High Temperature SOI CMOS Band-Gap Voltage Reference

By

Jerry Brewer

Master of Science

Oklahoma State University

Stillwater, Oklahoma

December, 2004

Submitted to the Faculty of the

Graduate College of the

Oklahoma State University

in partial fulfillment of

the requirements for

the Degree of

Master of Science

December, 2004

HIGH TEMPERATURE SOI CMOS BAND-GAP VOLTAGE REFERENCE

Thesis Approved:

Dr. Chris Hutchens

Advisor's Approval

Dr. Louis Johnson

Dr. Keith Teague

Dr. Gordon Emslie

Dean of Graduate College

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The voltage reference is one of the key building blocks in electronic circuit design. A voltage reference is a circuit that provides a fixed output voltage. A voltage reference is often used to compare input or output voltages to a certain level. The voltage reference is widely used in sensors, data converters, biomedical circuits, and power supplies [1]. A simple application where a voltage reference is used to compare the voltage of a system to a safe level is shown in Figure 1.1. This paper details the design, fabrication, and measurement of a voltage reference in CMOS technology.

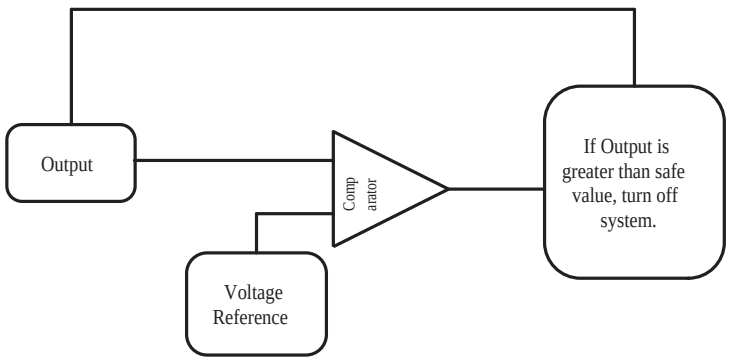


Figure 1.1: Safety Detector

The voltage reference has three main figures of merit: the output voltage's dependence on temperature, voltage supply, and the circuit's power consumption. Other factors, such as area, are important, but these three are the dominant design criteria. Unfortunately, the criteria often work against each other. For example, achieving low dependence on temperature and voltage supply can easily lead to increased power consumption [2]. Current published efforts are summarized in the following table.

	Leung [2]	Jiang [3]	Annema [4]	Malcovati [5]
Technology	.6 μm CMOS	1.2 μm CMOS	.35 μm CMOS	.8 μm CMOS
Minimum Supply Voltage	.98 V	1.2 V	.85 V	.95 V
Supply Current	18 μA	500 μA	1.2 μA	92 μA
Tempco	15 ppm/°C	100 ppm/°C	57 ppm/°C	19 ppm/°C

Table 1.1: Comparison of Band-Gap References

For reasons to be discussed later, voltage references have historically been integrated in bipolar or BiCMOS processes [5]. This presents a challenge for designers using CMOS fabrication process. Recently CMOS technology has emerged as a dominant technology due largely to CMOS's benefits in digital circuits. CMOS transistors exhibit far greater scaling properties than comparable technologies. This allows circuits to occupy less space and operate at higher speeds due to the decreased parasitics (parasitic capacitance is a function of area). Also, digital CMOS circuits benefit from very small static power dissipation [6]. But since the voltage reference is such an important element, circuit designers must find new methods of achieving success in the CMOS process. More information on the CMOS process is included in the **CMOS Review** chapter.

The voltage reference is implemented as follows. Two quantities that are predictable with temperature are formed. These two values are used to cancel each other or another quantity to produce a output voltage that is independent of temperature. The dependence of the output voltage to the power supply is largely set by the topology (which is usually optimized for temperature performance), but several steps can be taken to ensure a constant output voltage for a wide range of supply values. These include using large transistors, which typically operate better than minimum

geometry devices, and using feedback to stabilize the circuit [3]. As mentioned, efforts to decrease the dependence of the output voltage on temperature and supply voltage can increase the power consumption. Adding complexity to the temperature compensation scheme will undoubtedly lead to more power required. Obviously, one would wish to have the voltage reference operate on a small as power supply as possible. This becomes even more important as voltage supplies shrink with scaling. As devices become smaller, the power-per-area increases. This value can approach values where the devices operate at unsafe levels [7]. Thus, compromises have to be considered when designing the voltage reference. The design process of the voltage reference is discussed in the chapter **Band-Gap Voltage Reference**.

Historically, the majority of voltage references have been implemented in a bipolar process. The reason for this is the ability to fabricate a diode. Diodes can be used to establish both a current and a voltage that are proportional to temperature. This class of voltage references is called *band-gap* references, due to the band-gap diode employed in the circuit [6]. Unfortunately, in a CMOS process, diodes are not readily available. To circumvent this shortcoming, CMOS circuit designers typically use either a parasitic CMOS diode or a gated-diode structure to establish predictable voltages and currents. Both the parasitic CMOS diode and the gated-diode have advantages and disadvantages over the other, and the designer must choose the suitable device for the application. The choice and implementation of the band-gap device is discussed further in Chapter 4, **Band-Gap Core**.

As with many circuits, an amplifier is often used in voltage references. Amplifiers can be used to provide gain or can be used in feedback to stabilize a circuit. An amplifier can be used in the voltage reference to decrease the output sensitivity to temperature, supply, and process variations [2]. Adding an amplifier to the circuit will obviously add more power to the design and more area required at the IC level. So a compromise must be made when using an amplifier in the voltage reference to determine the performance versus power and area. The design and implementation of the amplifier used in the voltage reference is provided in Chapter 5, **Amplifier Design**.

Two voltage references was designed, simulated, and fabricated in the Peregrine SOS CMOS process. One reference employed an amplifier biased in the saturation region, while the other ref-

erence consists of an amplifier biased in the subthreshold region. The saturation design achieved simulation results of temperature dependence of 144 ppm/°C and a voltage supply dependence of 40 ppm/V. With a three volt supply, the voltage reference draws 945 μW. At the minimum supply of 1.75 V, the reference draws 940 μW. The subthreshold version achieved simulation results of a temperature dependence of 182 ppm/ °C, voltage supply dependence of 123 ppm / V, and consumes 180 μW of power. After fabrication, the two circuits were measured for comparison to simulated results. While the subthreshold reference performed as expected, the saturation version did not function properly. The measurements of the IC indicate a temperature dependence of 120 ppm/ °C, voltage supply dependence of 975 ppm/ V, and a power consumption of 310 μW. The design occupies 1.215 mm^2 of area. The difference between the simulation and measurement are explained in Chapter 6, **Measurements and Conclusions**. The details of the design of the voltage reference are discussed in this paper.

Chapter 2

CMOS Review

2.1 Introduction

The section is a review of CMOS theory. Nomenclature used throughout this thesis is explained here. The sources of error, such as noise, mismatch, and thermal degradation are discussed in this section.

2.2 SOI Fabrication

A layout schematic of a bulk silicon transistor is shown in Figure 2.1.

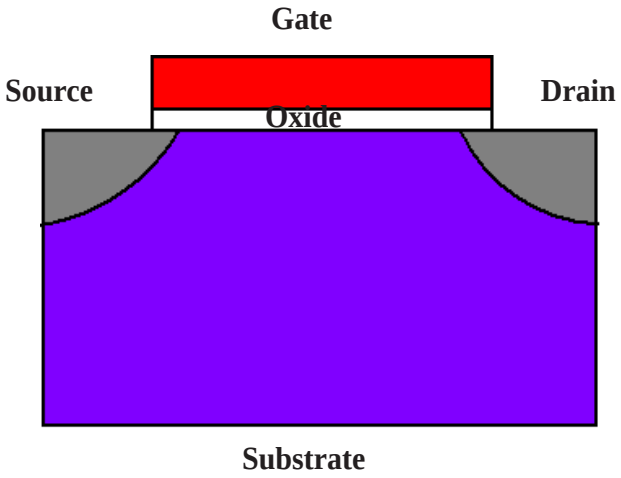


Figure 2.1: MOS Transistor

This type of transistor is known as a NMOS transistor. The device consists of a p-type (holes are major carrier) substrate (body) with two n-type implants. One implant is referred to as the source and the other the drain. The MOS device is a symmetrical device, as a result the source is defined as the terminal at the lower potential. On the surface of the silicon, a oxide layer is formed. To provide a electrical path to the oxide, a gate layer is formed using metal or polysilicon. If a potential is applied to the gate with respect to the source, the free electrons in the body will be attracted to the positive voltage, forming an induced n-type channel (hence the name NMOS). Thus, electrons can flow from the n-type source through the n-type induced channel to the n-type drain (a voltage drop must be present to induce the current flow). A PMOS device can be formed replacing the p-type substrate with a n-type substrate and the n-type implants with p-type. Often in bulk processes, a PMOS transistor is formed by implanting a n-type *well* in the p-type substrate or a p-type *well* in the n-type substrate [7].

Several problems exist with this fabrication process. First, the large substrate results in large parasitic junction capacitances. The capacitances result in both higher delay and increased power dissipation. To alleviate this problem, the SOI (Silicon-On-Insulator) device was introduced. Silicon-on-Insulator was first introduced to improve scaling and radiation hardening of transistors. The SOI

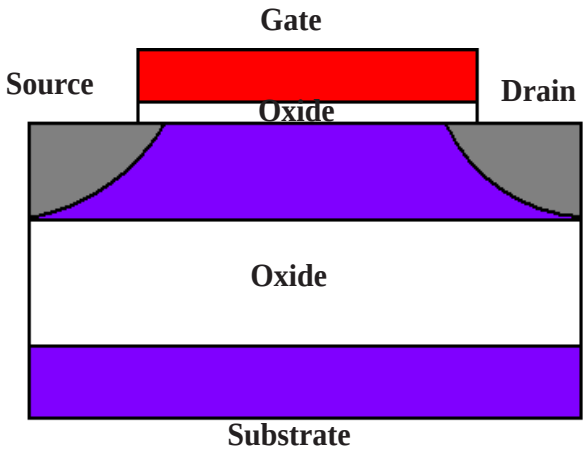


Figure 2.2: SOI MOS Transistor

transistor replaces the large substrate with a small *island*, greatly reducing the parasitic effects [8]. The SOI transistor is shown in Figure 2.2. Due to lower capacitance and other beneficial properties, SOI transistors have emerged as a leading fabrication process for high performance devices. The band-gap voltage reference was fabricated in the Peregrine SOI process.

2.3 DC Characteristics

The symbol for a MOS transistor is shown in Figure 2.3.

If voltages are applied to the transistor's terminals, the following relations are realized:

$$I_{DS} = \begin{cases} \frac{W}{L} q X D_n n_{po} exp(\frac{\psi_s}{V_T}) & : \quad V_{GS} < V_T \\ \mu \cdot C_{OX} \cdot \frac{W}{L} \cdot (2 \cdot (V_{GS} - V_T) \cdot V_{DS} - V_{DS}^2) & : \quad V_{GS} > V_T \ \& \ V_{DS} < (V_{GS} - V_T) \\ \frac{1}{2} \cdot \mu \cdot C_{OX} \cdot \frac{W}{L} \cdot (V_{GS} - V_T)^2 \cdot (1 + \lambda \cdot V_{DS}) & : \quad V_{GS} > V_T \ \& \ V_{DS} > (V_{GS} - V_T) \end{cases} \tag{2.1}$$

where V_T is the device threshold voltage (positive for a NMOS device, negative for a PMOS device), μ is carrier mobility (typically NMOS transistors have two times greater mobility than PMOS), C_{OX} is the oxide capacitance, W and L are the width and length, and λ is the channel length modulation

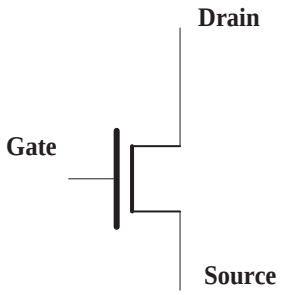


Figure 2.3: MOS Transistor Symbol

parameter or $\frac{1}{V_A}$ [7].

If the gate-to-source voltage is below the device threshold voltage, only subthreshold (or leakage) current flows from source-to-drain. When the gate-to-source voltage is approaching the value of V_T, the transistor begins to enter *inversion*. As the gate potential rises, it attracts the minority carriers to the surface of the substrate toward the gate voltage. Thus a n-type channel (in a p-type substrate) is created to allow current to flow from the two n-type implants at the source and drain. This state is known as weak inversion. If the gate-to-source voltage is greater than the threshold voltage, the channel becomes more inverted and enters the moderate inversion. As the gate-to-source voltage increases the channel becomes greatly inverted and the transistor is in strong inversion. The value of gate-to-source voltage required to operate a device in strong inversion is typically 200 milli-volts beyond the threshold voltage. Only in the moderate and strong inversion states does non-leakage currents flow [9].

The difference in the gate-to-source voltage and the threshold is defined as the overdrive voltage or

$$V_{OVERDRIVE} = V_{EFF} = \Delta V = V_{GS} - V_T \tag{2.2}$$

If the device is conducting and the drain-to-source voltage is less than the overdrive voltage, the

transistor is operating in the triode region. Likewise, if the drain-to-source voltage is greater than the overdrive voltage, the transistor is operating in the saturation region. The requirements for each operating region are shown below.

V_{EFF}	V_{DS} - V_{EFF}	Region
< 0	X	Subthreshold
> 0	< 0	Triode
> 0	> 0	Saturation

Table 2.1: Operating Regions

The resulting transfer characteristics are shown in Figure 2.4. The current is plotted versus the drain-to-source voltage for several gate-to-source voltages.

