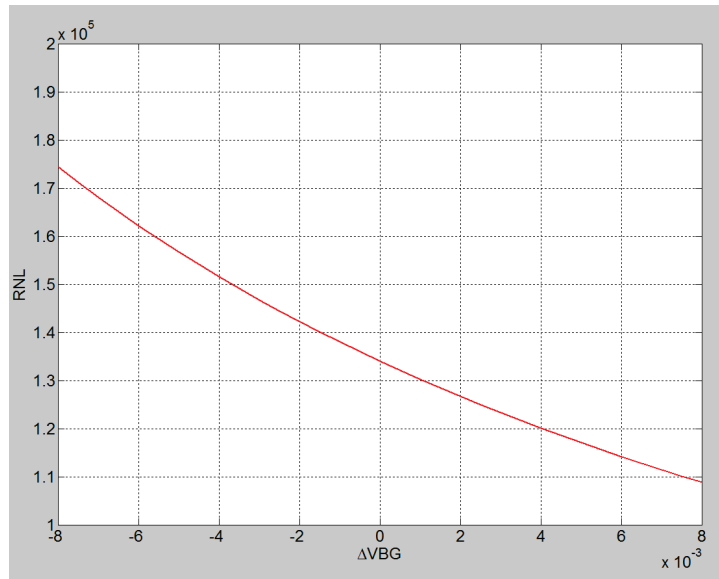


Figure 4.18:  $R_{NL}^* = f(\Delta V, T_1, T_2)$

$R_{NL}^*$  can't be written as a linear function of  $\Delta V$ ,  $T_1$  and  $T_2$ : this non-linearity will give some errors (voltage step is not constant).

It is possible to compare the analytical relation  $R_{NL}^* = f(\Delta V, T_1, T_2)$  with some experimental points achieved by a Cadence parametric analysis:

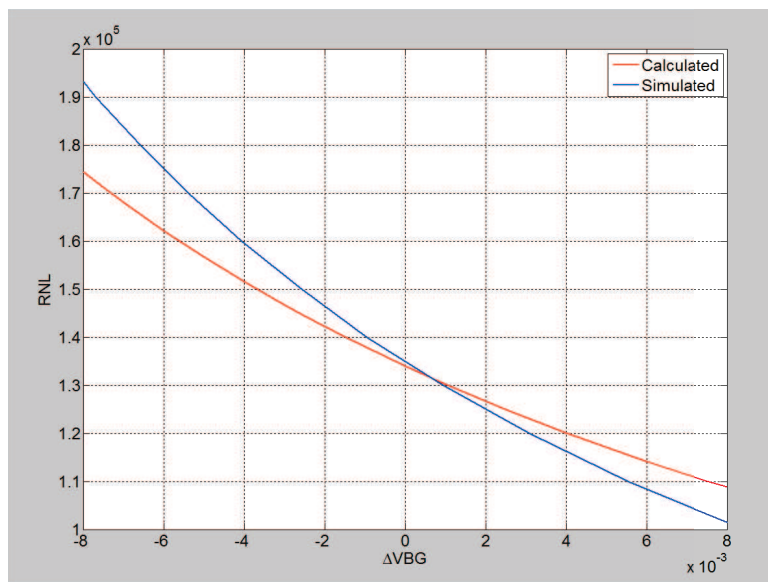


Figure 4.19: Comparison between  $R_{NL}^* = f(\Delta V, T_1, T_2)$  and Cadence values

Analytical model does not fit completely with Cadence values: this is due to the approximation done ignoring  $I_{NL}$  in 4.33 and in 4.34. So  $R_{NL_{min}} = 105K\Omega$  and  $R_{NL_{max}} = 180K\Omega$ .

$$\Delta R_{NL} = R_{NL_{max}} - R_{NL_{min}} \simeq 75K\Omega \tag{4.36}$$

$\Delta R_1$  could be calculated as follows for a trimming network with  $2^3 = 8$  trimming words and binary weighted trims:

$$\Delta R_{NL} = \Delta R_1 + 2\Delta R_1 + 4\Delta R_1 \tag{4.37}$$

Eq. 4.37 yields  $\Delta R_1 = 11K\Omega$ .

Instead if trimming network has  $2^2 = 4$  trimming words:

$$\Delta R_{NL} = \Delta R_1 + 2\Delta R_1 \tag{4.38}$$

Eq. 4.38 yields  $\Delta R_1 = 26K\Omega$ .

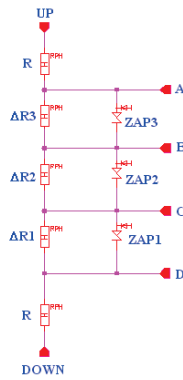


Figure 4.20: Trimming network with 8 trimming words  $R = 52.5K\Omega$ ,  $\Delta R_1 = 11K\Omega$ ,  $\Delta R_2 = 22K\Omega$ ,  $\Delta R_3 = 44K\Omega$

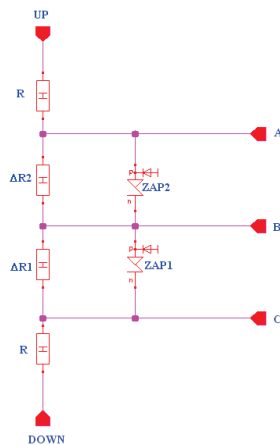


Figure 4.21: Trimming network with 4 trimming words  $R = 52.5K\Omega$ ,  $\Delta R_1 = 26K\Omega$ ,  $\Delta R_2 = 52K\Omega$

#### 4.4.2 Trimming $R_{LOAD}$

In order to trim  $R_{LOAD}$ , it is also possible to design two networks:

- first network:  $N = 3$  trimming bits (i.e.  $2^3 = 8$  trimming words) will be used. referring to fig. 4.15  $\Delta V \simeq 40mV$ . Using this network, the minimum voltage step is  $\frac{\Delta V}{2^N - 1} = \frac{40mV}{2^3 - 1} = 5.71mV$  and the goal is to centre the statistical distribution between  $\pm \frac{5.71mV}{2} = \pm 2.86mV$ .
- second network:  $N = 4$  trimming bits (i.e.  $2^4 = 16$  trimming words) will be used. referring to fig. 4.15  $\Delta V \simeq 40mV$ . Using this network, the minimum voltage step is  $\frac{\Delta V}{2^N - 1} = \frac{40mV}{2^4 - 1} = 2.67mV$  and the goal is to centre the statistical distribution between  $\pm \frac{2.67mV}{2} = \pm 1.33mV$ .

It is necessary to find an analytical relation to evaluate  $\Delta R_1$ . Now only two bounds are known::

- upper bound: the smallest resistance has to shift down bandgap output voltage:  $V_{BG}(T = 27^\circ C) - 800mV \simeq -20mV$ .
- lower bound: the biggest resistance has to shift up bandgap output voltage:  $V_{BG}(T = 27^\circ C) - 800mV \simeq +20mV$ .

The equations described in the previous Chapter can be used to get an analytical relation that describes new values of  $R_{LOAD}^*$  as a function of its known value  $R_{LOAD} = 344K\Omega$  (nominal value), the temperature  $T$  and the desired bandgap output voltage variation  $\Delta V_{BG}$ :

$$R_{LOAD}^* = f(R_{LOAD}, \Delta V_{BG}, T) \quad (4.39)$$

referring to eq. 4.31, it is possible to write the following system of equations:

$$\left\{ \begin{array}{l} V_{BG_1} = R_{LOAD} \left[ \frac{V_{G0}}{R_{CTAT}} + \left( \frac{k_B}{q} \frac{\ln Q}{R_{PTAT}} - \frac{V_{G0} - V_{BE}(T_R)}{R_{CTAT} T_R} \right) T - \frac{k_B}{q R_{PTAT}} T \ln \frac{T}{T_R} (\eta - 1) + \frac{1}{R_{NL}} \frac{k_B}{q} T \ln \left( \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right) \right] \\ V_{BG_2} = R_{LOAD}^* \left[ \frac{V_{G0}}{R_{CTAT}} + \left( \frac{k_B}{q} \frac{\ln Q}{R_{PTAT}} - \frac{V_{G0} - V_{BE}(T_R)}{R_{CTAT} T_R} \right) T - \frac{k_B}{q R_{PTAT}} T \ln \frac{T}{T_R} (\eta - 1) + \frac{1}{R_{NL}} \frac{k_B}{q} T \ln \left( \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right) \right] \end{array} \right. \quad (4.40)$$

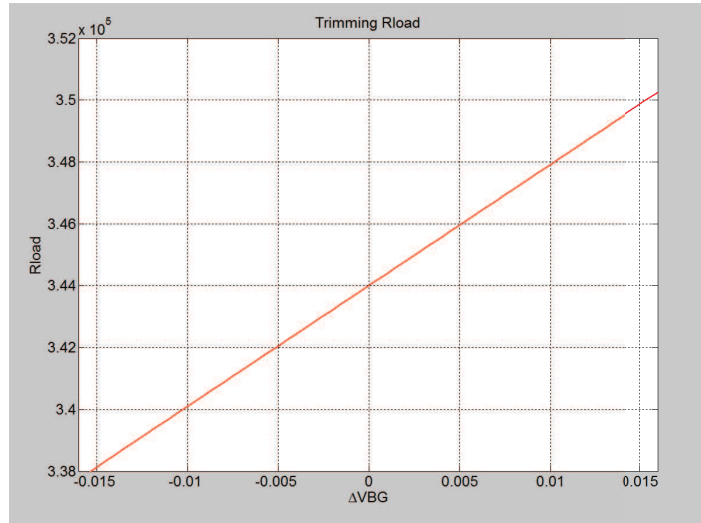
Eq. 4.40 can also be written as

$$\left\{ \begin{array}{l} V_{BG_1} = R_{LOAD} \alpha \\ V_{BG_2} = R_{LOAD}^* \alpha \end{array} \right. \quad (4.41)$$

where  $\alpha = \frac{V_{G0}}{R_{CTAT}} + \left( \frac{k_B}{q} \frac{\ln Q}{R_{PTAT}} - \frac{V_{G0} - V_{BE}(T_R)}{R_{CTAT} T_R} \right) T - \frac{k_B}{q R_{PTAT}} T \ln \frac{T}{T_R} (\eta - 1) + \frac{1}{R_{NL}} \frac{k_B}{q} T \ln \left( \frac{2I_{PTAT}}{I_{NL} + I_{const}} \right)$ ,  $T_R = 338^\circ K$ ,  $T = 300^\circ K$ ,  $V_{BE}(T_R) = 594.653mV$ .

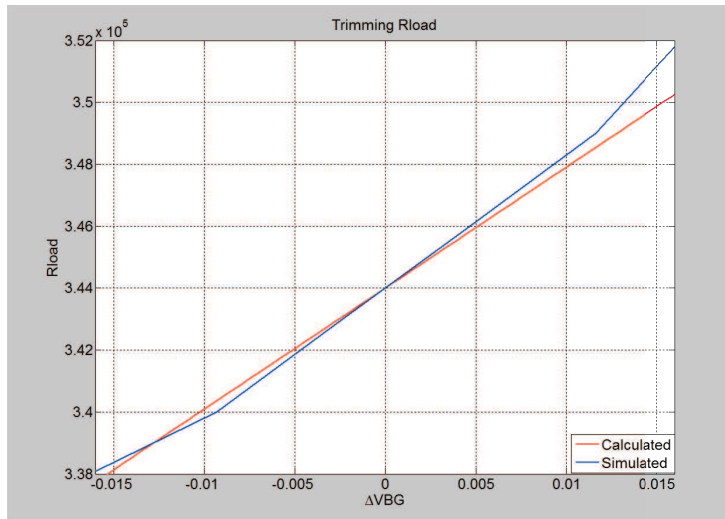
The solution of system 4.41 is:

$$R_{LOAD}^* = R_{LOAD} - \frac{V_{BG_1} - V_{BG_2}}{\alpha} = R_{LOAD} - \frac{\Delta V_{BG}}{\alpha} \quad (4.42)$$

Figure 4.22:  $R_{LOAD}^* = f(R_{LOAD}, \Delta V_{BG}, T)$ 

$R_{LOAD}^*$ , unlike  $R_{NL}^*$ , can be written as a linear function of  $\Delta V_{BG}$ ,  $T$  and  $R_{LOAD}$ : the minimum voltage trimming step of  $R_{LOAD}$  is constant.

It is possible to compare the analytical relation  $R_{LOAD}^* = f(R_{LOAD}, \Delta V_{BG}, T)$  with some experimental points achieved by a Cadence parametric analysis:

Figure 4.23: Comparison between  $R_{LOAD}^* = f(R_{LOAD}, \Delta V_{BG}, T)$  and Cadence's values

The model fits with Cadence values.

So  $R_{LOAD_{min}} = 336K\Omega$  and  $R_{LOAD_{max}} = 352K\Omega$ .

$$\Delta R_{LOAD} = R_{LOAD_{max}} - R_{LOAD_{min}} \simeq 16K\Omega \quad (4.43)$$



If the trims are binary weighted and the trimming network has  $2^3 = 8$  trimming words,  $\Delta R_1$  can be calculated as follows:

$$\Delta R_{LOAD} = \Delta R_1 + 2\Delta R_1 + 4\Delta R_1 \quad (4.44)$$

Eq. 4.44 yields  $\Delta R_1 \simeq 2.5K\Omega$ .

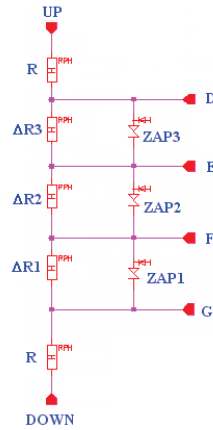


Figure 4.24: Trimming network with 8 trimming words  $R = 52.5K\Omega$ ,  $\Delta R_1 = 2.5K\Omega$ ,  $\Delta R_2 = 5K\Omega$ ,  $\Delta R_3 = 10K\Omega$

Instead if the network has  $2^4 = 16$  trimming words:

$$\Delta R_{NL} = \Delta R_1 + 2\Delta R_1 + 4\Delta R_1 + 8\Delta R_1 \quad (4.45)$$

Eq. 4.45 yields  $\Delta R_1 = 1.1K\Omega$ . With the used technology the smallest resistance is about  $R_{smallest} = 2.2K\Omega$  and so  $\Delta R_1$  can be realized with the parallel of two  $R_{smallest}$ .

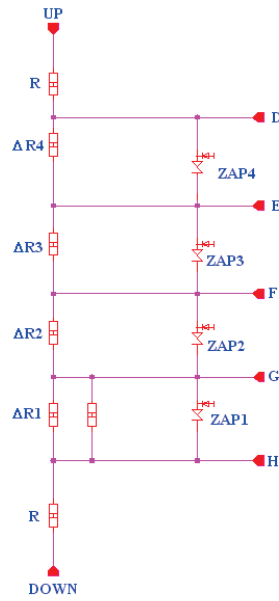


Figure 4.25: Trimming network with 16 trimming words  $R = 52.5K\Omega$ ,  $\Delta R_1 = 1.1K\Omega$ ,  $\Delta R_2 = 2.2K\Omega$ ,  $\Delta R_3 = 4.4K\Omega$ ,  $\Delta R_4 = 8.8K\Omega$



## Chapter 5

# TRIMMING VERIFICATION

### 5.1 Residual curvature of nominal run

The random effects of process and mismatch cause a slope inaccuracy and amplitude spread of bandgap output voltage. Trimming options let reduce the inaccuracy and the spread using networks with variable resistances. However, what may occur is that a significant nonlinear systematic errors may be present. The effects of these errors yields to the “residual curvature”: that is not possible to compensate. The nonlinear compensation is never perfect and so bandgap output voltage is never completely flat. Trimming networks may shift up and down or rotate around  $T_R$  bandgap output voltage, however they can't reduce this kind of error. If process and mismatch give to bandgap output voltage a PTAT or a CTAT behaviour, bandgap output voltage can be flattened up to its residual curvature adjusting  $R_{NL}$ .

It is possible to implement a “rectification algorithm” to investigate the residual curvature. The nominal bandgap output voltage is drawn in fig. 5.1

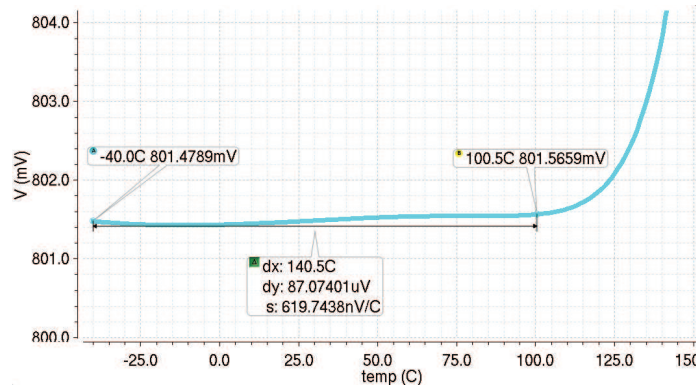


Figure 5.1: The nominal bandgap output voltage  $R_{PTAT} = 60k\Omega$ ,  $R_{CTAT} = 470k\Omega$ ,  $R_{NL} = 134k\Omega$ ,  $R_{LOAD} = 344k\Omega$

In order to implement correctly the rectification algorithm it is necessary to avoid leakage effects: bandgap output voltage has to be evaluated between  $T_{min} = -40^\circ C$  and  $T_{max} = 100^\circ C$ . Using a MATLAB script it is possible to find the coefficients of a polynomial  $P(X)$  of degree  $N = 1$  (i.e. a trendline) that fits the bandgap output voltage best in a least-squares sense. The trendline  $L$  to fit the bandgap output voltage is calculated with the coefficients of  $P(X)$  for  $T \in [T_{min}, T_{max}]$ . If bandgap output voltage is completely flat  $L$  is parallel to the temperature axis (i.e.  $L$  is parallel to the x axis).

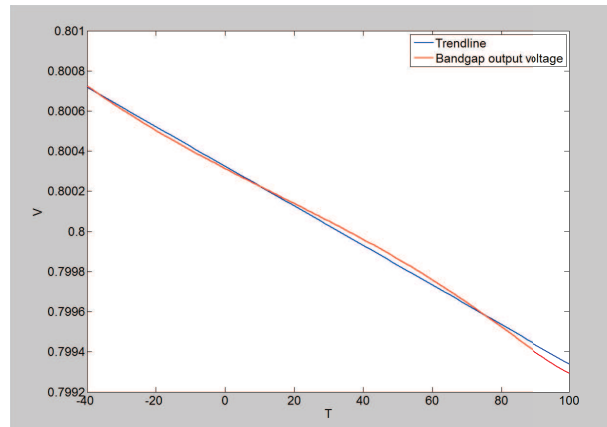


Figure 5.2: Fitting bandgap output voltage with  $L = mx + q$

The residual curvature can be calculated as the difference between bandgap output voltage and  $L$ . The residual curvature is caused by the residual errors like base currents, channel-length-modulation effect and the nonlinear-order compensation. This kind of error cannot be reduced by the designed trimming networks of  $R_{NL}$  and  $R_{LOAD}$ . The random effects of mismatch and process will increase this error, but these effects may be mitigated by trimming networks of  $R_{NL}$  and  $R_{LOAD}$ .