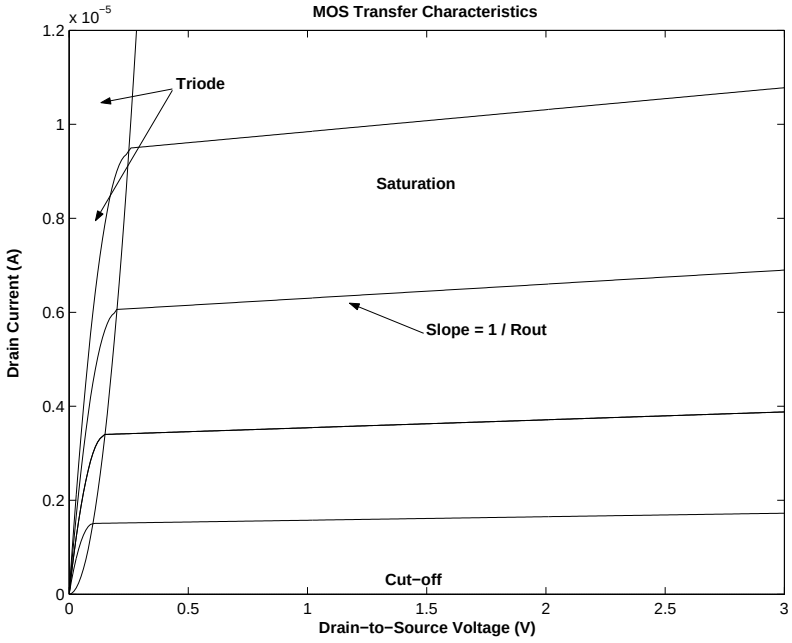


Figure 2.4: MOS Voltage-Current Relationships

2.4 Small-Signal Analysis

If the device is biased in the saturation region, voltage gain is possible. The gain of an amplifier can be expressed as a product of its transconductance, g_m, and output impedance, or

$$A = g_m \cdot r_{OUT} \tag{2.3}$$

Transconductance is defined as the change in output current with respect to a change in input voltage. For a MOS transistor, this results in

$$g_m = \frac{\partial I_{DS}}{\partial V_{GS}} = \frac{W}{L} \cdot \mu \cdot C_{OX} \cdot (V_{GS} - V_T) = \beta \cdot \Delta V \tag{2.4}$$

where $\beta = \frac{W}{L} \cdot \mu \cdot C_{OX}$

The output impedance of a device is defined as the change in output voltage with respect to change in output current, or

$$r_{out} = \frac{\partial V_{OUT}}{\partial I_{OUT}} \tag{2.5}$$

For a MOS transistor, this evolves into

$$r_{out} = \frac{\partial V_{DS}}{\partial I_{DS}} = \frac{V_A}{I_{DS}} = \frac{1}{\lambda \cdot I_{DS}} \tag{2.6}$$

where V_A is the Early voltage. Early voltage is a quantity which describes the output impedance of a transistor. Typically, the longer the channel length of transistor, the larger the Early voltage (this trend slows greatly as the channel length reaches very large values, typically 30 μ) [9].

Finally, the theoretical maximum voltage gain, or self-gain is defined as

$$\mu = g_m \cdot r_{out} = \frac{2 \cdot V_A}{\Delta V} \tag{2.7}$$

2.5 Temperature Effects

All of the MOS parameters are temperature dependent. As it will be shown, several of these parameters will lead to a degradation of performance at elevated temperatures. The two major temperature dependent terms are threshold voltage and mobility.

Threshold voltage is commonly known to decrease linearly with temperature. Current literature approximates the shift as

$$V_T = V_{T0} - \alpha \cdot (T - T_0) \tag{2.8}$$

where α is parameter specific to a fabrication process. Typically the value of α is between -0.5 and 1.5 mV/°C.

The current literature suggests mobility degrades with temperature in an power series effect. The trend is approximated as

$$\mu = \frac{K}{T^{\beta}} \tag{2.9}$$

with K and β again being dependent on the fabrication process. The values of K range from 20 to 200 $\mu A/V^2$. Both threshold and mobility shifts are modelled as shown in current SPICE decks [10].

2.6 Noise Characteristics

The two major noise sources of MOS are $\frac{1}{f}$ or flicker noise and thermal noise. Flicker noise is due to random trappings of carriers in the oxide channel, while thermal noise is due to the excitation of carriers in the channel. The resultant noise voltage is approximated by a Gaussian distribution with a mean of zero. While both of the noise voltages have a random value, their probability functions do fit known functions. The flicker noise is modelled as a spectral density with constant amplitude over all frequencies. The flicker noise is approximated by a density function that is inversely proportional to frequency, hence the term $\frac{1}{f}$ noise. Figure 2.5 shows the equivalent noise voltages referred to the

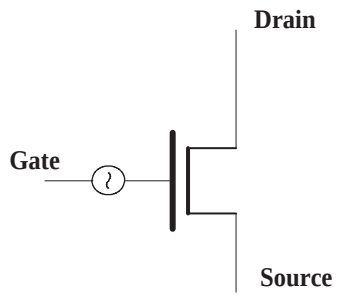


Figure 2.5: MOS Noise Voltage

input of a MOS transistor.

where the equivalent input voltage has a spectrum given by

$$V_n^2(f) = 4kT \cdot \frac{2}{3} \cdot \frac{1}{g_m} + \frac{K}{WLC_{OX} \cdot f} \tag{2.10}$$

Several design decisions can be made to keep the noise voltage low, including using long channel transistors and large transconductances. Figure 2.6 plots the frequency of the noise voltage. At lower frequencies, the noise is dominated by the $\frac{1}{f}$ noise, and at higher frequencies the noise is dominated by thermal contribution.

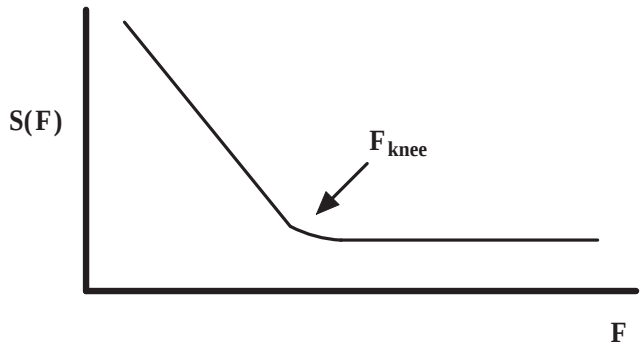


Figure 2.6: MOS Noise Spectrum

It should be noted that since noise voltage is a integral function of frequency, the higher the bandwidth of the circuit, the higher the input referred voltage [9].

2.7 Transistor Mismatch

Another source of error in MOS circuits is mismatch. Area mismatch occurs when the device is not properly etched, resulting in a geometry error occurs. The other major type of mismatch is threshold voltage. Threshold voltage is determined by the doping in gate. Variations in the doping result in shifts in the threshold voltage. As expected, these variations increase as transistor size decreases.

Although the area and threshold mismatch are random values, their probability does form a Gaussian distribution. The first order expected value of these mismatch can be found by the following distributions

$$V_{ERR(Area)} = \frac{A_{WL}}{\sqrt{W \cdot L}} \tag{2.11}$$

$$V_{ERR(V_T)} = \frac{A_{V_T}}{\sqrt{W \cdot L}} \tag{2.12}$$

Thus, the larger the area of the transistor, the smaller the error [11]. The quantity A_{VT} has a value of 3.5 to 4.2 mV / μm

2.8 Conclusion

As shown, designing analog circuits can be a challenging endeavor. Many tradeoffs are presented in determining the optimal solution. For example, steps in achieving high bandwidth in a circuit often decrease the gain and increase the power budget. These tradeoffs and compromises will ultimately play a large roll in designing the voltage reference.

Chapter 3

Band-Gap Reference Circuit

3.1 Introduction

This section details the design of the band-gap voltage reference. The design process included selection of a band-gap core and reference topology. Based on measurement results, the gated-diode was chosen as the band-gap device. The reference architecture chosen was the improved bipolar voltage reference. Two references were designed and simulated. The first reference consists of an amplifier biased in the saturation region, the other with an amplifier biased in the subthreshold region. The design of the reference with a saturation amplifier would result in simulation performance of 144 ppm/ °C temperature dependence and 40 ppm/V voltage dependence with a minimum supply of 1.8 V. The subthreshold design would simulate temperature dependence of 182 ppm/ °C, voltage supply dependence of 123 ppm / V, and operate on a 1.7 V supply.

3.2 Goals

Voltage references are a class of electronic circuits in which the circuit is designed to provide a constant output voltage. Voltage references are largely graded on three figures-of-merit: temperature behavior, voltage supply dependence, and power consumption.

- **Temperature Behavior -** A voltage reference would ideally provide a constant voltage regardless of temperature. Obviously, this is physically not feasible. The goal of the voltage reference is to supply the nominal output voltage over the largest temperature range possible. Temperature dependence is often quantified by a term called the temperature coefficient, or tempco. Tempco is determined by dividing the output voltage deviation by the allowable operating range. This quantity is often divided by the nominal voltage and multiplied by one million to express the value in parts-per-million (ppm).

$$TC = \frac{\Delta V_{NOM}}{\Delta T \cdot V_{NOM}} \cdot 10^6 \quad (\frac{ppm}{C}) \tag{3.1}$$

  Current literature report results of voltage references achieving temperature coefficients of 200 ppm [2].

- **Voltage Dependence -** The voltage dependence of a voltage reference is graded in two areas: minimal supply and supply effect on the output voltage. With supply voltages steadily decreasing, voltage references must be designed to operate accordingly. Fortunately, for this application, the band-gap reference is required to operate on a modest three volt supply. But, the reference should be made to function on a low as supply as possible to avoid errors from supply deviation and ease scaling efforts. Reducing the effect of supply variation is a major concern. If the reference is powered by a noisy or wavering supply, the output voltage can not deviate. This dependence of supply voltage is often expressed as Power-Supply-Rejection-Ration (PSSR). Current efforts have reported PSSR values of 200 ppm [2].

- **Power Consumption -** Obviously, the voltage reference would be preferred to operate with very little power consumption. Power consumption is becoming more critical as voltage supplies and area requirements shrink to meet high speed and portable application requirements requirements. Current voltage references draw less than 600 μW [2].

3.3 Architecture Selection

The vast majority of voltage references fall into two categories: CMOS and Bipolar. The advantages and disadvantages of each type are discussed in this section.

3.3.1 CMOS Voltage Reference

The block diagram of the CMOS voltage reference (often referred to as a β multiplier) is shown in Figure 3.1.

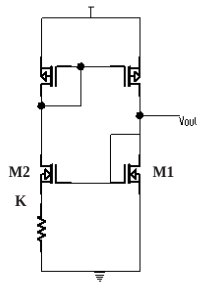


Figure 3.1: CMOS Voltage Reference

A current mirror is used to keep the currents in both legs equal. A voltage drop is developed from the scaling the width of transistor M2. Solving for the output voltage provides the following relation

$$V_{OUT} = V_T + V_{EFF1} = V_T + \frac{2}{R \cdot \beta_2}\left(1 - \frac{1}{K}\right) \tag{3.2}$$

In order to achieve temperature independence, the derivative of the output voltage with respect to temperature must be zero. Unfortunately, this circuit falls well short. If the classical temperature dependence of the transistor's mobility degradation is used (Equation 2.9), the derivative of the output voltage (ignoring the small effect of the resistor) is

$$\frac{\partial V_{OUT}}{\partial T} = \frac{\partial V_T}{\partial T} + \frac{2}{R \cdot \beta_2} \cdot \left(1 - \frac{1}{K}\right) \cdot \left(\frac{3}{2 \cdot T}\right) \tag{3.3}$$

In order to force this derivative to zero, the resistor and variable K must be set to match the temperature dependence of the threshold voltage. Since the threshold varies directly proportional to temperature, matching the derivative is only possible at one temperature. Although the derivative can not be set to zero at all temperatures, the temperature dependence can be set to a relatively small value. Published efforts report temperature dependence of 1000 ppm.

The voltage supply dependence of the CMOS voltage reference is also relatively praise-worthy. The minimum supply voltage required is the threshold voltage, the overdrive voltage, plus the supply needed for the current mirror. The power-supply-rejection-ratio can also be increased by using a current mirror with a large output impedance. A high performance current mirror will also be needed to ensure the analysis shown before is valid.

The power consumption of the CMOS voltage reference is also beneficial. Since there are only two legs to the circuit, the current drawn from the supply will be at a minimum. Obviously, the values of the resistor and scaling constant will have an impact on the power consumption. Although, to achieve a high performance current mirror, more circuitry may be required, resulting in higher power consumption.

There are several properties of the CMOS voltage reference that are not desirable. First, the output voltage can not be selected. The value of the output voltage is set in an effort to minimize the derivative. In fact, the output voltage will actually vary from production to production. The threshold voltage of transistors varies from devices, typically with a Gaussian distribution. Since the output voltage is dependent on the threshold voltage, the output voltage will vary with the error of the threshold voltage. Also, the temperature dependence is highly dependent on the temperature effects of the transistor. This translates to a critical dependence on the simulation model. Unfortunately, the temperature parameters are often modelled incorrectly. In fact, recent efforts indicate the temperature models used in SPICE simulations are inaccurate [13].

To summarize, the CMOS voltage reference is a relatively inexpensive design choice. The circuit consists entirely of resistors and transistors, operates at low voltage and power, and is very simple. Unfortunately, the simplicity of the circuit results in a sacrifice for performance. The temperature dependence is fair and the output voltage cannot be set. To achieve higher accuracy, other alternatives should be investigated.

3.3.2 Bipolar Voltage Reference

A circuit often employed in bipolar and BiCMOS (bipolar and CMOS) process is shown in Figure 3.2.

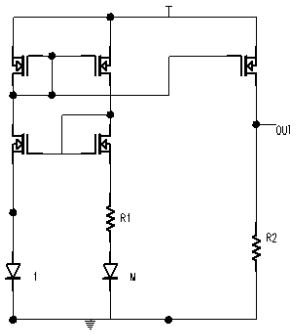


Figure 3.2: Bipolar Voltage Reference

By assuming the current mirror is ideal, the output voltage can be found as

$$V_{OUT} = V_{D1} + I \cdot R = V_{D1} + n \cdot v_T \cdot ln(M) \cdot \frac{R_2}{R_1} \tag{3.4}$$

where v_T is the thermal voltage of diode. The derivative of this equation with respect to temperature is

$$\frac{\partial V_{OUT}}{\partial T} = \frac{\partial V_{D1}}{\partial T} + \frac{R_2 \cdot n \cdot k * ln(M)}{R_1 \cdot q} \tag{3.5}$$

where k is Boltzmann's constant, q is the charge of an electron, and n is the emission coefficient or

ideality factor.

The quantity $\frac{\partial V_{BE}}{\partial T}$ is very nearly constant over a large range of temperatures. Also, the quantity is a negative value. Thus, by adjusting the value of M and the ration $\frac{R_2}{R_1}$, the derivative can be minimized resulting in small temperature dependence. This is a tremendous improvement from the previous voltage reference [12].

Like the CMOS voltage reference, the power supply rejection ratio depends on the current mirror. The higher the output impedance of the mirror, the larger the PSSR. The bipolar voltage reference can be expected to consume more power than the CMOS circuit due to the added third branch of current.

Unfortunately, the bipolar voltage reference is not ideal. Like the CMOS voltage reference, the output voltage cannot be selected by the designer; it is a result of trying to minimize the temperature derivative. And again the designer is dependent on the diode parameters (I_S, n) for an expected output voltage.

3.3.3 Improved Bipolar Voltage Reference

To alleviate the last problem, a slight variation of the bipolar voltage reference circuit is shown in Figure 3.3.

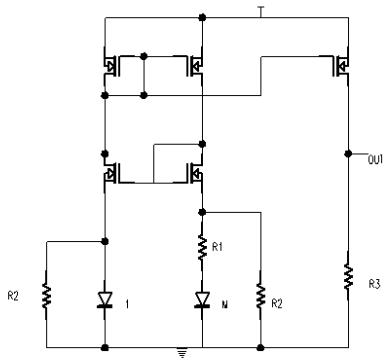


Figure 3.3: Improved Bipolar Voltage Reference

Again assuming the current mirror is ideal, the output voltage can now be found as

$$V_{OUT} = \frac{R_3}{R_2} \cdot \left[V_D + \left(\frac{R_2}{R_1} \cdot ln(M) \cdot v_T \right) \right] \tag{3.6}$$

By adding the two resistors R_3 and R_2, the designer is now equipped with ability to scale the output voltage to any desired value. This cost in area and power is far less than adding an amplifier to scale the output to a needed value. Also, since the output effects of the resistors is seen as a ration, the temperature behavior of the resistors is essentially cancelled [2]. Although the matching of the resistors is still a major concern.