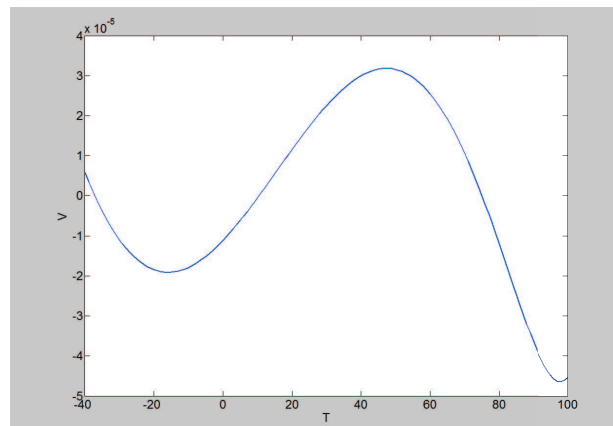


Figure 5.3: Residual curvature for nominal run of nonlinear-order compensated bandgap voltage reference

It is possible to compare the residual curvature of the nonlinear compensated voltage reference with the residual curvature of a first order compensated voltage reference, realized with the same technology:

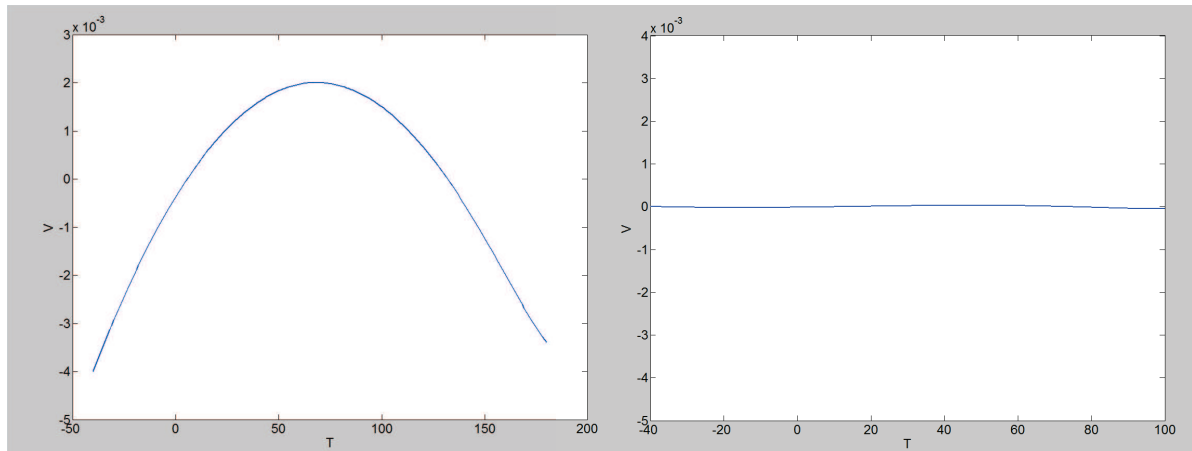


Figure 5.4: Comparison between residual curvatures of the nonlinear-order compensated voltage reference and a first-order compensated voltage reference

The residual curvature of first-order bandgap voltage reference can be evaluated between  $T_{min} = -40^{\circ}\text{C}$  and  $T_{max} = 160^{\circ}\text{C}$  because a leakage compensation circuit has been developed for this circuit.

First-order bandgap voltage references compensate the linear component, but does not compensate the non-linear component of the base-emitter voltage. Instead curvature-corrected bandgap references attempt to approximately cancel the nonlinear component of the diode voltage, in addition to first-order terms cancellation. In this way the residual curvature of a first-order is bigger than a nonlinear-order.

Nominally, the residual curvature for the first-order is about  $6 \times 10^{-3}\text{V}$ , while for a nonlinear-order is about  $7 \times 10^{-5}\text{V}$ .

## 5.2 Residual curvature of Monte Carlo Analysis

It is possible to use the rectification algorithm to investigate the residual curvature after a Monte Carlo Analysis. This is useful in order to have informations about the statistical distribution of the residual curvature. Monte Carlo Analysis predicts the statistical behavior of a circuit when values are varied within tolerance. Running a Monte Carlo Analysis, it is always necessary to specify the number of runs, i.e. the number of times the selected simulation profiles will be run. Component parameters with tolerances will be randomly varied for each run. The maximum number of runs is primarily limited by the amount of available memory.

Each Monte Carlo Analysis has 200 runs. In this way the Analysis guarantees that the results have a good statistical distribution. Using a Matlab script it is possible to fit each run with a trendline and subtract it from bandgap output voltage, in order to get the residual curvature. The rectification algorithm proposed is like an ideal trimming: bandgap output voltage can be flattened if it has a PTAT or a CTAT behaviour. The result is a “beam”: it represents how process and mismatch may change the accuracy of bandgap reference, after trimming.

Moreover, referring to fig. 5.5 and to fig. 5.6, difference between maximum and minimum in each run of the beam has been calculated in order to represent the residual curvature as a statistical distribution.

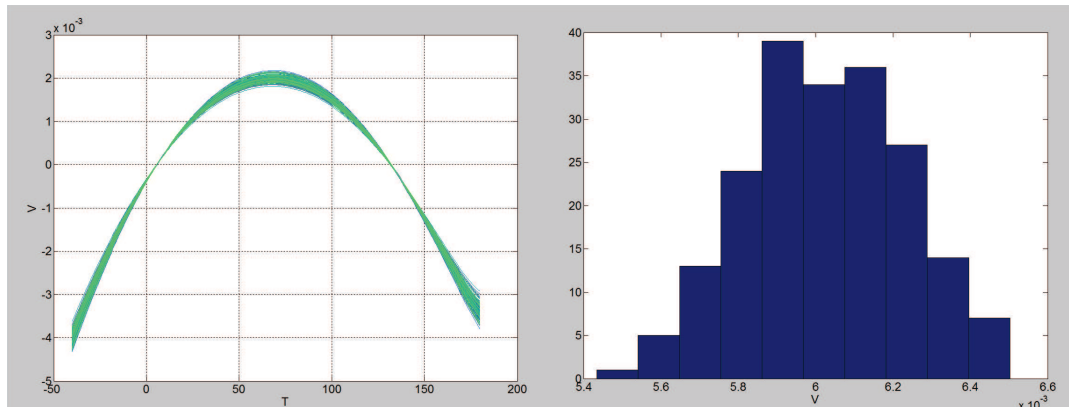


Figure 5.5: Evaluation of the residual curvature of a first-order using Monte Carlo Analysis  $\mu = 6.03 \times 10^{-3}V$   
 $var = 4.22 \times 10^{-8}$

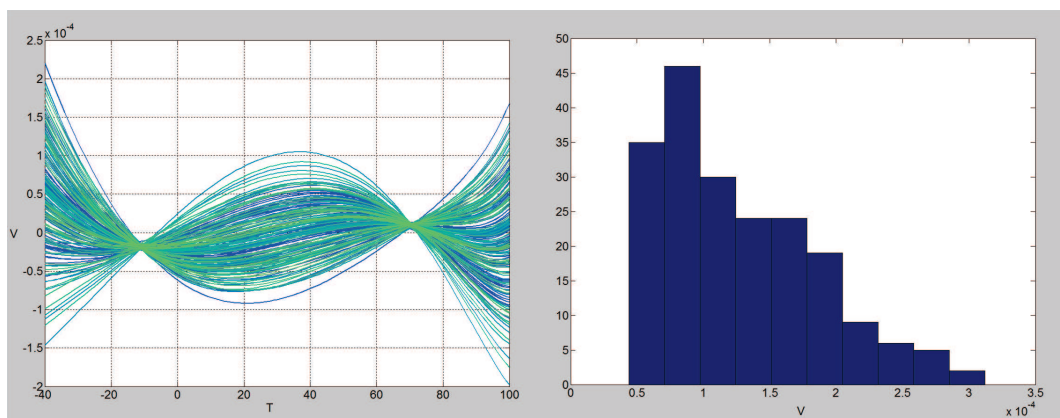


Figure 5.6: Evaluation of the residual curvature of the nonlinear-order using Monte Carlo Analysis  $\mu = 1.28 \times 10^{-4}V$   
 $var = 3.46 \times 10^{-9}$

The residual curvature of first-order is about  $10^{-3}V$ , while the residual curvature of non-linear order is about  $10^{-4}V$ . It is possible to compare the residual curvature of the designed nonlinear-order bandgap voltage reference with the residual curvature in another nonlinear-order bandgap voltage reference realized with another technology:

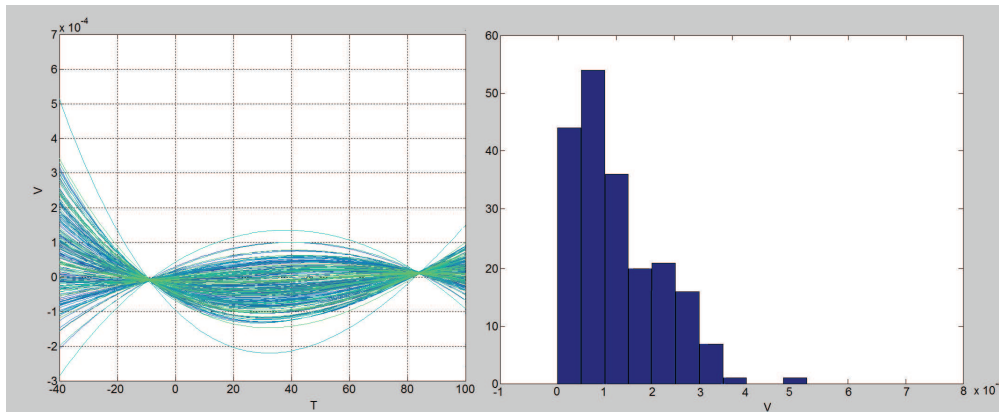


Figure 5.7: Evaluation of the residual curvature of another nonlinear-order using Monte Carlo Analysis  $\mu = 1.28 \times 10^{-4}V$   $var = 3.46 \times 10^{-9}$

referring to fig. 5.6 and to fig. 5.7 the residual curvature has the same order of magnitude  $10^{-4}V$  and the residual curvature has been calculated in the same range of temperatures. The residual curvature of fig. 5.6 is narrower than the residual curvature of fig. 5.7. Bandgap voltage reference of fig. 5.6 is less sensitive than 5.7 to process and mismatch variations.

### 5.3 Residual curvature of Corner Analysis

Corners define differences due to process inaccuracies, temperature and other parameter variations. It is clear that simulations that take these differences into considerations will differ one from another. Process parameters such as threshold voltage, mobility, etc. are different functions of temperature for each different corner.

Corners that describes differences due to process inaccuracies (such as doping variations) are supplied with the process kit and usually located in models library. For example the kit can include corners for fast N-Channel MOSFET fast P-Channel MOSFET, slow N-Channel MOSFET slow P-Channel MOSFET, fast N-Channel MOSFET slow P-Channel MOSFET, slow N-Channel MOSFET fast P-Channel MOSFET, typical N-Channel MOSFET typical P-Channel MOSFET.

There is also the possibility that corners will describe IC's behaviour in different temperatures and other parameters variations, such as  $V_{DD}$  variations. Each corner that will be simulated can contain one technology corner, one temperature value and one value for every parameter. During corner simulation, all available corners are simulated and thus influence of parameter variations on IC can be checked.

It is important to perform such simulations, because if a design meets all requirements for all technology corners available in the kit during simulation stage, the probability that all requirements will be met during chip test increases.