Due to the large performance improvements of the bipolar circuit over the CMOS voltage reference, the bipolar voltage reference was chosen as the preferred solution. The small increase in power is a small sacrifice for the improved temperature behavior and the ability to adjust the output voltage. Finally, the bipolar voltage reference is often referred to as a band-gap voltage reference due to the presence of the diode voltage V_D. This voltage is dependent on the band-gap semiconductor material used (silicon, gallium-arsenide, etc.), which leads to the name band-gap voltage reference. Further analysis and information on diode implementation is provided in chapter **Band-Gap Core**.

3.4 Current Mirror Design

Both major categories of voltage references employ a current mirror to keep the currents in the two main branches equal. In fact, in solving for the output equations, the analysis above assumed the currents were perfectly equal. Unfortunately, this is an unrealistic goal. All current sources have some finite output impedance, which will cause deviation from the expected results. Obviously, by creating a current mirror with extremely high output impedance (> 10 MΩ), the effects of the finite output impedance will be greatly reduced. There are several methods to achieving a current mirror with high output impedance, these methods are discussed in the next sections.

3.4.1 Basic Current Mirror

A basic current mirror is shown in Figure 3.4.

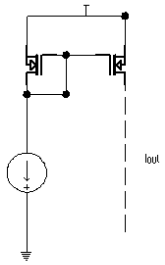


Figure 3.4: Basic Current Mirror

The circuit works as follows: A current source is provided in the left branch. The transistor in the left branch is *diode connected*, meaning a short is placed from the transistor's gate to drain. This keeps the gate-to-source and drain-to-source voltages at constant value for that current. Since the device is operating in the saturation region, the voltages are equal to (background information on CMOS devices is provided in Chapter 3)

$$V_{GS} = V_{DS} = \sqrt{\frac{2 \cdot I \cdot L}{W \cdot \mu \cdot C_{OX}}} + V_T \tag{3.7}$$

Since the two transistors share common gate and source connections, the two transistors will have the same gate-to-source potential. Using a more exact current model, the current in the second branch can be found as

$$I_{OUT} = \frac{W \cdot \mu \cdot C_{OX}}{L} \cdot (V_{GS} - V_T)^2 \cdot (1 + \lambda \cdot V_{DS}) \tag{3.8}$$

which equates to

$$I_{OUT} = I_{IN} \cdot (1 + \lambda \cdot V_{DS}) \tag{3.9}$$

Since λ is equal to the current divided by the output impedance, the ratio of the two currents becomes [9]

$$\frac{I_{OUT}}{I_{IN}} = (1 + \frac{I_{IN} \cdot V_{DS}}{r_{OUT}}) \tag{3.10}$$

Thus, there are three ways to reduce the error between the two branches. One, minimize the drain-to-source voltage of the second transistor. Unfortunately, by referring to the schematic of the band-gap voltage reference, it is clear that the drain-to-source voltage is equal to the supply voltage minus the diode voltage. The diode voltage is usually set at a value between two-tenths and eight-tenths of a volt. This voltage is setup to establish the correct temperature behavior of the circuit. Thus, minimizing the drain-to-source voltage is not a possibility.

Likewise, minimizing the current I_{IN} is not a valued solution. Since the diode's voltage is related to its current, by establishing a range of voltage to operate the diode, the current is fixed. This leaves the only possible way to reduce the error between the two devices is increasing the output impedance of the current mirror. Methods to accomplish this are discussed next.

3.5 Improving the Output Impedance

As discussed, a current mirror with large output impedance is needed to keep the currents in both diodes equal. If the currents are not equal, a large error will be seen at the output due to the diodes exponential current vs. voltage relationship. Thus, a current mirror with the highest output impedance and current balance should be implemented. The most common ways to achieve high output impedance are increasing the channel length of the mirroring devices, cascoding, and gain-enhancing.

3.5.1 Large Length Current Mirror

It has been shown that the output impedance of the basic current mirror is equal to

$$r_{OUT} = \frac{1}{I \cdot \lambda} \tag{3.11}$$

where the current is essentially fixed by the circuit. Fortunately, the designer of the current mirror has some control over the quantity λ. The value of λ is dependent on the channel length based on the following relation

$$\lambda = \frac{K}{L} \cdot \frac{\partial X_D}{\partial V_{DS}} \tag{3.12}$$

where X_D is the depletion depth.

Unfortunately, this relationship does have a limit. There exists a range in L where the designer sees diminished returns in increasing the channel length. As mentioned in the **CMOS Review**, by increasing the channel length, one slows down the device. In this case, the complex component of z_{OUT} begins to become a factor. This decreases the bandwidth of the circuit and increases the settling time of the circuit [9]. While bandwidth is not a critical requirement of the band-gap circuit (it is essentially a DC circuit), sufficient settling time is required in the event of a sudden increase in voltage. Also, increasing the channel length will also cost more area. Therefore, increasing the channel length indefinitely is not an useful solution. A realistic value of 10 μm was set for the current mirror.

Figure 3.5 and 3.6 show the temperature and power supply dependence of the band-gap circuit.

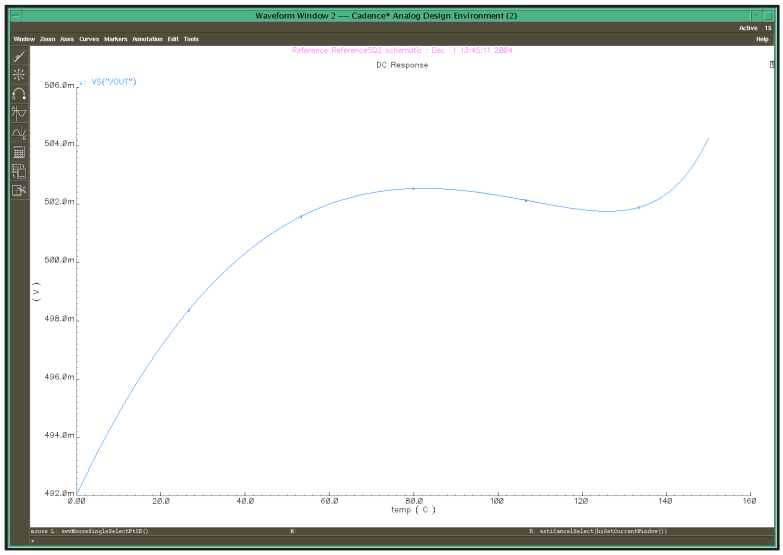


Figure 3.5: Basic Band-Gap Reference Temperature Dependence

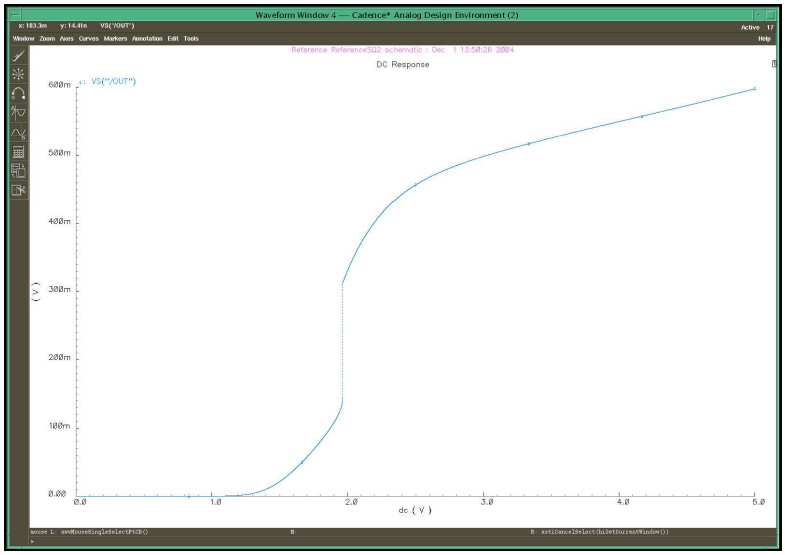


Figure 3.6: Basic Band-Gap Reference Power Supply Dependence

While these results offer a promising start, a more advanced solution should be considered. While the temperature dependence is respectable, the output voltage displays very high dependence on the supply voltage.

3.5.2 Cascode Current Mirror

Another method in increasing the output impedance is to use cascoding. Cascoding is employing a common gate amplifier to increase the output impedance of the current mirror by the self gain of the common gate stage [9]. This concept is shown in Figure 3.7.

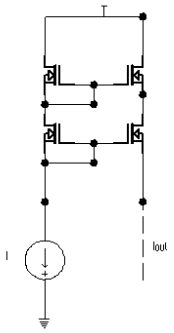


Figure 3.7: Cascode Current Mirror

While the output impedance is greatly increased, the gain is not achieved without cost. First, there will obviously be an increase in area. Since the increase is only two transistors, the increase is small. Also, the minimum supply voltage is increased. The supply voltage must now accommodate the diode voltage and two drain-to-source voltages. Fortunately, with a three volt supply, there is no harm in adding this stage.

The temperature and power supply dependence are shown in Figures 3.8 and 3.9 respectively. The result of the cascoding offers a large increase in performance for no increase in power.

The cascoding stage has decreased the tempco to and the PSSR to 250 and 560 ppm. Modest improvements from the basic current mirror. Although the minimum supply voltage is now larger to accommodate the extra transistor.

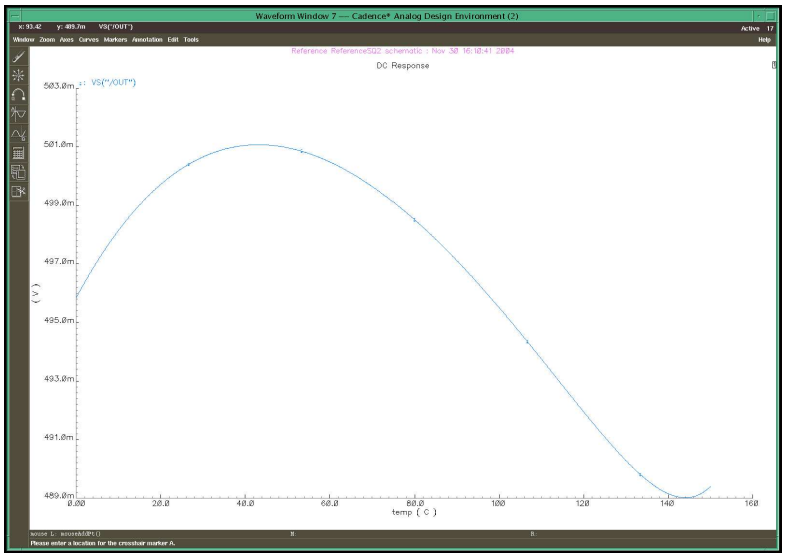


Figure 3.8: Cascode Band-Gap Reference Temperature Dependence

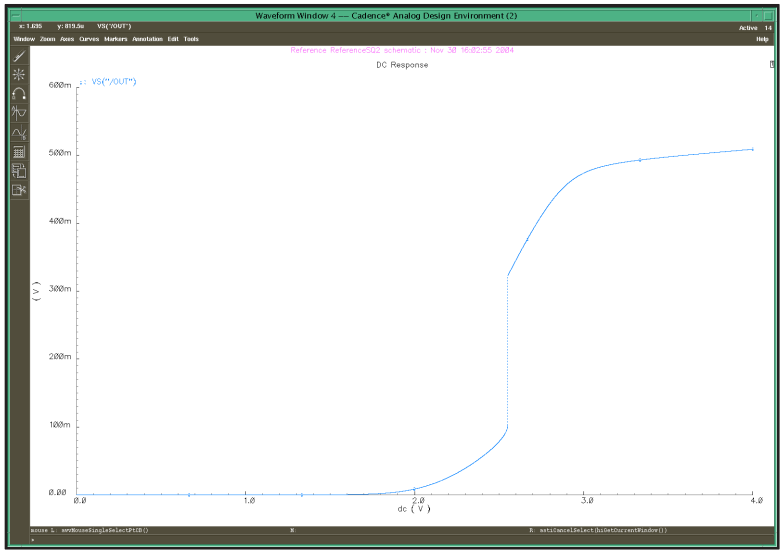


Figure 3.9: Cascode Band-Gap Reference Power Supply Dependence

3.5.3 Gain Enhanced Current Mirror

The last solution to increase the output impedance of the current mirror investigated is the gain enhanced of a gain boosted current mirror. Similar to the cascode current mirror, the gain enhanced mirror uses a full amplifier to achieve an even higher multiplication of the output impedance. A gain boosted cascode mirror is shown in Figure 3.10.

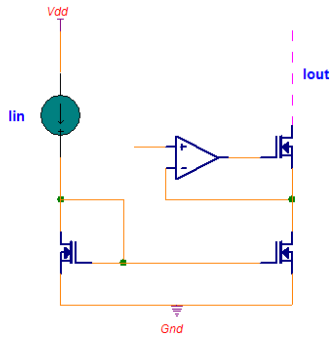


Figure 3.10: Gain Enhanced Cascode Current Mirror

By employing a full amplifier, the output impedance of the current mirror can be increased by a factor of 1 million. This is obviously greater than the increase on stage of cascoding can achieve (typically self-gain of one stage is less than 100) [12]. Again the designer sees the improvements have not come without cost. The added amplifier can be a very large area and power concern if not designed properly. Also, the addition of the amplifier greatly increases the complexity of the voltage reference, which could decrease its yield. These disadvantages will have to be considered when determining the preferred solution.

Figures 3.11 and 3.12 illustrate the temperature and supply dependence of the gain enhanced voltage reference.