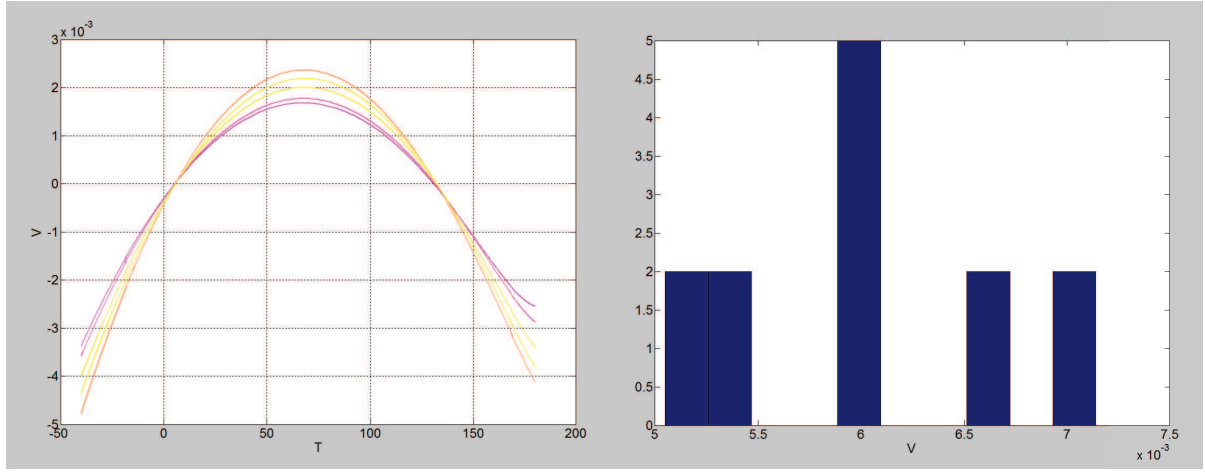


Figure 5.8: Evaluation of residual curvature of a first-order using Corner Analysis

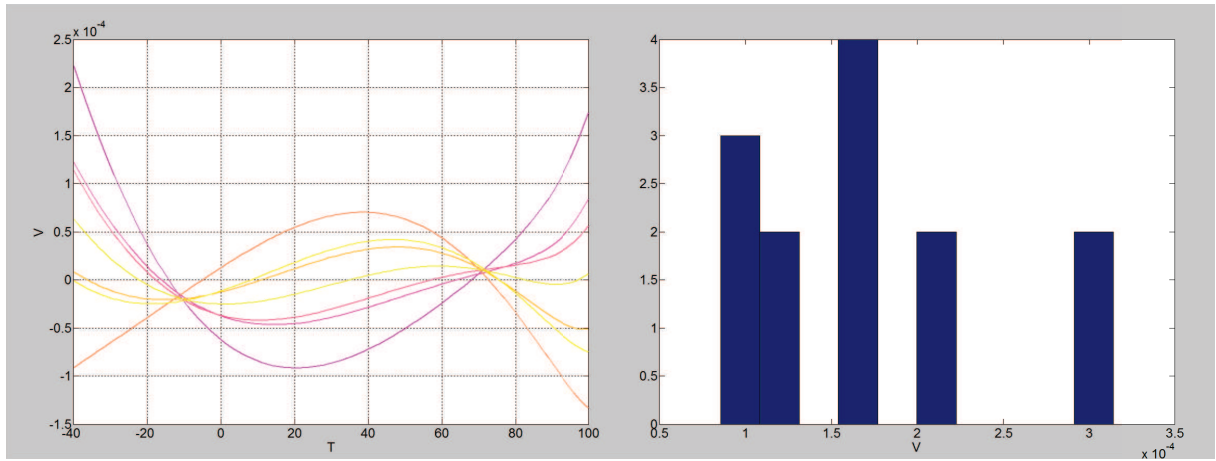


Figure 5.9: Evaluation of residual curvature of nonlinear-order using Corner Analysis

Corner Analysis investigates the effects of process only, while Monte Carlo Analysis can investigate the effects of process and mismatch: however figures 5.5 and 5.7 (residual curvature evaluated with Monte Carlo) are very similar to figures 5.8 and 5.9 respectively (residual curvature evaluated with Corners). In this way it is possible to presume that the residual curvature depends mainly on process effects.

It is possible to analyze the process-dependent parameters of base-emitter voltage.  $V_{BE}$  can be expressed as a function of temperature and collector current as follows:

$$V_{BE} = V_t \ln \left( \frac{I_c}{J_s A_e} \right) \quad (5.1)$$

$A_e$ , the emitter area, is a process-dependent parameter. Moreover  $J_s$  is a function of different process-dependent parameters:

$$J_s = \frac{q^2 n_i^2}{Q_B} \tilde{D}_n \quad (5.2)$$

where  $q$  is the electron charge,  $n_i$  is the concentration of carriers in an intrinsic semiconductor,  $Q_B$  is the total charge per unit area in base and  $\tilde{D}_n = V_t \mu_n = \frac{k_B T}{q} \mu_n$  is the diffusion coefficient and it depends on temperature and on the mobility of carriers  $\mu_n$ . In doped material, the mobility of carriers is a highly dependent process-parameters, since it depends on the crystal lattice scattering and on the ionized impurity scattering.

## 5.4 Trimming Algorithm

The trimming networks of  $R_{NL}$  and  $R_{LOAD}$  have been designed to compensate the random effects of process and mismatch. Their goal is to reduce the spread of bandgap output voltage and to improve its slope accuracy. In other words, trimming networks are useful to guarantee a flat and a close to nominal value (800mV) of the bandgap output voltage.

A MATLAB trimming algorithm can be used in order to verify the trimming networks: Cadence, in fact, doesn't implement this kind of analysis directly. This algorithm verifies the trimming words using as starting point a variable number of Monte Carlo Analysis.  $N$  matrixes of Monte Carlo Analysis are necessary, where  $N$  is the number of bits used to trim  $R_{LOAD}$ . For each  $R_{LOAD}$ 's trimming words (i.e for each configuration of trimming bits of  $R_{LOAD}$ ) it is necessary to run  $M$  Monte Carlo Analysis sets where  $M$  is the number of  $R_{NL}$ 's trimming words. Each Monte Carlo Analysis set has 200 runs.

The implemented algorithm (used 3 trimming bits for  $R_{NL}$  and 3 trimming bits for  $R_{LOAD}$ ) can be explained in three steps:

- first step: it is necessary to choose one of the trimming words of  $R_{LOAD}$ . Using this bits configuration, a Monte Carlo Analysis has to be run for each trimming words of  $R_{NL}$ . In this way the result is a matrix of Monte Carlo Analysis, which collects all the 200 runs for all the values of  $R_{NL}$  and for a given value of  $R_{LOAD}$ . Then it is necessary to repeat this sequence for the remaining trimming words of  $R_{LOAD}$ . At the end of this step, 8 matrixes of Monte Carlo Analysis have been collected.

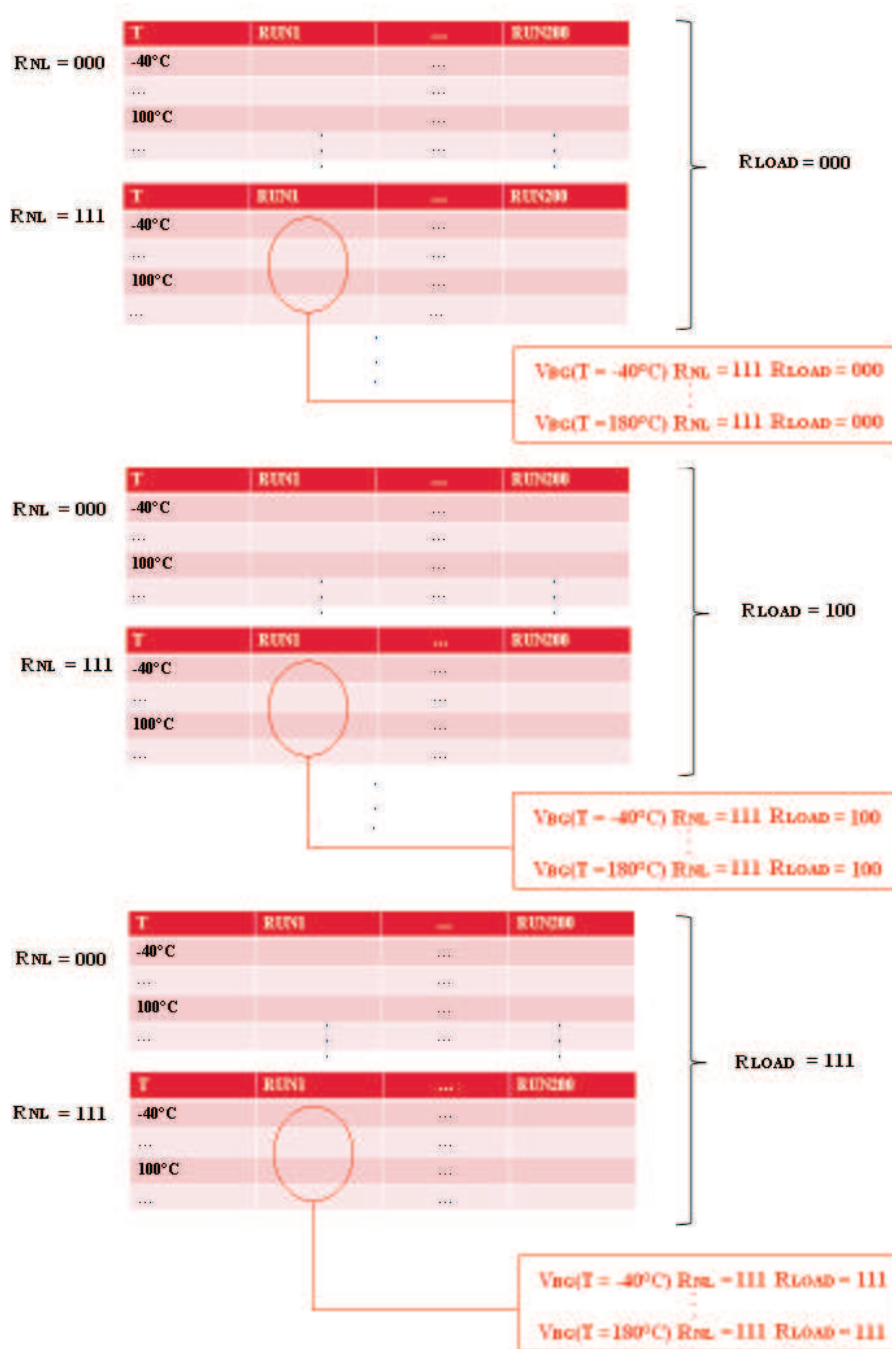


Figure 5.10: First step of trimming verification algorithm

- second step: for each group of matrixes with given  $R_{LOAD}$  trimming word (for example, referring to fig. 5.10, the group with  $R_{LOAD} = 111$ ) the difference  $\Delta V_{BG}$  between  $V_{BG}(T_{max} = 100^{\circ}\text{C})$  and  $V_{BG}(T_{min} = -40^{\circ}\text{C})$  has

been calculated, i.e.  $\Delta V_{BG} = V_{BG}(T_{max} = 100^{\circ}\text{C}) - V_{BG}(T_{min} = -40^{\circ}\text{C})$ .  $T_{max} = 100^{\circ}\text{C}$  is chosen instead of  $180^{\circ}\text{C}$  to avoid the effect of leakage. In this way it is possible to evaluate the difference  $\Delta V_{BG}$ , for each trimming words. The goal of  $R_{NL}$  trimming network is to minimize the slope, in order to get a flat bandgap output voltage. The slope  $\Delta V_{BG}$  can be minimized choosing the correct  $R_{NL}$  trimming word for each run.

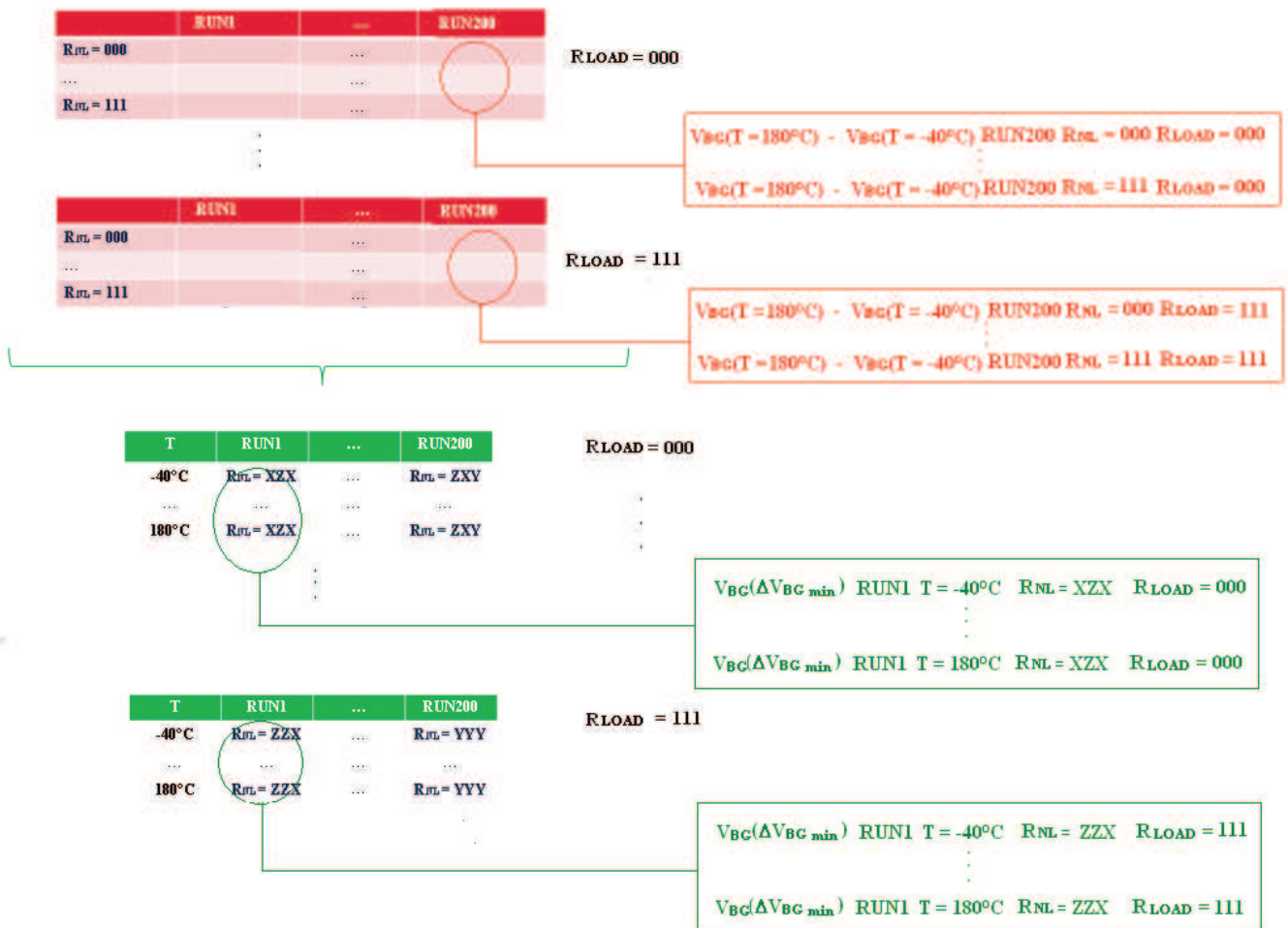


Figure 5.11: Second step of trimming verification algorithm

- third step: now each run of the 200 runs of the Monte Carlo Analysis matrixes is very flat, because  $R_{NL}$  trimming word to minimize its slope has been chosen. As result, there are 8 matrixes of Monte Carlo Analysis: each of them has 200 runs and these runs are all parallel. It is possible to choose the trimming word of  $R_{LOAD}$ , in order to minimize the difference  $V_{BG} - 800mV$ . As result, each run has its slope minimized (choosing a particular  $R_{NL}$  trimming word) and has minimized its spread from the nominal value (choosing a particular  $R_{LOAD}$  trimming word).

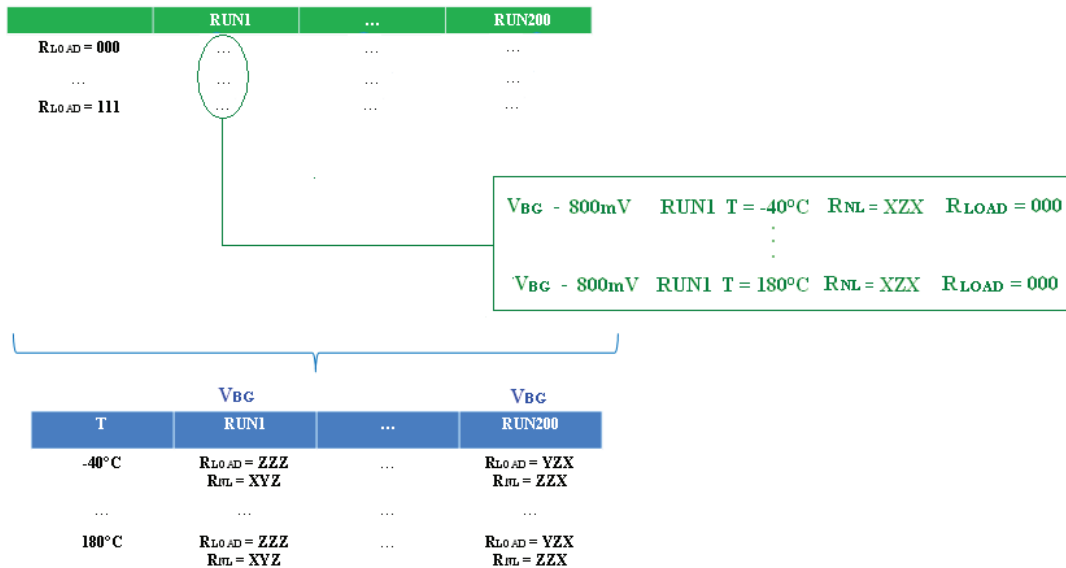


Figure 5.12: Third step of trimming verification algorithm

Trimming verification algorithm uses Monte Carlo Analysis to verify the trimming words. It is necessary that the statistical models used by Cadence in Monte Carlo Analysis are always the same, otherwise the algorithm doesn't work. For example, the run number  $x$  (where  $x$  is a number between 1 and 200) has to be the same both for trimming words  $R_{LOAD} = 111$   $R_{NL} = 000$  and trimming words  $R_{LOAD} = 111$   $R_{NL} = 100$ . In order to verify the trimming networks, it is necessary that the trimming word is the only modification in the netlist, i.e. the seed does not have to change. The seed determines the starting point from the random variables will be generated.

It is quite easy to verify that the seed is always the same, through trimming words change. Referring to the first step of trimming algorithm, it is possible to analyze a group of matrixes of Monte Carlo with a given  $R_{LOAD}$  (for example the group with  $R_{LOAD} = 111$ ). If the run number  $x$  (where  $x$  is random number between 1 and 200) has always to be generated from the same seed for each analysis, each run number  $x$  differ only for the value of  $R_{NL}$  and so they must have the same compensation temperature. For example, all the run number 1, 2, 3, ... 200 have the same compensation temperature, however the compensation temperature of run 1 is different from all the other (the same for the compensation temperature of run number 2, 3, ... 200):

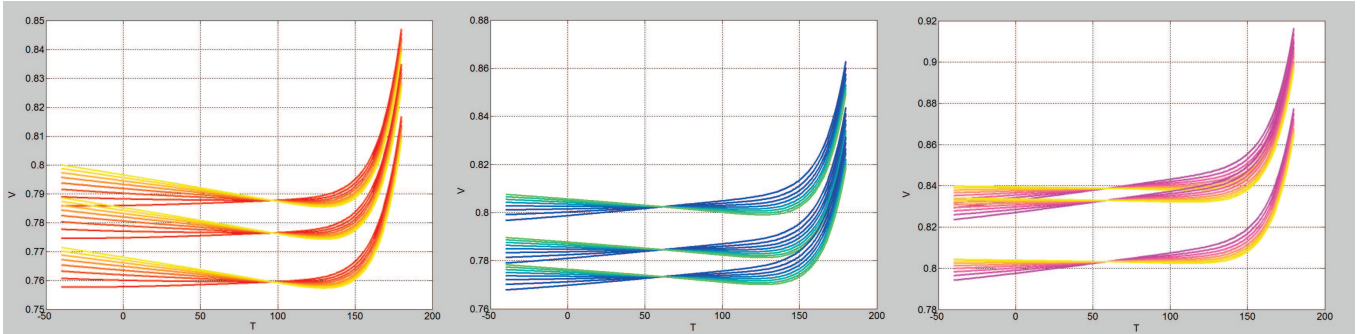


Figure 5.13: Verification of Monte Carlo seed

## 5.5 Bandgap output voltage after double trimming

### 5.5.1 First trimming option

First trimming option aims to use 3 bits to trim  $R_{NL}$  and 3 bits to trim  $R_{LOAD}$ . The two trimming networks are used to reduce the amplitude distribution (spread) and the slope distribution of bandgap output voltage. Running a Monte Carlo Analysis, it is possible to evaluate the untrimmed statistical distribution of bandgap output voltage slope  $\Delta_S = V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$  and the initial statistical distribution of bandgap output voltage amplitude  $\Delta_A = V_{BG}(T = 27^\circ\text{C}) - 800\text{mV}$ .