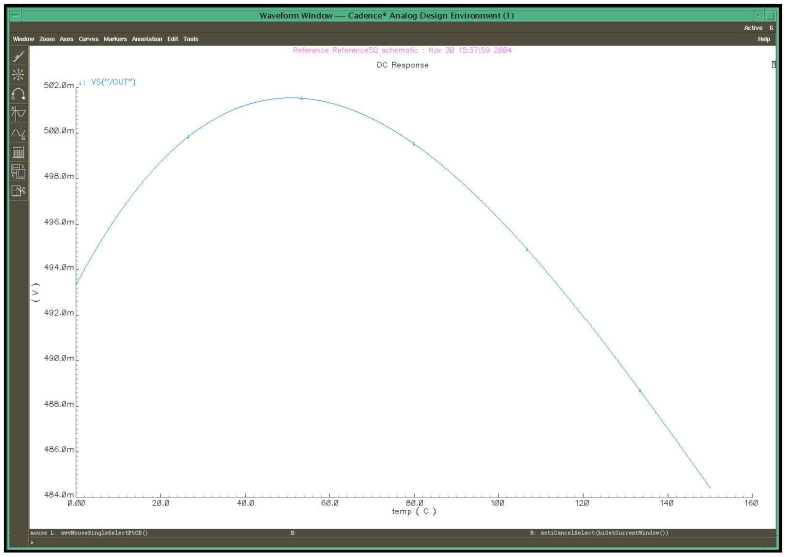


Figure 3.11: Gain Enhanced Band-Gap Reference Temperature Dependence

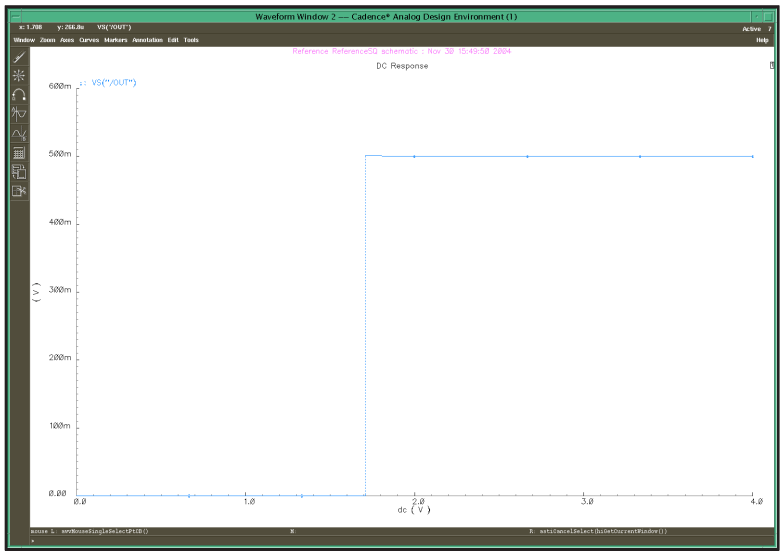


Figure 3.12: Gain Enhanced Band-Gap Reference Power Supply Dependence

3.6 Conclusion

Of the three voltage references designed and simulated, the band-gap reference employing an amplifier to increase the output resistance of the current mirror achieved the greatest results. Although the circuit is more complex and will require more power and area, the tradeoff in performance is acceptable for the cost in area and power. The following table demonstrates the temperature and voltage supply dependence of each topology. The best performance is achieved with the boosted band-gap voltage reference. Therefore, it was chosen as the circuit topology. The final topology is shown in Figure 3.13. Two versions of the band-gap reference were simulated and prepared for layout. The first consists of an amplifier biased in the saturation region, the other biased in subthreshold. More information on the amplifier design is provided in Chapter 5.

Topology	**TC**	**PSSR**	**Power Consumption**
Basic Voltage Reference	260 ppm	99000 ppm	56 (μW)
Cascode Voltage Reference	170 ppm	58000 ppm	56 (μW)
Amplifier (Saturation) Reference	144 ppm	40 ppm	945 (μW)
Amplifier (Subthreshold) Reference	182 ppm	123 ppm	180 (μW)

Table 3.1: Comparison of Band-Gap Topologies

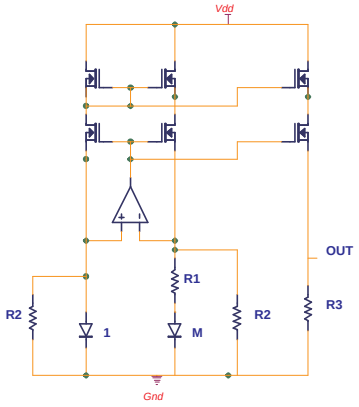


Figure 3.13: Gain Enhanced Band-Gap Reference

Chapter 4

Band-Gap Core: Background, Measurement, and Model Extraction

4.1 Background

This section discusses the selection process for band-gap core for the voltage reference circuit. The band-gap core is used to establish a current that is predictably proportional to temperature. The two choices available in CMOS technology are a parasitic diode and a gated-diode structure. Both devices were measured to determine the optimal solution. Due to area and biasing constraints, the gated-diode was chosen as the band-gap core.

4.1.1 Bipolar Technology

Bipolar technology is currently the workhorse for temperature insensitive voltage and current references. One can understand this by examining the temperature dependence of the diode. The current equation of the diode is commonly known to be

$$I_D = I_S \cdot (e^{\frac{q \cdot V_D}{n \cdot k \cdot T}} - 1) = I_S \cdot (e^{\frac{V_D}{n \cdot V_T}} - 1) \tag{4.1}$$

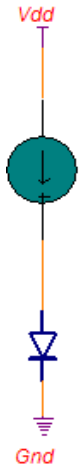


Figure 4.1: Forward Biased Diode

where I_S is the reverse-bias or saturation current, V_D is the forward-bias diode voltage, k is Boltzmann's constant, q is the charge of an electron, T is temperature in Kelvin, and n is the ideality factor or emission coefficient. The product

$$v_T = \frac{k \cdot T}{q} \tag{4.2}$$

is referred to as the thermal voltage of the diode junction.

The thermal contributions are found in the quantities I_S and v_T. As shown from equation 4.2, the thermal voltage is directly proportional to temperature. The temperature dependence of the saturation current, I_S, is far more complicated.

Despite the diode's exponential relationship with temperature, diodes can be used to develop voltages and currents approximately linear to temperature. An example of this is the circuit shown in Figure 4.1. The diode is forward-biased by a constant current source.

Simple analysis of this circuit shows the diode voltage is approximately linear with temperature. The diode voltage is equal to

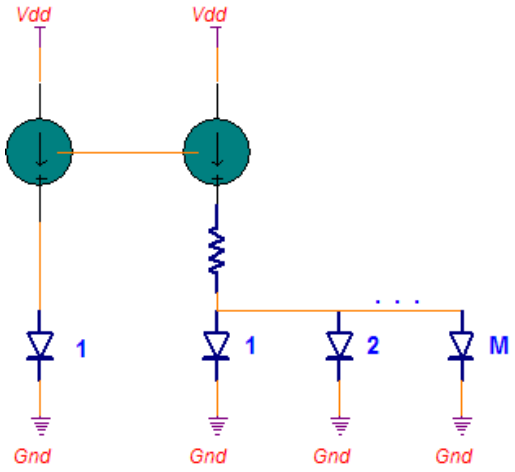


Figure 4.2: PTAT Circuit

$$V_D = n \cdot v_T \cdot ln(\frac{I}{I_S}) \tag{4.3}$$

The diode's voltage temperature dependence can be found by taken the derivative with respect to temperature.

$$\frac{\partial V_D}{\partial T} = n \cdot \frac{k \cdot T}{q} \cdot \frac{\partial ln(\frac{I}{I_S})}{\partial T} + n \cdot \frac{k}{q} \cdot ln(\frac{I}{I_S}) \cong -\frac{k \cdot a}{q} \cdot ln(\frac{I}{I_{S0}}) \tag{4.4}$$

Thus, the diode voltage is approximately linear with temperature with a negative slope.

Consider the circuit shown in Figure 4.2. The diode current is mirrored by the ideal current mirror. The integer M is the number of diodes placed in parallel. Simple analysis of this circuit produces the following relationship:

$$I = \frac{v_T \cdot ln(M)}{R} \tag{4.5}$$

By choosing a resistor with small temperature dependence, a current is produced that is proportional-to-absolute-temperature or PTAT.

Figure 4.3: BJT Diode

Finally, it should be noted that in processes where pn junction diodes are not readily available, such as CMOS, a diode can be constructed from a BJT as shown in Figure 4.3. The base-collector diode is shorted, leaving only the base-emitter diode. Ignoring the small base current, the current reduces to Equation 4.1. This configuration is commonly known as a Widlar diode [12].

4.1.2 CMOS Technology

To fully understand why bipolar technology is so popular for voltage references, one should analyze the temperature dependence of the other dominant process, CMOS. To do this, consider the circuit shown in Figure 4.4. This circuit is the similar to the circuit shown in Figure 4.1; the only difference is the diode is replaced with a CMOS transistor. This configuration is commonly known as a MOS diode.

Using simple circuit analysis, the diode voltage can be found as

$$V_D = \sqrt{\frac{2 \cdot I \cdot L}{W \cdot \mu \cdot C_{OX}}} + V_T \tag{4.6}$$

where μ is mobility, C_{OX} is the oxide capacitance, V_T is the threshold voltage, and W and L are the width and length, respectively.

The temperature dependent terms in Equation 4.6 are the mobility and threshold voltage. By substituting classical models for these terms, Equation 4.6 evolves to

$$V_D = \sqrt{\frac{2 \cdot I \cdot L}{W \cdot C_{OX}}} \cdot \sqrt{\frac{T^{1.5}}{K}} + V_{T0} \cdot (1 + \alpha(T - T_0)) \tag{4.7}$$

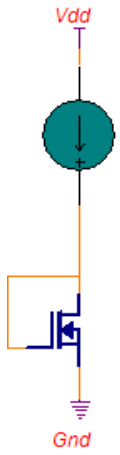


Figure 4.4: MOS Diode

where K and μ are parameters specific to the fabrication process. Again taking the derivative of the diode voltage with respect to temperature yields

$$\frac{\partial V_D}{\partial T} = -\sqrt{\frac{2 \cdot I \cdot L}{W \cdot C_{OX}}} \cdot \frac{T^{\frac{3}{4}}}{2 \cdot K^{\frac{3}{2}}} + \alpha \cdot V_{T0} \tag{4.8}$$

Analysis of Equations 4.7 and 4.8 illustrate several problems using this type of circuit for a band-gap core. First, the diode voltage in Equation 4.7 is very dependent on fabrication parameters. For example, the threshold voltage varies linearly with temperature. This change alone would produce a shift in voltage that is undesirable in many high precision applications. Similar challenges arise from Equation 4.8. Instead of a nearly constant derivative with temperature as with the bipolar devices, the derivative of the CMOS device changes with temperature dramatically. These downfalls are the major reasons bipolar technology is used nearly exclusively in discrete high-precision voltage references [12].

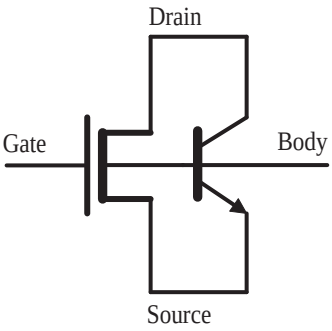


Figure 4.5: Parasitic BJT

4.2 Diode Implementation

After selecting the diode as the preferred core for the voltage reference, correct implementation of the diode is required. As mentioned, in a purely CMOS technology a BJT or diode is not readily available. Few options are available, and these options are discussed in this section. These choices are a parasitic BJT and a gated-diode structure.

4.2.1 Parasitic Bipolar-Junction-Transistor (BJT)

A common choice for implementing a diode in CMOS processes is to take advantage of the parasitic BJT in parallel with a CMOS transistor. The BJT is formed by the npn arrangement of the drain, body, source. The drain effectively becomes the collector, source becomes emitter, and the body becomes the base. The gate must be biased to avoid operating the transistor in the linear or saturation regions. This is usually accomplished by biasing the gate voltage at the one of the power supplies (negative supply for NMOS, positive for PMOS). If the threshold voltage of the transistor is sufficiently large ($|V_T| > .5V$), the gate can be tied to the source. The equivalent circuit for the parasitic BJT is shown in Figure 4.5. A diode can be formed by simply shorting the body to the source of the transistor [7].

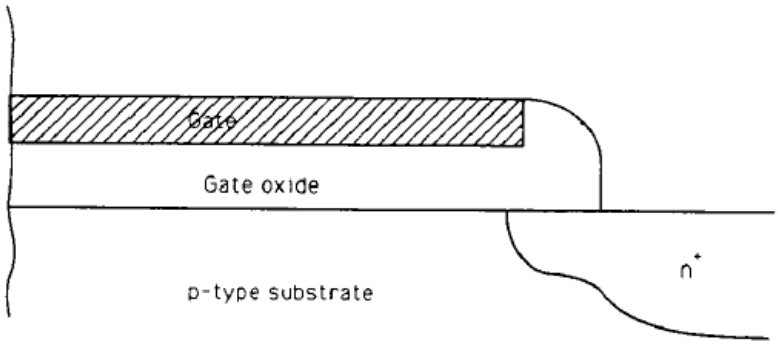


Figure 4.6: Gated-diode

4.2.2 Gated-Diode

Another alternative for a diode solution is the gated-diode. The gated-diode structure is a popular choice for electrostatic discharge (ESD) protection circuits. The gated-diode is formed by placing a p implant in the source or drain of a NMOS transistor, or a placing a n implant in the source or drain of a PMOS transistor. Thus a pn junction is formed through the source, substrate, and drain. The gate voltage must be biased to keep the transistor from operating in the conductive region. A fabrication level drawing of the gated-diode is shown in Figure 4.6 [14].

4.3 Diode Measurement and Selection

The two options where then measured to determine which would be used as the band-gap device. The devices would be selected on matching, temperature performance, and bias compatibility. Twenty of each device was measured to determine the preferred solution. The gated-diode emerged as the only feasible solution.

4.3.1 Measurement

A common method to determine the voltage-current characteristics are to simply vary the voltage across the diode and measure the current. This is done easily with the HP4156 Parameter Analyzer.

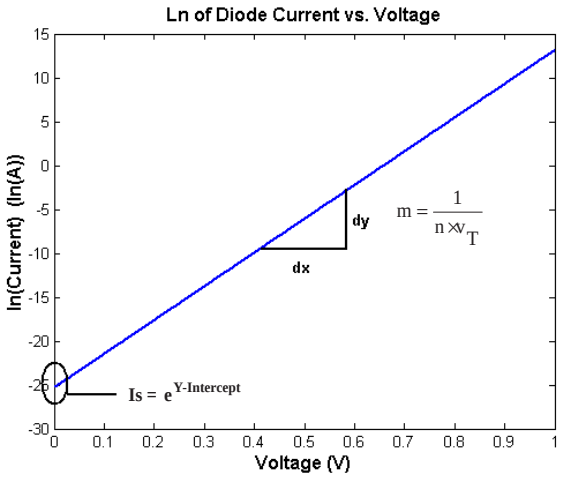
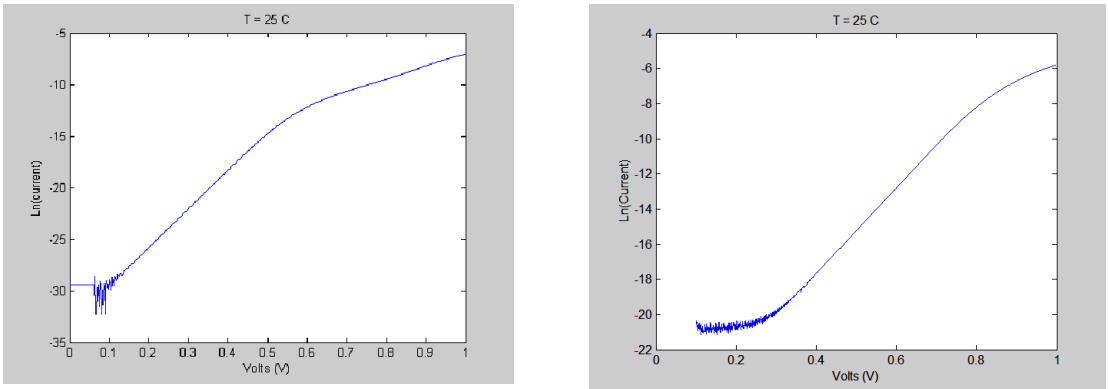


Figure 4.7: Ideal Diode

Parameters *n* and *Is* can determining by examining the natural logarithm of Equation 4.1

$$ln(I_D) = \frac{V_D}{n \cdot v_T} + ln(I_S) \tag{4.9}$$

Figure 4.7 shows an ideal plot natural logarithm of diode current vs. diode voltage (T = 30° C). where the slope is equal to $\frac{1}{n \cdot v_T}$ and the y-intercept is equal to $ln(I_S)$. Thus, if the temperature is known, the ideality factor, *n*, and saturation current, *Is*, can be found. Figure 4.8 shows the natural logarithm of the diode current for the parasitic BJT and gated-diode, respectively.