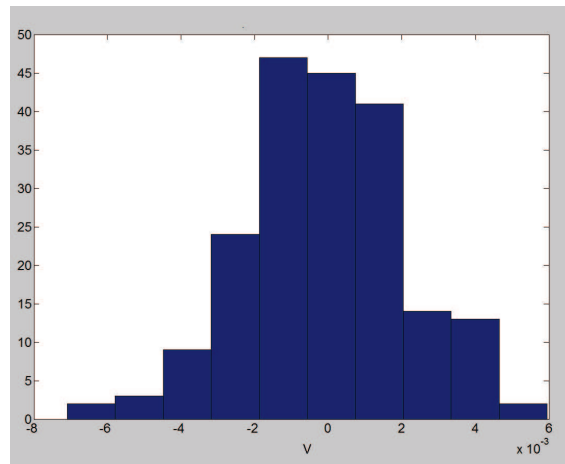


Figure 5.14: Untrimmed statistical distribution of bandgap output voltage's slope  $\Delta_S = V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$   $\bar{\mu}_S = 1.311 \times 10^{-4}\text{V}$   $\sigma_S = 2.192 \times 10^{-3}$

$N = 3$  trimming bits (i.e.  $2^3 = 8$  trimming words) are used to trim  $R_{NL}$ . referring to fig. 5.14  $\Delta V \simeq 13\text{mV}$ . Using this network, the step is  $\frac{\Delta V}{2^N - 1} = \frac{13\text{mV}}{2^3 - 1} = 1.86\text{mV}$  and the goal is to centre the statistical distribution between  $\pm \frac{1.86\text{mV}}{2} = \pm 0.926\text{mV}$ . In order to have binary weighted trims, the three resistances to trim  $R_{NL}$  are:  $\Delta R_1 = 11\text{K}\Omega$ ,  $\Delta R_2 = 22\text{K}\Omega$  and  $\Delta R_3 = 44\text{K}\Omega$ .

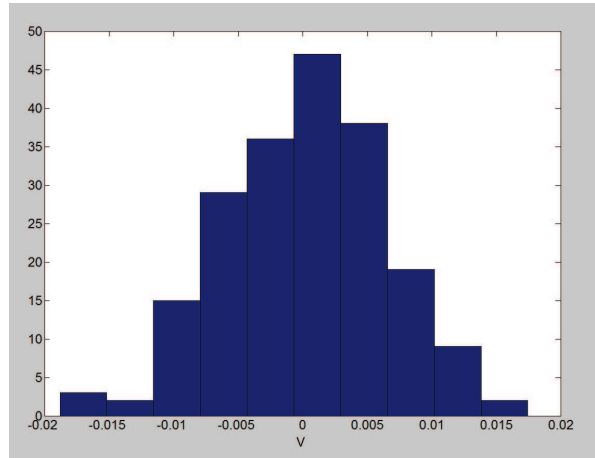


Figure 5.15: Untrimmed statistical distribution of bandgap output voltage's amplitude  $\Delta_A = V_{BG}(T = 27^\circ C) - 800mV$   $\bar{\mu}_A = 2.8383 \times 10^{-4}V$   $\sigma_A = 5.731 \times 10^{-3}$

$N = 3$  trimming bits (i.e.  $2^3 = 8$  trimming words) are used to trim  $R_{LOAD}$ . referring to fig. 5.15  $\Delta V \simeq 40mV$ . Using this network, our step is  $\frac{\Delta V}{2^N - 1} = \frac{40mV}{2^3 - 1} = 5.71mV$  and the goal is to centre the statistical distribution between  $\pm \frac{5.71mV}{2} = \pm 2.86mV$ . In order to have binary weighted trims, the three resistances to trim  $R_{LOAD}$  are:  $\Delta R_1 = 2.5K\Omega$ ,  $\Delta R_2 = 5K\Omega$ ,  $\Delta R_3 = 10K\Omega$ . Choosing these values, the minimum trimming voltage step becomes  $6mV$ . The final amplitude distribution has to be centered between  $\pm \frac{6.0mV}{2} = \pm 3.0mV$ .

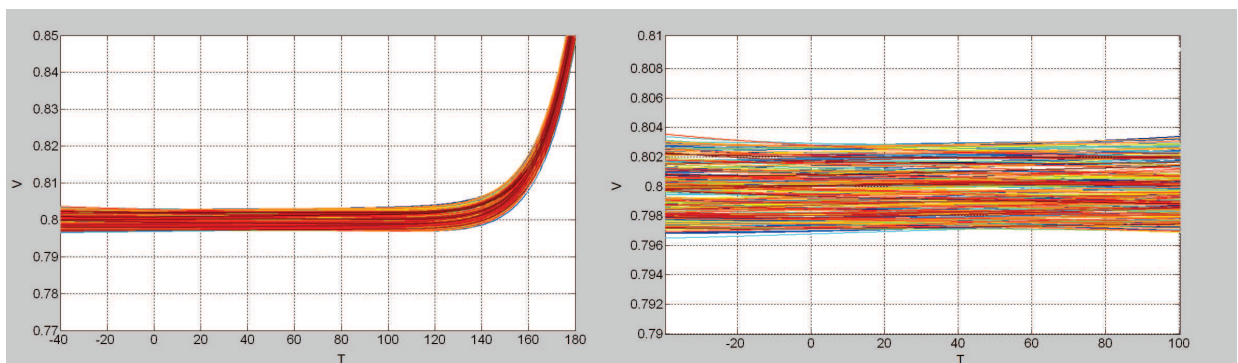


Figure 5.16: Bandgap output voltage after double trimming (first trimmig option)

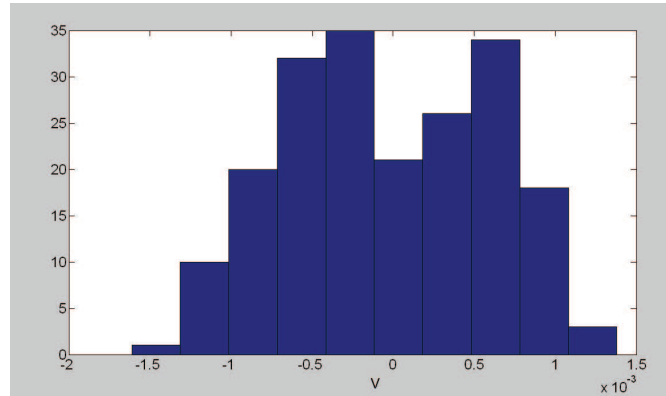


Figure 5.17: Statistical distribution of  $\Delta_S = V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$  after double trimming  $\bar{\mu}_S = -3.067 \times 10^{-5}\text{V}$  (first trimming option)

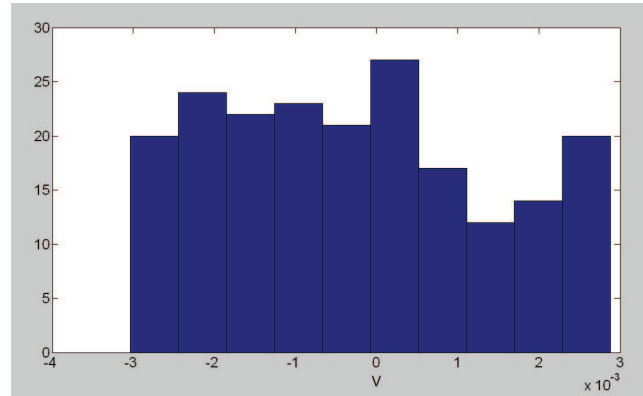


Figure 5.18: Statistical distribution of  $\Delta_A = V_{BG}(T = 27^\circ\text{C}) - 800\text{mV}$  after double trimming  $\bar{\mu}_A = -2.677 \times 10^{-4}\text{V}$  (first trimming option)

Fig. 5.17 shows that the distribution is centered between  $\pm 1.5\text{mV}$  instead between  $\pm 0.926\text{mV}$ . This error is due to the nonlinear relation  $V_{NL} = f(I_{NL}, R_{NL})$ . The voltage drop across  $R_{NL}$  isn't a linear function and so the binary weighted trims don't produce a constant trimming voltage steps.

### 5.5.2 Second trimming option

In this case,  $N = 2$  trimming bits (i.e.  $2^2 = 4$  trimming words) are used to trim  $R_{NL}$ . referring to fig. 5.14  $\Delta V \simeq 13\text{mV}$ . Using this network, the step is  $\frac{\Delta V}{2^N - 1} = \frac{13\text{mV}}{2^2 - 1} = 4.33\text{mV}$  and the goal is to centre the statistical distribution between  $\pm \frac{4.33\text{mV}}{2} = \pm 2.17\text{mV}$ . However, the nonlinear relation  $V_{NL} = f(I_{NL}, R_{NL})$  will give an error and the resulting slope distribution will be larger than  $\pm 2.17\text{mV}$ . In order to have binary weighted trims, the two resistances to trim  $R_{NL}$  are:  $\Delta R_1 = 26\text{K}\Omega$ ,  $\Delta R_2 = 52\text{K}\Omega$ .

$N = 4$  trimming bits (i.e.  $2^4 = 16$  trimming words) are used to trim  $R_{LOAD}$ . referring to fig. 5.15  $\Delta V \simeq 40\text{mV}$ . Using this network, minimum voltage step is  $\frac{\Delta V}{2^N - 1} = \frac{40\text{mV}}{2^4 - 1} = 2.67\text{mV}$  and the goal is to centre the statistical distribution between  $\pm \frac{2.67\text{mV}}{2} = \pm 1.33\text{mV}$ . In order to have binary weighted trims, the four resistances to trim  $R_{LOAD}$



are:  $\Delta R_1 = 1.1K\Omega$ ,  $\Delta R_2 = 2.2K\Omega$ ,  $\Delta R_3 = 4.4K\Omega$ ,  $\Delta R_4 = 8.8K\Omega$ . Choosing these values, the minimum trimming voltage step becomes  $3mV$ . The final amplitude distribution has to be centered between  $\pm \frac{3.0mV}{2} = \pm 1.5mV$ .

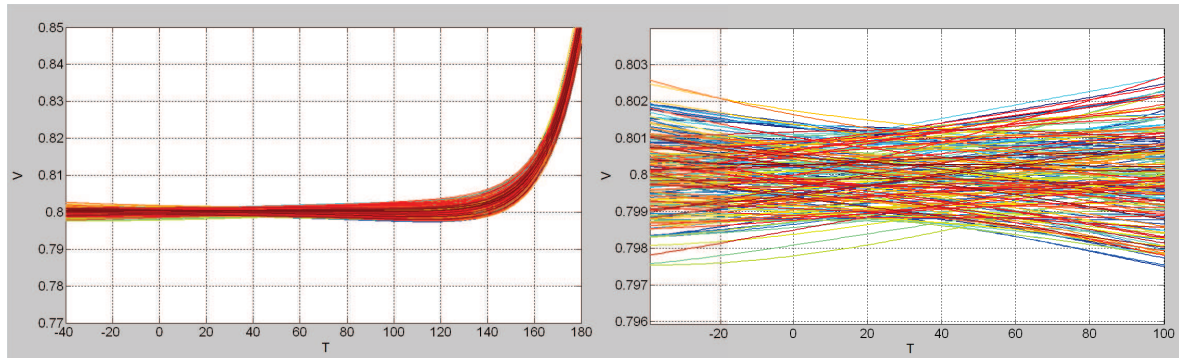


Figure 5.19: Bandgap output voltage after double trimming (second trimmig option)

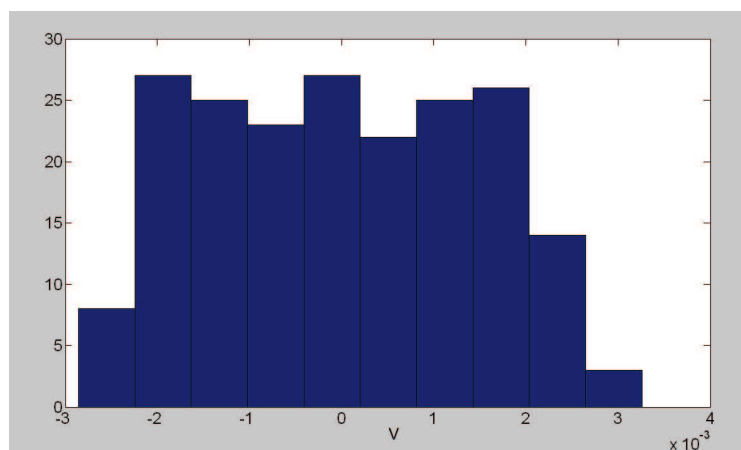


Figure 5.20: Statistical distribution of  $\Delta_S = V_{BG}(T = 100^\circ\text{C}) - V_{BG}(T = -40^\circ\text{C})$  after double trimming  $\bar{\mu}_S = -4.929 \times 10^{-6}V$  (second trimming option)

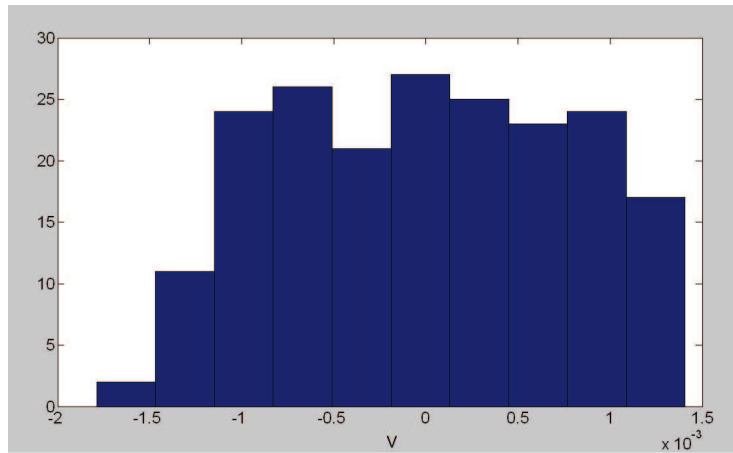


Figure 5.21: Statistical distribution of  $\Delta_A = V_{BG}(T = 27^\circ\text{C}) - 800\text{mV}$  after double trimming  $\bar{\mu}_A = -6.576 \times 10^{-6}\text{V}$  (second trimming option)

### 5.5.3 Comparison of trimming options

Fig. 5.19 shows that the second trimming option produces a bandgap output voltage not as flat as the first trimming option. In this way, the statistical distribution  $\Delta_A$  changes if it is evaluated at different temperatures. Instead, the first trimming option produces a bandgap output voltage flatter (there is one bit more to reduce the slope) and therefore the statistical distribution  $\Delta_A$  is quite constant at different temperatures.

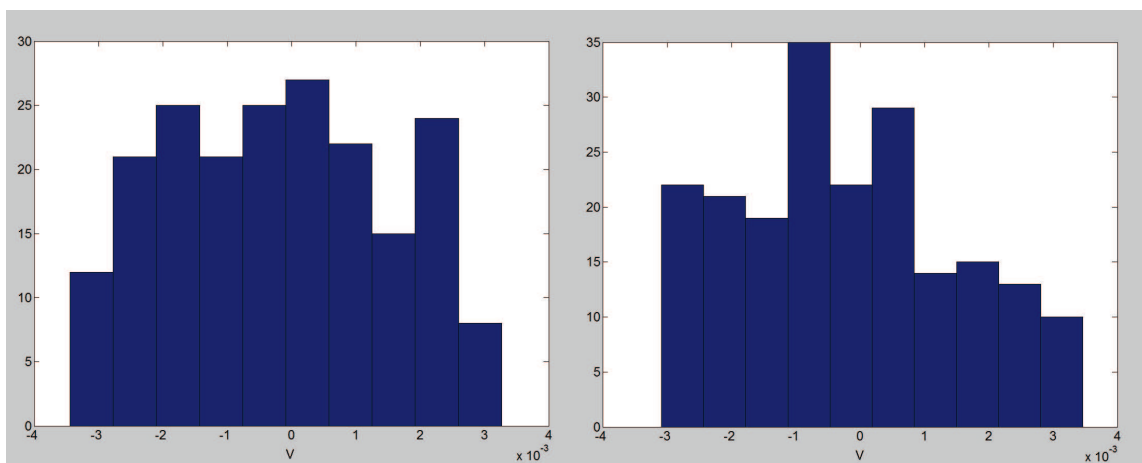


Figure 5.22: Statistical distributions for first trimming option of  $\Delta_{A1} = V_{BG}(T_1 = -40^\circ\text{C}) - 800\text{mV}$   $\bar{\mu}_{A1} = -2.190 \times 10^{-4}\text{V}$  (left) and of  $\Delta_{A2} = V_{BG}(T_2 = 100^\circ\text{C}) - 800\text{mV}$   $\bar{\mu}_{A2} = -2.496 \times 10^{-4}\text{V}$  (right)

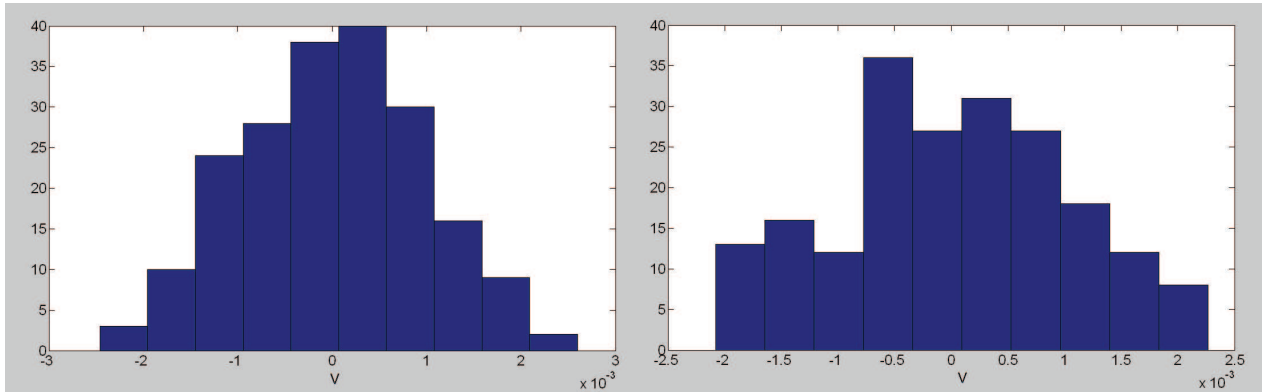


Figure 5.23: Statistical distributions for first trimming option of  $\Delta_{A_1} = V_{BG}(T_1 = -40^\circ\text{C}) - 800\text{mV}$   $\mu_{A_1} = 2.450 \times 10^{-5}\text{V}$ (left) and of  $\Delta_{A_2} = V_{BG}(T_2 = 100^\circ\text{C}) - 800\text{mV}$   $\mu_{A_2} = 1.809 \times 10^{-5}\text{V}$  (right)

The first trimming option produces a bandgap output voltage with an error margin of  $\pm 3\text{mV}$  on amplitude and of  $\pm 1.5\text{mV}$  on slope for a temperature range of  $140^\circ\text{C}$  ( $T_{min} = -40^\circ\text{C}$  and  $T_{max} = 100^\circ\text{C}$ ). Instead second trimming option produces a bandgap output voltage with an error margin of  $\pm 1.5\text{mV}$  on amplitude and of  $\pm 3.0\text{mV}$  on slope. However amplitude margin of error depends on the temperature: at  $T = -40^\circ\text{C}$  it is  $\pm 2.5\text{mV}$ , at  $T = 27^\circ\text{C}$  it is  $\pm 1.5\text{mV}$  and at  $T = 100^\circ\text{C}$  it is  $\pm 2\text{mV}$

## 5.6 Residual curvature after double trimming

It is interesting to evaluate how the trimming networks influence the residual curvature. The residual curvature can be calculated with the rectification algorithm using the Monte Carlo Analysis of fig. 5.16 and the Monte Carlo of fig. 5.19.

Residual curvature should not be changed by trimming networks. However, this is true as long as there is a linear relation between the trimmed parameter and the drop voltage across it.

referring to fig. 5.5, the first-order bandgap voltage reference has a three bits to trim  $R_{PTAT}$ : there is a linear relation between the trimmed resistance and the drop voltage across it. In this case, trimming network doesn't change the residual curvature.