Parasitic Diode **Gated Diode**

Figure 4.8: Natural Logarithms of Diode Currents

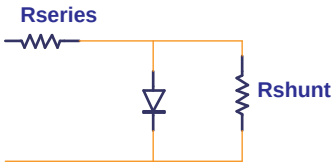


Figure 4.9: Diode Parasitic Resistances

Interestingly, both diodes deviate from the expected trend. At the lower and upper regions of the curve, the curve flattens with a slope of approximately zero. Only in center region do the diodes perform as desired. This behavior can be explained by the parasitic resistances present from the substrate and contacts. These distributed resistances can be approximated by a shunt and series resistor as shown in Figure 4.9 [7].

At low voltages where the actual diode current (Equation 4.1) is essentially zero, but the shunt resistor contributes a current. Conversely, at higher voltages, the series resistor becomes a considerable voltage drop, reducing the exponential relationship. Only in the middle region is the current dominated by the diode. Thus, the total current of the diode circuit can be found by

$$I_D = \frac{V}{R_{S\,hunt}} + I_S \cdot e^{\frac{V-V_D}{n \cdot v_T}} \tag{4.10}$$

where V_D is given by Equation 4.3. Analysis of this transidental equation requires iteration, therefore is use not straight-forward.

4.3.2 Selection

As mentioned above, the diode chosen as the band-gap device was selected on three criterion: bias feasibility, temperature behavior, and matching.

- **Bias Compatibility -** As shown in Equation 4.5, the current produced in the diode PTAT circuit in Figure 4.2 is proportional to the thermal voltage and log of the number of diodes in parallel. Unfortunately, the current is not a free parameter. The current must be biased in the log-linear range as shown in Figure 4.8. Since the thermal voltage is set by temperature, the only free parameters are M and R. The value of M would ideally be set as large as possible to increase matching, but will ultimately be limited by area restraints. Obviously, the larger the log-linear range in Figure 4.8, the larger the value of M can be. Based on this, the gated-diode is a better choice. The log-linear range of the parasitic diode is approximately 280 milli-Volts, while the gated-diode is only 250 milli-Volts (log-linear range was defined by the portion of the data with a 99.9 percent correlation with a straight line).

  The value of the resistor is also set by the log-linear properties of the diode. The value of the resistor can be found by choosing M based on the above statement, and solving for the resistor value that provides the current in the middle of the log-linear region of Figure 4.8. A problem arises when completing this process with the parasitic diode: the value of the resistor becomes very large. Even in the unrealistic case where $ln(M) = 1$, R is over 20 $M\Omega$! Unfortunately, this size of a resistor is not pleasable on integrated circuits. The area consumed would simply be too large. For the gated-diode, a resistor of one to 20 $k\Omega$ is required (with a M of two to ten). Since the gated-diode is operating a higher current, M must be kept small to prevent a large voltage drop across the resistor. A conservative value of four was selected. Although the parasitic diode operates at a lower power (a goal of voltage references) and has

a larger log-linear range, the device can not be biased at a desirable operating point due to its large parasitic resistance. The gated-diode was the only option despite its DC shortcomings compared to the parasitic diode.

It should be noted that the problem discussed above for the parasitic diode can be solved by adding multiple copies of the parasitic diode in parallel to effectively increase the saturation. Although, to make up the difference between the actual saturation current and saturation current needed for successful bias would require scaling the parasitic diode by more than 100 times! This cost in area is obviously unacceptable.

- **Temperature Performance -** With the first criterion investigated, the next property to be examined is the temperature performance. As mentioned, the diode's current relation consists of two temperature dependent terms: saturation current I_S and thermal voltage v_T. Thus, the saturation current and ideality factor can be easily be determined by measuring the diode current versus voltage a several temperatures. This process was completed for the gated-diode for temperatures 25 to 185 °C at ten degree intervals. To ensure the devices reached thermal equilibrium, the wafer was allowed to soak three minutes before measurement. The extraction of the temperature dependence of the saturation current produced a very interesting results. The temperature dependence did not fit the expected trend. A far more accurate model is the following

$$I_S(T) = I_S(T_0) \cdot e^{(\alpha \cdot T^2 + \beta \cdot T)} \tag{4.11}$$

where α and β are constants, $I_S(T_0)$ is the saturation current at the nominal temperature (T = 30 °C was selected).

An even more interesting phenomenon was observed when extracting the ideality factor n. By taking the natural logarithm of the diode current versus voltage, the slope now becomes $\frac{1}{n \cdot v_T}$. The ideality factor can be found through simple division. The ideality factor should

be a temperature independent quantity. In fact, SPICE models do not have a temperature dependent term for the ideality factor. But contrary to expectations, the ideality factor for the gated-diode exhibited the temperature dependence shown in Figure 4.10.

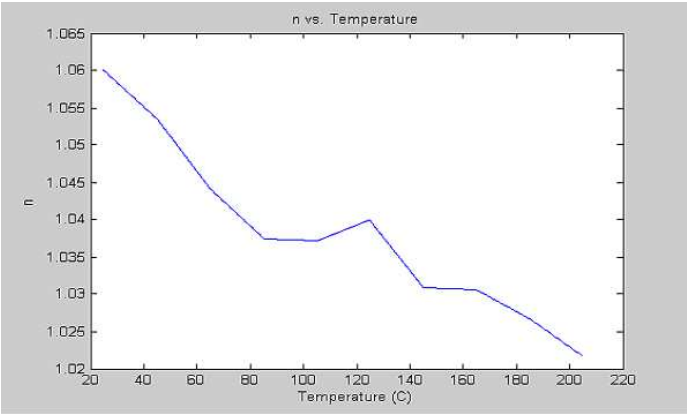
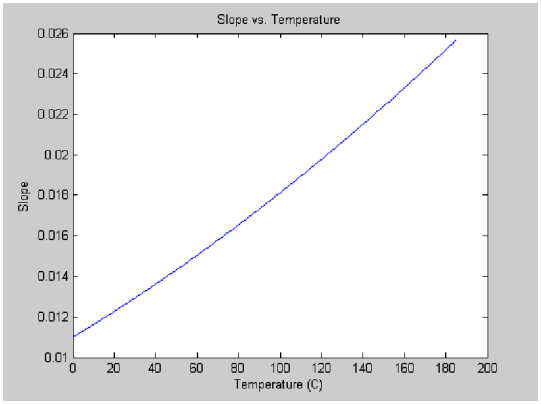


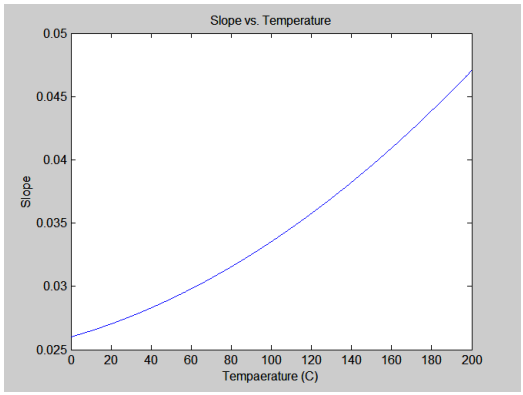
Figure 4.10: Diode Ideality Factor

Not only does the ideality factor seem to vary with temperature, it produces a nonlinear curve that is very sporadic. The same testing was completed on the parasitic diode to investigate this finding. A very similar curve was found for the parasitic diode: a nonlinear relation with no simple dependence with temperature. This barrier is a serious problem in the design of the voltage reference. If the diode's temperature behavior cannot be modelled accurately, the voltage reference's performance will degrade.

As shown before, by taking the natural logarithm of the diode's current, the exponential term can be determined as the slope of the log-linear portion of Figure 4.8. Current literature suggests this term should simply be $\frac{1}{n \cdot v_T}$, or its inverse is $n \cdot v_T$ Thus, plotting the exponential term's inverse versus temperature should result in a straight line with constant slope of $\frac{n \cdot k}{q}$ and y-intercept of zero where k is Boltzman's constant and q is the charge of an electron. Interestingly, the exponential term did not fit a straight line or have a zero y-intercept. The plots of the inverse of the exponential terms of the gated-diode and parasitic diode are shown in Figure 4.11.



Parasitic Diode



Gated Diode

Figure 4.11: Inverse of Exponential Terms

Simple curve fitting shows these curves have a high correlation with a second order polynomial. Thus, the curves shown can be fit to a model of

$$\alpha \cdot T^2 + \beta \cdot T + C \tag{4.12}$$

This yields the actual exponential contribution as

$$\frac{1}{\frac{k}{q} \cdot n_0 \cdot T + (\frac{k}{q})^2 \cdot n_1 \cdot T^2 + n_2} \tag{4.13}$$

$n_1 \cdot T^2 + +n_2$ are the unexpected contributions. n_0 is unitless, n_1 has units of $\frac{1}{T}$, and n_2 has units of T.

Table 4.1 shows the values of the coefficients of Equation 4.13 for both the parasitic diode and the gated-diode.

Device	n_0	$n_1(\frac{1}{T})$	n_2 **(T)**
Parasitic Diode	3.87304E-6	6.1305E-6	5.741213E-1
Gated Diode	5.3199239E-6	9.481423E-8	1.97479233513E-2

Table 4.1: Coefficients of Exponential Term

Again the parasitic diode's performance is preferred. The unexpected coefficients (n_1 and n_2 for the parasitic diode are far less than those of the gated-diode. But due to the incompatibility with the chosen band-gap architecture, the gated-diode must be chosen. The cause of the higher temperature dependence is not currently known.

To summarize, both band-gap options showed unexpected higher order temperature dependence than expected. These contributions can be taken into account using the following equation

$$I_D = I_{S0} \cdot e^{\alpha \cdot T^2 + \beta \cdot T} \cdot e^{\frac{cdotV_D}{\frac{k}{q} \cdot n_0 \cdot T + (\frac{k}{q})^2 \cdot n_1 \cdot T^2 + n_2}} \tag{4.14}$$

Although the parasitic diode's temperature was more as expected, the bias-area compatibility does not allow use of the parasitic diode. Since the temperature dependence of the diode is not modelled in SPICE decks, a Verilog-A module was created for simulation purposes. The Verilog-A code is provided in Appendix A.

- **Matching -**Device matching is a very critical parameter in the design of the band-gap voltage reference. Matching is especially important for the diode. Since the diode current-voltage relation is exponential, the smallest deviation results in large swings in currents. Since the log-linear region shown in Figure 4.8 is relatively small, the mismatch could push the operating point beyond the acceptable region.

  Mismatch is a result of several non-ideal properties of the device process. Mismatch is a largely random phenomenon. Studies have shown that the mismatch of devices is highly Gaussian, and can be minimized by including multiple copies, or fingers of devices. Since the transistors that make-up the gated-diode are not being biased in the triode or saturation regions, the threshold mismatch is of no concern. The largest source of error is the area mismatch [11].

  Direct comparison of the matching properties of the two devices is not straightforward. The

dominant matching terms for each device are different. The dominant term for matching for the parasitic diode is doping in the channel, while the matching for the gated-diode is area mismatch. While the sources of mismatch are different, a general guideline is that matching increases with area and channel length. This trend benefits the gated-diode. The gated-diode consists of four transistors of width of 50 μm and length 5 μm for a total area of 250 μm^2. While the parasitic diode is formed with a transistor of width 10 μm and length 1 μm for a total area of 10 μm^2.

The measurements proved the gated-diode matching was higher. All devices measured fit the relation provided in Equation 4.12 and Table 4.1 to within half a percent over all temperatures measured. The error for the parasitic diode was beyond a percent. This ensures an accurate model was constructed and simulation data can be used with confidence.

4.4 Conclusion

Two band-gap devices, the parasitic bipolar-junction-transistor and gated-diode, were proposed and measured for possible solutions for the voltage reference. Due to bias constraints, the gated-diode was chosen as the preferred solution. The gated-diode DC and temperature behavior does not fit the established SPICE diode model; therefore, a Verilog-A model was created for simulation purposes.

Chapter 5

Amplifier Design

5.1 Introduction

This chapter details the design process of the amplifier used to improve the performance of thr band-gap voltage reference. As explained in the **Band-Gap Voltage Reference** chapter, the amplifier is used to keep the currents in reference equal. The amplifier was designed to achieve high gain using low power and a low input offset voltage. The amplifier also must have high performance at elevated temperature. These design criterion would lead to selection of a boosted cascode amplifier.

5.2 Motivation

To illustrate the importance of the amplifier in the reference, refer to Figure 5.1. The amplifier is used to keep the diode voltages approximately equal. The output of an amplifier is equal to the voltage difference of its input terminals multiplied by the gain or

$$V_{out} = (V_+ - V_-) \cdot A \tag{5.1}$$

If the output voltage is relatively small and the gain is large, the voltage difference between the two input terminals becomes very small. In the case of the band-gap circuit of Figure 5.1, this

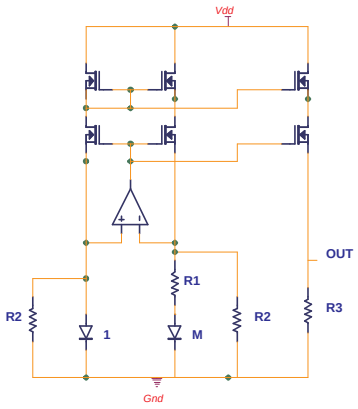


Figure 5.1: Band-Gap Voltage Reference

translates to keeping the diodes at approximately the same voltage. This is extremely important due to the exponential relationship of diode current with voltage [5]. A small difference in voltage will produce a large swing in current. From this point on the difference between the input terminals of the amplifier will be referred to as voltage offset, or V_{OS}.

The amplifier is also used to increase the performance of the current mirror. Since the drain voltages of both transistors are held at approximately the same voltage, the currents in the current mirror will track perfectly. Only the mismatch in the two transistors will cause a deviation in current. The use of an amplifier also increases the output impedance of the current mirror. Small signal analysis can be used to show that the output impedance increases by a factor of $(1 + A)$. An high output impedance current mirror is important in the event the supply voltage changes. In small signal terms, the change of drain current is equal to the change of drain-to-source voltage divided by the output impedance or in the case of the band-gap circuit

$$\Delta I = \frac{\Delta V_{DS}}{r_{OUT}} = \frac{\Delta V_{SUPPLY}}{r_{OUT}} \tag{5.2}$$

Thus, by employing a current mirror with high output impedance, the currents in the voltage reference can become nearly insensitive to supply voltage [2].

5.3 Design Criteria

Many design variables must be considered in designing an amplifier. For this application, several considerations emerged as most important. These include

- **Gain -** Gain of the amplifier is the most critical characteristic for the band-gap circuit. As mentioned, the amplifier's gain increases the balance of the currents in the voltage reference. The amplifier will also have to be designed to achieve high gain at elevated temperatures.

- **Offset -** Amplifier offset is the major contributor to difference in diode voltage. Amplifier offset also creates an error at the output of the band-gap circuit. There are three main contributors to amplifier offset: systematic offset, noise, and device mismatch. Systematic offset refers to the input referred offset due to the architecture of the amplifier. The noise offset is also an input referred voltage due to the noise voltage of the amplifier. Noise voltage has two major sources: thermal and flicker noise. The mismatch offset is a result of the inaccuracy of the transistors' expected sizes and actual sizes in the amplifier.

- **Power Consumption -** One figure-of-merit of a voltage reference is its power consumption. Obviously keeping the power down in the amplifier is a step in the right direction. The power consumed by the amplifier is determined by the topology of the design and transistor size.

- **Frequency Response -** Every amplifier must have sufficient bandwidth for the application. Although the band-gap circuit is a DC application, some bandwidth for settling proper settling time. Frequency response is a product of architecture and transistor size.

Unfortunately, these design criteria often oppose one another. For example, a low power amplifier often has poor bandwidth. But a lower bandwidth amplifier often has lower noise voltage and can achieve high gain. As with most amplifier design, a suitable compromise is required to optimize as many requirements as possible.

5.4 Amplifier Design

This section outlines the steps taken in selecting an amplifier for the band-gap reference. The amplifier architecture was selected based on the criteria listed in Section 5.3. These criteria would lead to selection and design of a boosted-cascode amplifier.

By analyzing the requirements listed in Section 5.3, several results can be concluded. First, the amplifier topology should be relatively uncomplicated. While many more advanced topologies are aimed at achieving elevated bandwidth, a large bandwidth is not required for the band-gap voltage reference. In fact, larger bandwidth than is required leads to higher input referred noise. Also, since the band-gap reference is a precision circuit, a reliable, high yield design should be used. This again points to a more elementary as opposed to an advanced architecture. Last, the amplifier should achieve the largest voltage gain possible. Unfortunately, most basic amplifier designs do not achieve large voltage gains. This leads to the goal of achieving large gain with a low power, reliable design.

Several architectures of amplifiers were considered before selection. These amplifiers include the two-stage amplifier, cascode amplifier, and boosted cascode amplifier. The benefits and drawbacks of each are listed here.

5.4.1 Bias Generator

Before discussing the design of the amplifier, an understanding of the bias generator is needed. As shown in the **CMOS Review**, to achieve voltage gain, a transistor must be biased in the saturation region. To do this, a bias generator must be constructed. A constant transconductance bias generator is shown in Figure 5.2. A voltage drop is developed across the bias resistor due to the imbalance in transistor sizes, and the current is mirrored by transistors M7 and M8.

Solving for the transconductance of device M1 yields the following equation

$$g_{m1} = \frac{2}{R_B}\left(\sqrt{\frac{(W/L)_1}{(W/L)_2}} - 1\right) = \frac{K}{R_B} \tag{5.3}$$

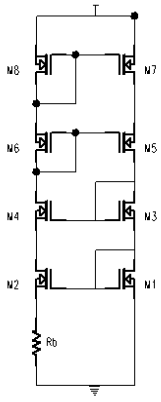


Figure 5.2: Bias Generator

As evident from the equation, the transconductance is set simply by the device geometry and bias resistor. Also, the transconductances can be set simply mirroring the current of bias generator to another leg of a circuit. The resultant transconductance is [9]

$$g_{mi} = g_{m1} \cdot \left(\sqrt{\frac{\mu_i \cdot I_i \cdot (W/L)_i}{\mu_1 \cdot I_B \cdot (W/L)_1}} \right) \tag{5.4}$$

An interesting property of this bias generator is the thermal properties. As evident from Equation 5.14, the only temperature dependent term in the transconductance is the bias resistor. Resistors typically have a positive linear temperature, resulting in an increase in impedance as temperature increases. The increase in resistance is approximated by [10]

$$R(T) = R_{T0} \cdot (1 + \alpha \cdot (T - T_0)) \tag{5.5}$$

where α is a parameter specific to the type of material used to construct the resistor. Thus, g_m will actually decrease as the temperature increases. But, if the temperature coefficient of the resistor is quite small, than the effect is negligible. This gives rise to the name constant transconductance circuit [9].

Unfortunately, this is not a perfect solution. As presented earlier, transconductance is a function

of parameters that are temperature dependent. Equation 5.6 can be rearranged to form the following

$$g_m = \sqrt{2 \cdot \mu \cdot C_{OX} \cdot (W/L) \cdot I_D} \tag{5.6}$$

The only parameters in this equation that can deviate with temperature are the mobility and current. Equation 5.11 illustrates that mobility decreases with temperature with temperature. To keep the transconductance constant, the current must perform the inverse of the mobility's shift. For example, if mobility degrades 50 percent, the current must double to compensate. This increase in current will obviously increase power, and from Equations 5.8 and 5.9, cause a reduction in self gain. In fact, the self gain will reduce at the same rate as mobility [9].

A circuit that attempts to keep self gain constant is shown in Figure 5.3.

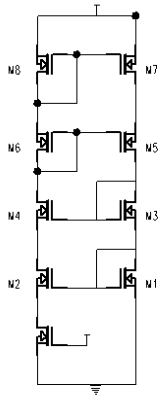


Figure 5.3: Constant Gain Bias Generator

The only difference between this circuit and the constant transconductance bias generator is resistor R_B is replaced by a transistor. The transistor M_R is operating in the triode region. By operating in the triode region, the transistor behaves as a voltage controlled resistor. The resistance of the transistor is simply

$$R_M = \frac{1}{\beta \cdot V_{EFF}} \tag{5.7}$$

Unfortunately, this type of bias generator is not feasible choice. The band-gap circuit must be designed to perform over a wide range of supply voltages. If the supply voltage is allowed to vary, the bias arrangement will shift. This will result in large swings in power consumption, gain, frequency response, and many others. This is not a acceptable circumstance, resulting in elimination of the constant gain bias generator shown in Figure 5.3.

Finally, the roles of transistors M5 and M6 need explanation. These transistors are used to cascode, or multiply the output impedance of the current mirror formed by M7 and M8. The output impedance of the current mirror is multiplied by the self gain of transistors M6 and M6. The output impedance boost greatly increases the power-supply-rejection-ration, making the bias generator less sensitive to supply variations. M3 and M4 serve a similar purpose [9]. Also, the increase output impedances results in a stronger match between the currents in both legs of the circuit. If the currents in the legs to setup equal, deviations from the transconductance expected will result.

As with all improvements in electronic design, the increase in performance is not without cost. The addition of the transistors increases the minimal supply voltage. The supply voltage must now accommodate the drain-to-source voltages of transistors M3 and M5, or and M4 and M6, which ever is the largest.

5.4.2 Subthreshold Bias Generator

Another bias generator configuration can be achieved by simply keeping the scaling term K very small. This keeps the transistors in the weak inversion region. By further increasing the resistor, the transistors can be biased into subthreshold. Subthreshold operation offers very large voltage gains at low power [12]. Unfortunately, due to the exponential relationship of the current, any mismatch in the transistors will have a very large impact [11].

5.4.3 Two-Stage Amplifier

The first amplifier analyzed was a two-stage amplifier. The two-stage amplifier is one of the most common OTAs (Operational Transconductance Amplifier). In fact, the two-stage amplifier is the

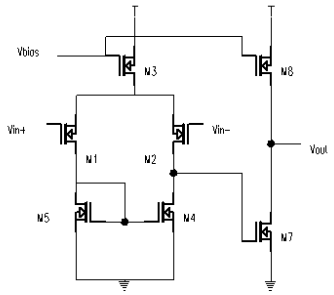


Figure 5.4: Two-Stage Amplifier

most popular choice in reported band-gap circuits. A basic two stage amplifier is shown in Figure 5.4.

The two-stage amplifier is simply two single stage amplifiers multiplied together. Simple small signal analysis finds the gain to be

$$A_V = (g_{m1} \cdot R_{DS1} || R_{DS5}) \cdot (g_{m7} \cdot R_{DS7} || R_{DS8}) = \mu_1 \frac{V_{A5}}{(V_{A1} + V_{A5})} \cdot \mu_7 \frac{V_{A8}}{(V_{A7} + V_{A8})} \tag{5.8}$$

This equation shows that a relatively large gain can be achieved by using a two-stage amplifier. If a long channel length is selected, a self gain of 200 can be achieved. Assuming the PMOS and NMOS Early voltages are identical, the voltage gain of the two-stage amplifier can approach 10000 [9]. While these may not approach the gain of commercial operational amplifiers, it is a respectable quantity in CMOS integrated circuits.

Another benefit of the two-stage design in the relatively small supply voltage needed. The minimum supply voltage required is the largest of the drain-to-source voltages of M1, M5, and M3, or M2, M4, and M3. The output stage is also very convenient. Only one transistor on each supply side needs to stay in saturation. This can result in achieving large gain even when the output is close to the rails.

Also, the systematic offset can be made very small. If the sizes of the input stage and output stage are properly sized, the output voltage will setup at $(V_{SUPPLY+} - V_{SUPPLY-})/2$. Minimizing the

total offset is one of the major design goals.

Unfortunately, the temperature behavior of the two-stage amplifier does not meet the temperature criteria. The gain of the amplifier is proportional to the self-gain of the input stage multiplied by the self-gain of the output stage. Since self-gain is proportional to mobility, the total gain is proportional to the temperature degradation of the first stage multiplied by the second stage, or using Equation 5.11 [10]

$$A_V \propto \frac{K_N}{T^{\alpha_N}} \cdot \frac{K_P}{T^{\alpha_P}} \tag{5.9}$$

α_P and α_N are often nearly identical, resulting in

$$A_V \propto \frac{K_N \cdot K_P}{T^{2 \cdot \alpha_N}} \tag{5.10}$$

This is an alarming conclusion. The gain of the amplifier decreases at the square of the rate of the mobility degradation! This circumstance is unacceptable since the amplifier is required to operate at elevated temperatures. Other alternatives will have to be considered.

5.4.4 Cascode Amplifier

Another common amplifier is the cascode amplifier. The cascode amplifier consists of a low gain first stage, followed by an output stage of very high gain.

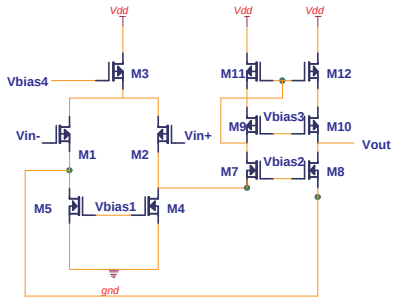


Figure 5.5: Cascode Amplifier

The output stage's gain can be made very high by using a cascode current mirror. Small signal analysis of the cascode amplifier results in an expected gain of

$$A_V = g_{m1} \cdot \frac{\mu_{10} \cdot r_{DS\,12}}{2} \tag{5.11}$$

Typical voltage gains of this amplifier type are around 5000 [9].

Several factors suggest this amplifier is not better choice than the two-stage amplifier. First, the gain is not as high as the two-stage amplifier. Since voltage gain is the main objective in the design of the amplifier, the cascode design does not look promising.

Also, thermal properties of the cascode amplifier are identical to the two-stage. The gain can again be found proportional to self-gain squared resulting in the following temperature dependence

$$A_V \propto \frac{K_N \cdot K_P}{T^{2 \cdot \alpha_N}} \tag{5.12}$$

Not only will the cascode have less gain than the two-stage, but will decrease with temperature at the same rate.

But several characteristics of the cascode are superior to the two-stage. One of these is these is frequency response. Since the amplifier consists mainly of one high gain stage, the gain-bandwidth-product (GBP) of the amplifier can be found through small signal analysis to be

$$GBP = \frac{g_{m1}}{C_{LOAD}} \tag{5.13}$$

This is a very attractable feature. To a first order, the GBP product can be set simply by adjusting the load capacitor. This process is far easier than the lead compensation required for the two-stage amplifier. The GBP can be set easily to maximize settling time and noise voltage [9].

The most important feature of the two-stage amplifier is the ability to boost the amplifier. In the same fashion a current mirror's output impedance can be elevated by cascoding, gain can be farther increased by multiplying the output impedance by a amplifier. Where a current mirror's

output impedance can be increased by μ which typically is somewhere between 50 to 200, a boosted current mirror is multiplied by values of up to 5000. This concept is discussed further in the next section [9].

5.4.5 Boosted-cascode Amplifier

A common method to achieve very high gain is using a concept called boosting. Boosting uses an amplifier to increase the output impedance of current mirror by a the gain of an amplifier. The output stage of a boosted-cascode amplifier is shown in Figure 5.6.

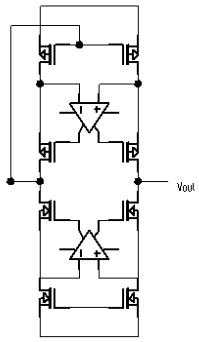


Figure 5.6: Boosted Cascode Amplifier

Small signal analysis can be used to show the output impedance to be

$$r_{OUTBoost} = (1 + A) \cdot r_{OUT} \tag{5.14}$$

If the cascoding transistors are boosted, the gain of the cascode stage is increased by the gain of the amplifier. Thus, by boosting both rail sides of the amplifier, the can be increased by a factor of more than A^2 [9]!

Ideally the gain of the boost amplifiers would be selected as large as possible. Unfortunately, there are constraints on the gain and bandwidth of boosting amplifiers. The bandwidth of the boosting amplifiers must be selected to ensure proper frequency response and stability [15].

To satisfy these constraints, the boosted amplifiers' transconductance are usually scaled back by a factor of four. This can be done using the same bias generator and reducing the widths of the devices by four.

As with all advances in electronics, the benefits of boosting are not achieved without cost. Obviously, the total power of the amplifier is increased. Although this goes against the goal of operating at the lowest power possible, the boosted amplifiers are scaled by a factor of four, thus reducing the power consumed by four. Thus, adding the boost amplifiers has the same power effect as doubling the widths on the main amplifier. The small power increase results in very large benefits in gain. Also, common-mode feedback circuits must be added for each boosting amplifier to ensure stability and correct DC biasing. This adds to the already growing power budget and area requirement [16].

Of course, the gain of the boost amplifiers will be temperature independent. Since the boost amplifiers use cascode topology, they will suffer the same degradation seen earlier. Although, even seriously degraded boosted amplifiers will be beneficial. For example, even if the gain of the boosts drop down to ten, the results gain increase will be over one hundred [15].