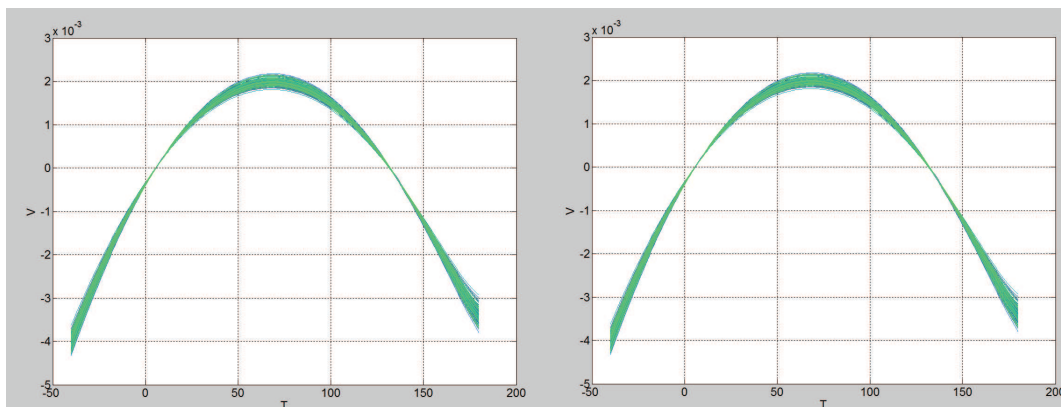


Figure 5.24: Residual curvature of first-order voltage reference: before trimming and after trimming

Instead the residual curvature of nonlinear-order bandgap voltage reference is different after double trimming, because  $R_{NL}$  trimming network does not trim a linear voltage (the voltage across  $R_{NL}$  is the difference of two base-emitter voltage). Moreover trimming  $R_{NL}$  value, the base currents of T3 and T4 will change: in this way will change the bias of these transistors. Both trimming options increase the residual curvature, however it remains of the same order of magnitude ( $10^{-4}V$ ).

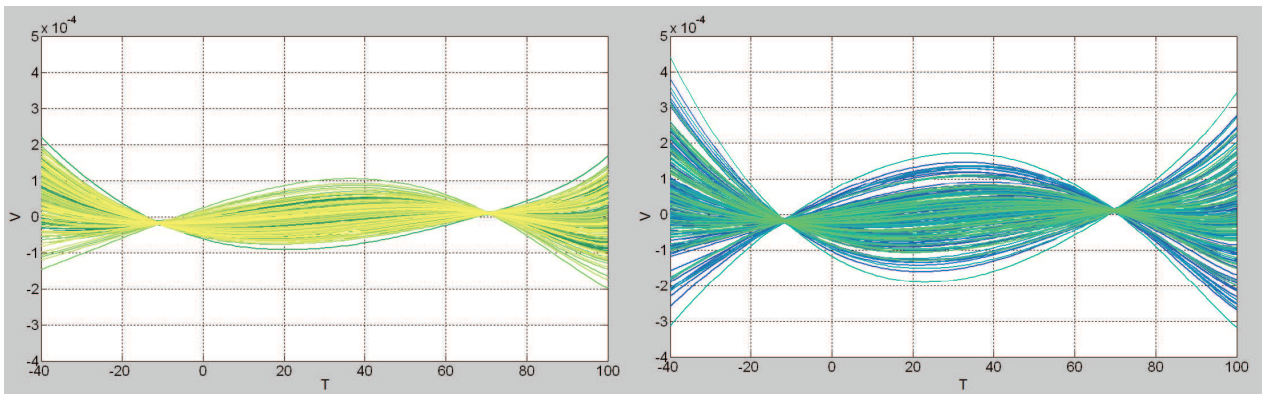


Figure 5.25: Residual curvature of nonlinear-order bandgap voltage reference (first trimming option): before trimming (left) and after trimming (right)

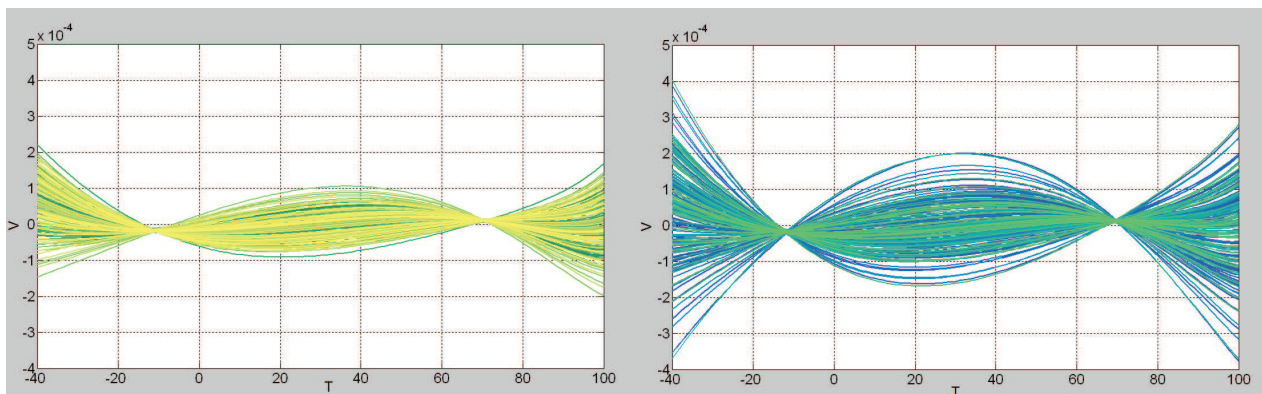


Figure 5.26: Residual curvature of nonlinear-order bandgap voltage reference (second trimming option): before trimming (left) and after trimming (right)



## **Chapter 6**

# **CONCLUSIONS**

In this thesis, two bandgap output voltage references have been studied. Their accuracy and their asymptotic stability over the whole temperature range have been analyzed with Cadence simulations. Each stage of the first-order bandgap voltage reference and each stage of diode loop topology have been studied and optimized, in order to improve their accuracy. Using Monte Carlo Analysis and Corner Analysis, the effects of process and mismatch have been evaluated. In order to avoid these negative effects, two trimming networks have been designed and compared. A Matlab script has been developed in order to verify the trimming words and to compare the different behaviour of the trimming networks. The trimming options have shown that they can effectively limit the effects of process and mismatch, because they can reduce the inaccuracy of slope and the amplitude spread of bandgap output voltage.



# Acknowledgements

To conclude, I would like to thanks:

- Prof. Andrea Gerosa, for his supporting and interest in my work.
- Infineon Technologies, especially Bartolomeo Cottone and Pietro Piersimoni, for letting me live this experience.
- Marco Piselli, for his constant presence and for the many things he taught me.
- Stefano Zampieri, for his many suggestions and constant support during this research.
- Infineon STD team and Infineon DC-DC team, for the precious help and happy moments during these months.
- Matteo Giarin and Tommaso Pendini, who helped me during the laboratory verifications.
- My mother, my father, Chiara and Enrico who encouraged and helped me during these years. Without them this would not have been possible.
- Juan Carlos and my friends of San Francesco group, in particular Davide, Federico, Francesco, Gianmaria and Giulio. During these years, we have passed a lot of time together and happy moments.





# Bibliography

- [1] A. P. BROKAW, *A Simple Three-Terminal IC Bandgap Reference*, IEEE J. of Solid-State Circuits, vol. sc-9, No. 6, pp. 388 - 393, December 1974.
- [2] DONALD T. COMER, *Zener Zap Anti-Fuse Trim in VLSI Circuits*, VLSI DESIGN, vol.5, No. 1, pp. 89 - 100, April 1995.
- [3] PAUL R. GRAY, ROBERT G. MAYER, *Analysis and Design of Analog Integrated Circuits*, John Wiley and Sons, p. 389, New York 1984.
- [4] V. GUPTA, G. A. RINCÓN MORA, *Predicting and Designing for the Impact of Process Variations and Mismatch on the Trim Range and Yield of Bandgap Reference*, IEEE Proceedings of the Sixth International Symposium on Quality Electronic Design, pp. 45 - 107, 2005.
- [5] HAMILTON KLIMACH, *A 4-Bits Trimmed CMOS Bandgap Reference with an Improved Matching Modeling Design*, Proceedings of ISCAS, pp. 1911 - 1914, 2007.
- [6] G. C. M. MEIJER, J. FONDERIE, J. H. HUIJSING, *A Curvature-Corrected Low-Voltage Bandgap Reference*, IEEE J. of Solid-State Circuits, vol. 28, No. 6, pp. 667-670, June 1993.
- [7] G. C. M. MEIJER, P. C. SCHIMALE, K. VAN ZALINGE, *A New Curvature-Corrected Bandgap Reference*, IEEE J. of Solid-State Circuits, vol. sc-17, No. 6, pp. 1139 - 1143, December 1982,
- [8] G. C. M. MEIJER, K. VINGERLING, *Measurement of the Temperature Dependence of the  $I_C(V_{be})$  Characteristics of Integrated Bipolar Transistors*, IEEE J. Solid-State Circuits, vol. sc-15, No. 2, pp. 237 - 240, April 1980.
- [9] R. S. MULLER, THEODORE I. KAMINS, *Device Electronics for Integrated Circuits third edition*, John Wiley and Sons, pp. 204 - 210, New York 2003.
- [10] P. RAJARSHJ, P. AMIT, B. SHAIENDRA, D. KAUSHIK, *Design of Second-Order Sub-Bandgap Mixed Mode Voltage Reference Circuit for Low Voltage Applications*, Proceedings of the 18th International Conference on VLSI Design, pp. 1 - 5, 2003.
- [11] G. A. RINCÓN MORA, *Voltage Reference*, Wiley Interscience, pp. 45 - 110, Montreal 2000.
- [12] Y. P. TSIVIDIS, *Accurate Analysis of Temperature Effects in  $I_C - V_{BE}$  Characteristics with Application to Bandgap Reference Sources*, IEEE J. Solid-State Circuits, vol. sc-15, No. 6, pp. 1076 - 1084, December 1980.
- [13] Y. P. VARSHNI, *Temperature dependence of the energy gap in semiconductors*, Physica, vol. 34, pp. 149 - 154, 1967.

- [14] R. WIDLAR, *New Developments in IC Voltage Regulators*, IEEE J. of Solid-State Circuits, vol. sc-6, NO. 1, pp. 2 - 7, February 1971.
- [15] X. XIAOYU, *Low voltage low power bandgap circuit*, US 6,989,708 B2, pp. 1 - 4, August 2003.
- [16] S. ZAMPIERI, *Investigation, Analysis and Design of a Sub-Bandgap Voltage Reference for Ultra-Low Voltage Applications*, Padova 2012.