Due to the superior voltage gain and temperature performance of the boosted-cascode amplifier, it was chosen as the topology of the amplifier for the band-gap voltage reference. Although it is a more complex circuit and draws more power than other amplifiers analyzed, the need for high gain and strong temperature performance out-weigh this problems. The design and simulation results of the amplifier are discussed in the next section.

5.5 Amplifier Implementation and Simulation Results

5.5.1 Bias Generator Design (Saturation)

After selecting an architecture for the amplifier, the amplifier was designed to meet the requirements. The performance of the amplifier is largely set by the bias arrangement. Thus, the bias generator was developed first. The first design choice to be made was the selection of the transistor length. The transistor's channel length is directly proportional to gain and a large channel length also drives

down offset [9],[11]. But increasing L lowers transconductance, which in turn lowers bandwidth. Fortunately, high bandwidth is not a large requirement. With this in mind, a relatively conservative value of 5 μm was chosen for channel length.

The next parameter to set is the value of the overdrive voltage or V_{EFF}. Selecting overdrive voltage is a critical design choice. A small overdrive voltage increases self gain (Equation 5.12) and transconductance, but also increases the offset voltage. Also, the smaller the overdrive voltage, the higher the probability of a transistor dropping out of saturation, resulting in large degradation. Again a conservative choice, a value of approximately 250 milli-Volts was selected. While this value does not maximize gain, it does reduce the overall power (Equation 5.5 (c)) by keeping the current small.

Finally, the choice of the bias resistor and the nominal width must be chosen to set transconductance. A value of 20 μm and a resistor of 10 $k\Omega$ were selected for the nominal width and bias resistor, respectively. These values biased the transistors at an overdrive voltage of approximately 250 milli-volts to insure operation in the saturation region. The resulting in a bias transconductance for the NMOS transistors was approximately 103 ($\mu A/V$) and 92 ($\mu A/V$) for the PMOS transistors.

This bias generator used in this circuit had to be modified from the bias generator shown in Figure 5.2. Due to the goal of operating the circuit at as low supply voltage as possible, the NMOS cascode stage was removed. This saves the voltage of the threshold voltage and the overdrive voltage of the NMOS, easily over a volt (the threshold voltage alone is 700 milli-Volts). To create the second bias point for the NMOS cascode amplifier, a third leg of the bias generator was added. The overdrive voltage can be easily set by adjusting the current through the device and its width. Also, M5 and M6 were sized to allow for the second bias point for the PMOS cascode. With these adjustments, the bias generator was designed as shown in Figure 5.7.

5.5.2 Bias Generator Design (Subthreshold)

As mentioned before, a bias generator can become a subthreshold bias generator by decreasing the scaling term K or increasing the resistor. A subthreshold bias generator was developed by selecting

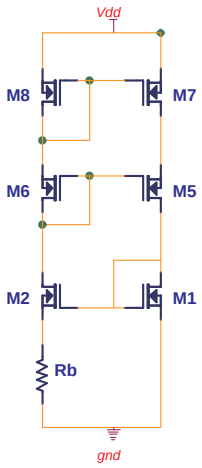


Figure 5.7: Bias Generator

a scaling term of two in Figure 5.7.

This allowed an overdrive voltage of -65 milli-volts and a transconductance of 83.75 ($\mu A/V$) in the NMOS transistors. The PMOS transistors setup at an overdrive of -50 milli-volts and a transconductance of 86.33 ($\mu A/V$).

Notice how despite the transistors are biased in the subthreshold region, the transconductances are nearly equal to saturation bias generator. This is due to the exponential relationship of the subthreshold region. Also, the subthreshold bias generator only draws 19 μA and operates on a minimum supply of 1.5 V! This is a big improvement of the saturation bias generator which requires a 2 volt supply.

5.5.3 Main Amplifier Design

With the bias generator designed, the main amplifier can be created. The first decision is to use a NMOS or PMOS input stage. The decision to use a NMOS or PMOS input stage is based on the expected input DC voltage. If the input voltage is expected to be close to the upper supply voltage, a NMOS input stage is preferred. This is due to the voltage needed to operate a transistor in the saturation region. If the input voltage is not less than the upper supply by at least a threshold voltage,

the input stage can not be in saturation. Since the input voltage of the amplifier is the diode voltage, a PMOS stage is preferred (the diode voltage must be biased lower than one volt) [9]. The size of the PMOS input transistors were also increased to drive down the input offset voltage.

The current of all legs of the amplifier were biased to be equal to the current of the bias generator. Transistors M4 and M5 have double the width to allow the current of both the input and output stage. Since the band-gap reference requires only a single ended amplifier, the output stage is converted from a differential gain stage to a single ended gain stage by tying the gate of M11 to the drain of M9. The resulting amplifier is shown in Figure 5.8 with transistor sizes shown in Table 5.1.

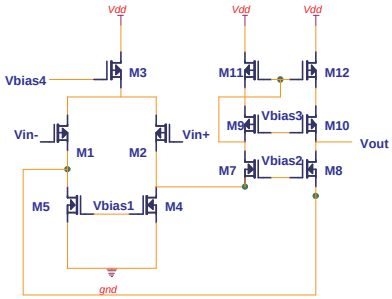


Figure 5.8: Cascode Amplifier

Transistor	W (μm)	L (μm)	Fingers
M1	10	5	32
M2	10	5	32
M3	10	5	64
M4	5	5	64
M5	5	5	64
M6	5	5	32
M7	5	5	32
M8	10	5	32
M9	10	5	32
M10	10	5	32
M11	10	5	32

Table 5.1: Amplifier Transistor Sizes

Figures 5.9 and 5.10 show the Cadence AC simulation results of the main amplifier with the two biasing arrangements. The saturation amplifier has a DC gain of approximately 52 dB (400) and a GBP of 2 MHz. The subthreshold amplifier achieves a gain of 67 dB (2200) and a GBP of 175 kHz. The small gain of the amplifier illustrates the need of boosting stages.

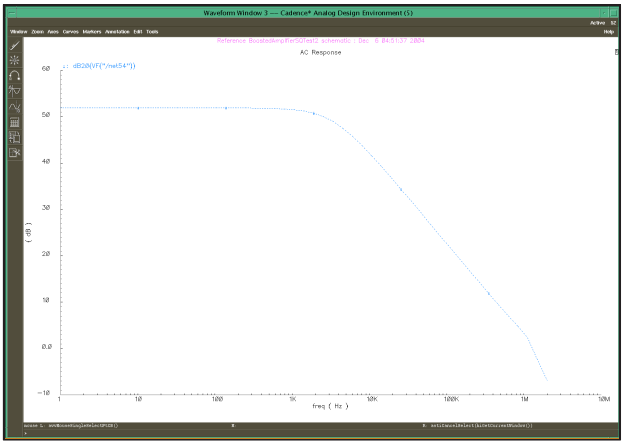


Figure 5.9: Amplifier simulation results (saturation)

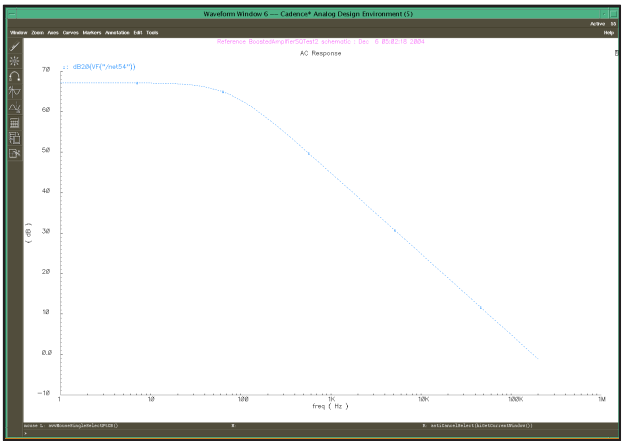


Figure 5.10: Amplifier simulation results (subthreshold)

5.5.4 Boosting Amplifiers

To maximize the effects of boosting, the gain enhancing amplifiers consist of entire cascode amplifiers. Since the application requires only a single sided amplifier, the boosted stages are of the same topology as the main amplifier. As discussed earlier, the transconductances are scaled down by a factor of four to avoid stability problems [15]. This was done simply by reducing the nominal width to 5 μm from 20 μm. The lower boost amplifier consists of a PMOS input stage, and a NMOS input stage was chosen for the upper boost.

5.6 Boosted Amplifier

5.6.1 Saturation Bias

Figure 5.11 demonstrates the simulated gain and frequency response of the boosted amplifier. As shown, the DC gain is approximately 110 dB (316000) and a GBP of 1.67 MHz. To maximize stability, the phase margin is approximately 80 degrees.

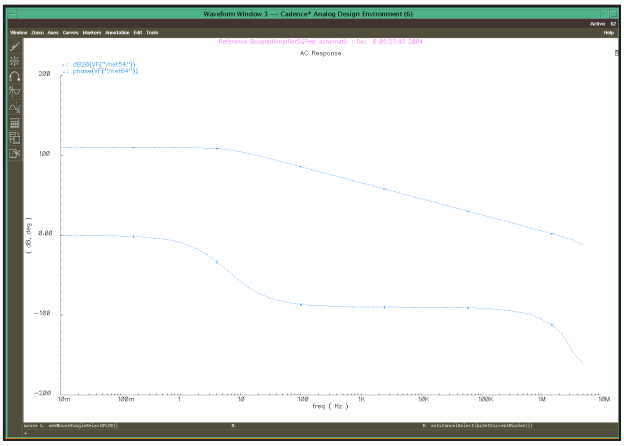


Figure 5.11: Amplifier simulation results (Saturation)

Figure 5.12 shows the simulated temperature dependence on the gain. As expected, the gain does decrease with temperature. But, the boosting keeps the gain very large at temperatures past 150 °C.

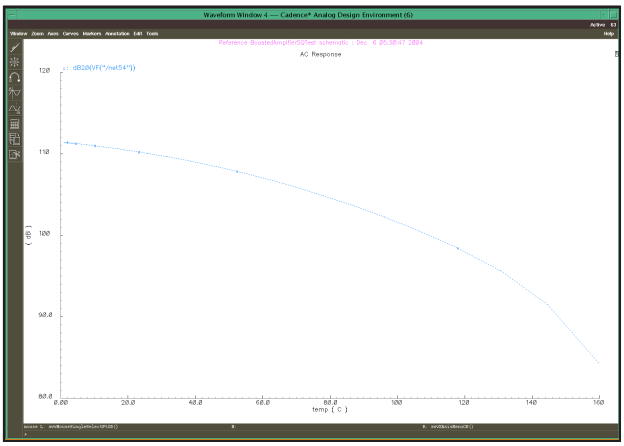


Figure 5.12: Amplifier simulation results (Saturation)

5.6.2 Subthreshold Bias

Figure 5.13 shows the simulated gain and frequency response of the boosted amplifier with subthreshold bias. As expected the amplifier achieves tremendous gain. The DC gain is approximately 136 dB (5.6 million) and the GBP is approximately 500 kHz.

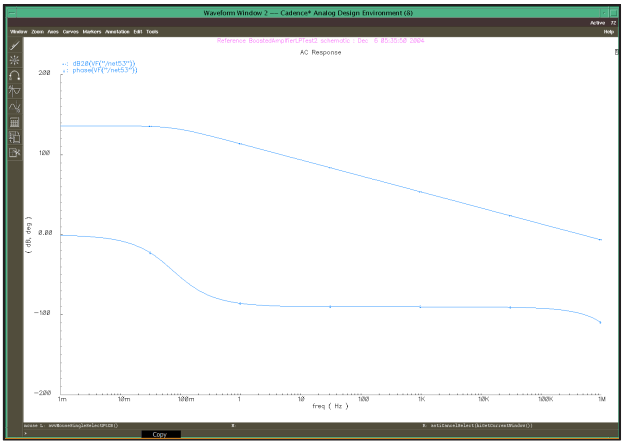


Figure 5.13: Amplifier simulation results (Subthreshold)

Figure 5.14 illustrates the simulated temperature performance of the boosted amplifier. The amplifier achieves elevated gain at high temperatures.

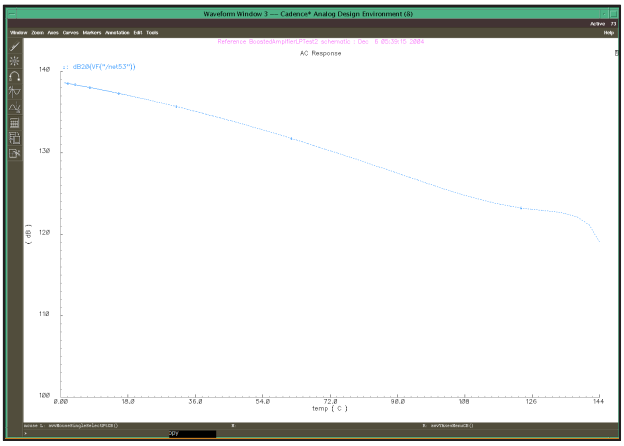


Figure 5.14: Amplifier simulation results (Subthreshold)

While these figures are very encouraging, it is worth mentioning that they are extremely dependent on the transistor models. These estimates should be considered a best case situation. Fabrication parasitics not modelled (resistance of metal contacts, variance in geometry) will have a negative effect on performance.

The simulation also predicts encouraging results for the goal of reducing the minimum supply voltage and power consumption. Table 5.2 shows the breakdown of the power budget for the boosted amplifier.

	Saturation Amplifier	**Subthreshold Amplifier**
Total Power	121 μW	900 μW
Bias Generator	46.3 %	5.7 %
Main Amplifier	33.38 %	42.6 %
Lower Boost	8.3 %	21.6 %
Upper Boost	8.4 %	20.1 %
CMFB	4 %	10 %

Table 5.2: Power Distribution

One might be surprised by the power consumed by the bias generator. But since the generator is biased in solid saturation using relatively large transistors, the large power contribution helps ensure the amplifier will setup at proper voltages. Designing the bias generator correctly goes far

in increasing the yield of the amplifier. The power breakdown also shows another advantage of the subthreshold design. The boosting amplifiers only contributed 17 percent of the power consumption, or only a 21 μW increase in power. This is a small price for an increase in gain of over 60 dB!

Chapter 6

Measurements and Conclusions

6.1 Introduction

This chapter summarizes the simulation and measurement results of the band-gap reference. The chosen architecture for the reference was a CMOS band-gap topology with a feedback amplifier. Also, the current mirror was casocoded to keep the currents in each branch equal. The complete design is shown in Figure 6.1.

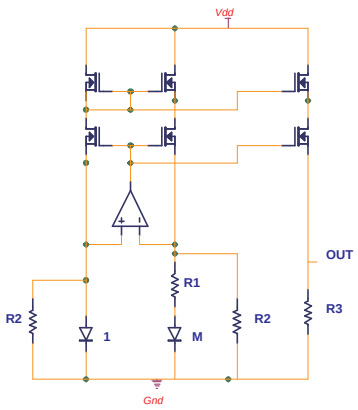


Figure 6.1: Gain Enhanced Band-Gap Reference

The design was simulated using the Cadence Design Suite. Finally, the circuit was fabricated in

the Peregrine SOS process. The simulation and measurement results are presented here.

6.2 Simulation Results

As described in the **Amplifier Design** chapter, two versions of the band-gap referenced were designed. The designs are identical, except one version uses an amplifier biased in the saturation region, the other with an amplifier biased in the subthreshold region.

6.2.1 Band-gap Reference with Saturation Amplifier

The reference was designed to achieve an output voltage of .5 V. The simulated reference consumes 945 μW with a three volt supply. The reference operates with a minimum supply of 1.75 V. The reference achieves a temperature dependence of 144 ppm/°C (over a 150 °range) and a voltage supply dependence of 40 ppm/V. The temperature and voltage supply dependence are illustrated in the following plots.

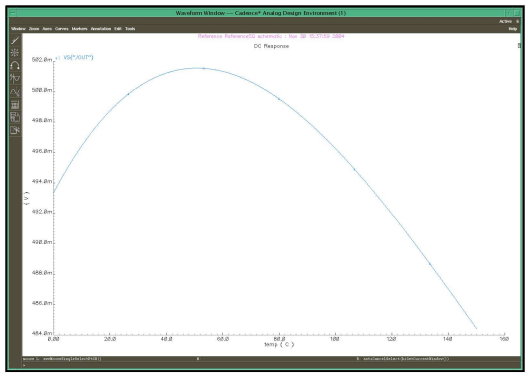


Figure 6.2: Gain Enhanced Band-Gap Reference Temperature Dependence with Saturation Amplifier

6.2.2 Band-gap Reference with Subthreshold Amplifier

A second band-gap reference was designed and simulated. The design is identical to the first with the exception that the amplifier is biased in the subthreshold region. The subthreshold amplifier

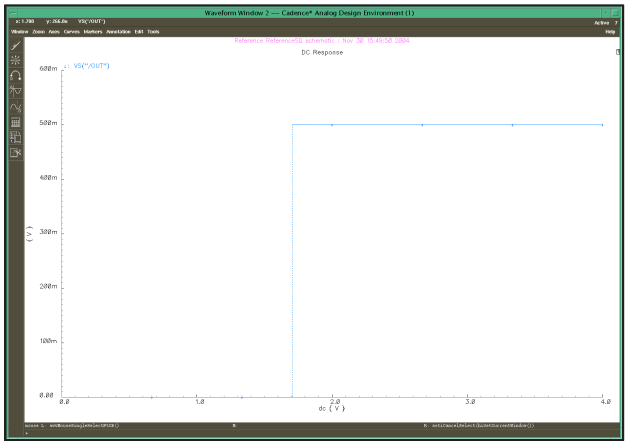


Figure 6.3: Gain Enhanced Band-Gap Reference Power Supply Dependence with Saturation Amplifier

achieved elevated gain with significantly less power.

The simulated reference consumes 945 μW with a three volt supply. The reference operates with a minimum supply of 1.75 V. The reference achieves a temperature dependence of 144 ppm/°C (over a 150 °range) and a voltage supply dependence of 40 ppm/V. The temperature and voltage supply dependence are illustrated in the following plots.

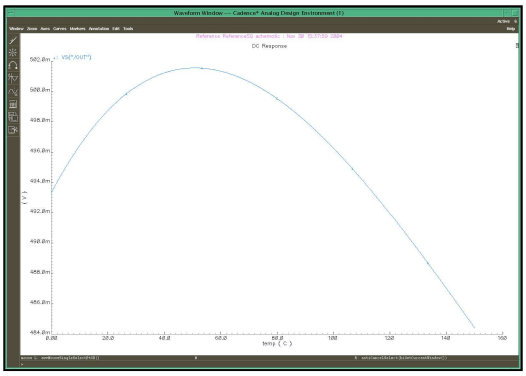


Figure 6.4: Gain Enhanced Band-Gap Reference Temperature Dependence with Subthreshold Amplifier

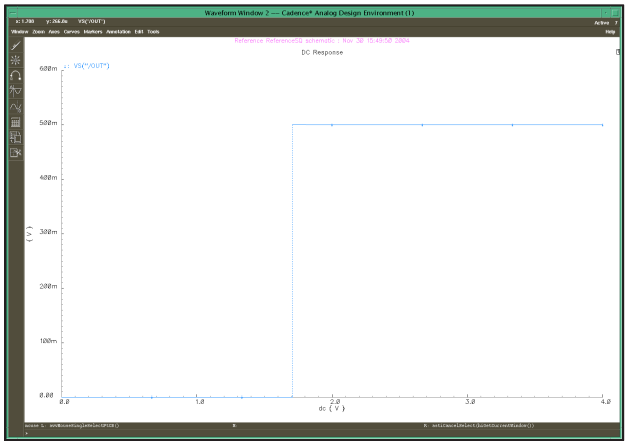


Figure 6.5: Gain Enhanced Band-Gap Reference Power Supply Dependence with Subthreshold Amplifier

6.3 Measurement Results

The two designs were prepared for fabrication using the Cadence Design Suite. The circuit was fabricated in the Peregrine SOS process. The circuit was measured using the HP4155 Semiconductor Analyzer to determine voltage supply dependence. The circuit was analyzed at elevated temperatures using the Cascade probe station and MSVLSI GPBI Controller Ver. 1.4 software for automated measurements.

6.3.1 Band-gap Reference with Saturation Amplifier

Unfortunately, the band-gap reference with a saturation amplifier did not function properly. Several possible explanations exist for the reference not functioning. These include device deviations from expected values. For example, the PMOS devices displayed different mobility and output impedances than modelled. The most probable reason for the circuit not functioning is the complexity of the layout.

6.3.2 Band-gap Reference with Subthreshold Amplifier

Unlike the band-gap reference with a saturation amplifier, the version with a subthreshold amplifier achieved good results. The reference achieved an average output voltage of .503 V (at room temperature) while drawing only 310 μW from a three volt supply. The average temperature dependence was 120 ppm/ °C (over a 120 °range), while the average voltage supply dependence was 975 ppm/ V. The following plots summarize the performance of the reference.

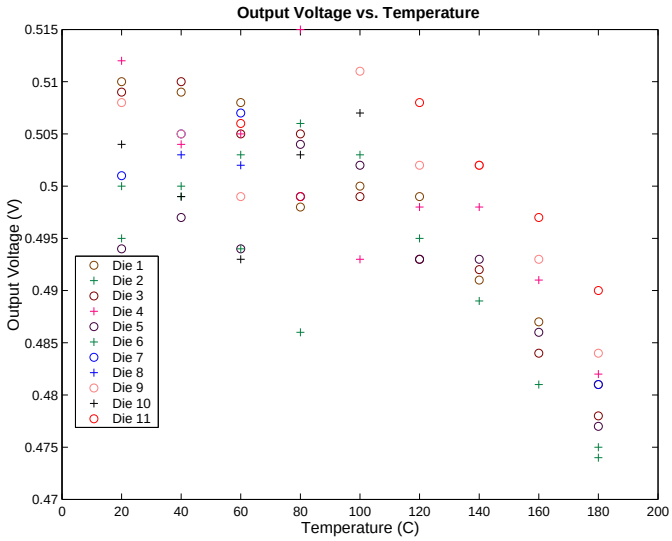


Figure 6.6: Gain Enhanced Band-Gap Reference Temperature Dependence with Subthreshold Amplifier

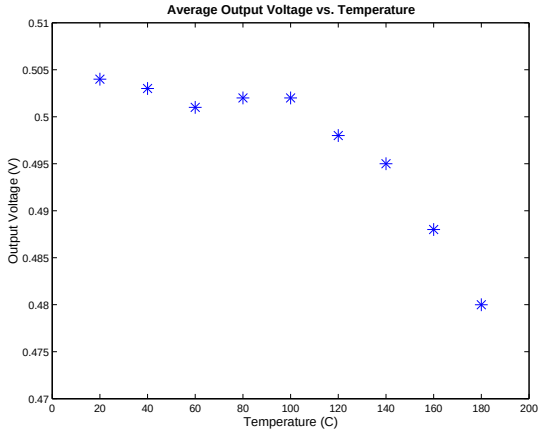


Figure 6.7: Gain Enhanced Band-Gap Reference Temperature Dependence with Subthreshold Amplifier (Average)

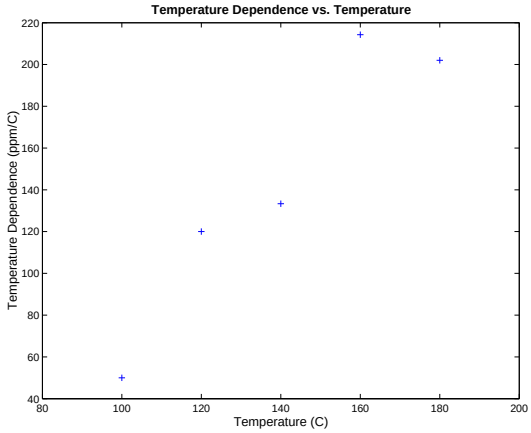


Figure 6.8: Gain Enhanced Band-Gap Reference Temperature Dependence with Subthreshold Amplifier

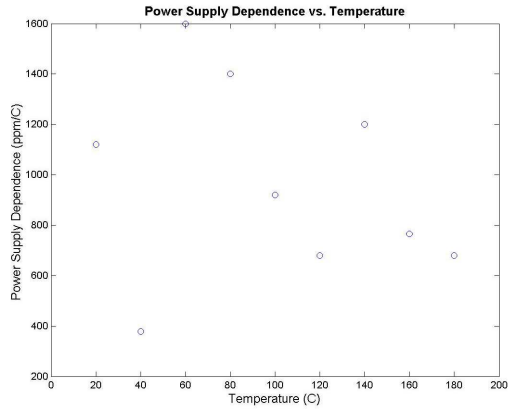


Figure 6.9: Gain Enhanced Band-Gap Reference Power Supply Dependence with Subthreshold Amplifier

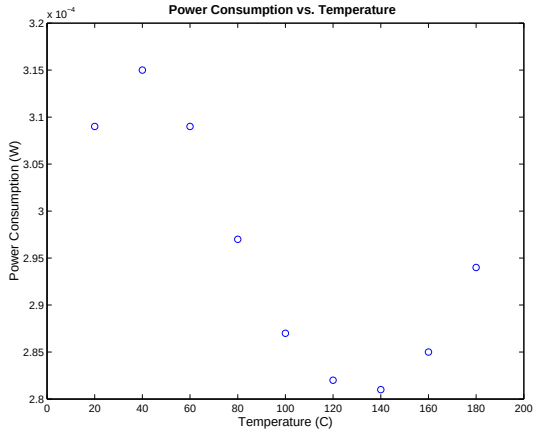


Figure 6.10: Gain Enhanced Band-Gap Reference Power Consumption with Subthreshold Amplifier

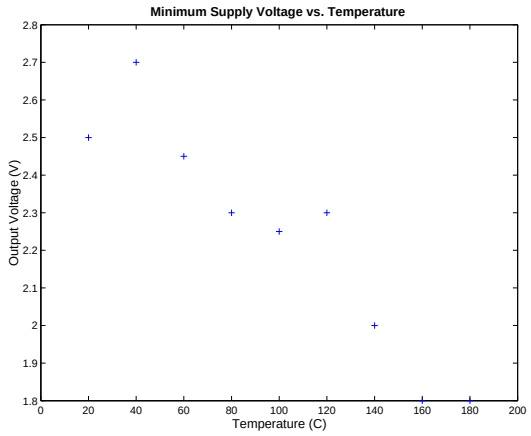


Figure 6.11: Gain Enhanced Band-Gap Reference Minimum Power Supply with Subthreshold Amplifier

6.4 Conclusion

While the results of the saturation design were not pleasing, the first run success of the subthreshold design was a modest achievement. The cause of the saturation reference incorrect performance is not fully understood at this time. Although, several factors could contribute to its failure. First, the design and layout are fairly complex. The layout was verified to be identical against the schematic, but layout parasitics could result in degraded performance or difficulty in testing the device. Also, several interesting characteristics were observed in both the NMOS and PMOS in testing the devices. Both transistors exhibited large degradation when the devices were biased at large overdrive and drain-to-source voltages. Although these problems cannot fully explain the failure at this time. The most interesting fact about the references is the layout for the saturation version was completed first, the modified only by changing the bias generator. Otherwise, the layouts are identical.

Fortunately, the results for the subthreshold voltage reference were very good. The average output voltage deviating only three milli-volts from the expected value. This displays the matching efforts in the amplifier, resistors, current mirror, and diodes was done properly. Also, the temperature dependence accurately following the simulated results shows the Verilog-A model for the gated-

diode was very accurate. A tempco of 55 ppm/°C over a range of 100 °C is comparable to current published work. Plus, the reference was able to achieve a tempco of 200 ppm °C upto 185 °C. The power supply dependence was not as expected at a modest 975 ppm/ V, but this is more-than-likely related to the drain-to-source voltage problem seen in the saturation version (as the power supply increases, the drain-to-source voltage increases). This effort conclude that subthreshold design can be used effectively in high temperature circuit applications. The following table is a comparison between simulated performance and measured performance of the circuit. The results match very well.

	Simulation	Measurement
Power Consumption	180 μW	310 μW
Minimum Supply Voltage	1.7 V	2.2 V
Tempco	182 ppm/ °C	133 ppm/ °C
PSRR	123 ppm / V	975 ppm/ V

Table 6.1: Comparison of Band-Gap References

Bibliography

[1] Y. Hu and M. Sawan, ”A 900 mV 25 μW High PSRR CMOS Voltage Reference Dedicated to Implantable Micro-Devices,” *Proceedings of the 2003 International Symposium on Circuits and Systems*, May 2003.

[2] K. Leung and P. Mok, ”A Sub-1-V 15-ppm/°C CMOS Bandgap Voltage Reference Without Requiring Low Threshold Voltage Device,” *IEEE J. Solid-State Circuits*, vol. 37, pp. 526-530, April 2002.

[3] Y. Leung and E.K.F. Lee, ”Design of low-voltage bandgap reference using transimpedance amplifier,” *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 552-555, June 2000.

[4] A.J. Annema, ”Low-power bandgap references featuring DTMOSTs,” *IEEE J. Solid-State Circuits*, vol. 34, pp. 949-955, July 1999.

[5] P. Malcovati, F. Maloberti, M. Pruzzi, and C. Fiocchi, ”Curvature compensated BiCMOS bandgap with 1-V supply voltage,” *Proc. ESS-CIRC’ 2000*, pp. 52-55.

[6] S. Laberge and G.W. Roberts , ”Temperature compensated CMOS voltage reference,” *2001 Southwest Symposium on Mixed-Signal Circuit Design*, vol. 1 pp. 26-29, May, 2001.

[7] S. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits Analysis and Design*, 2nd ed. McGraww-Hill: 1999.

[8] M. Lee and K. Asada , ”A High Speed and Low Power on CMOS/SOI Technology and its Modeling,” *Proceedings on the 36th Midwest Symposium on Circuits and Systems*, vol. 2 pp. 1040-1043, August, 1993.

[9] D.A. Johns and K. Marin, *Analog Integrated Circuit Design*, 1st ed. New York: 1997.

[10] P.E. Allen and D. Holberg, *CMOS Analog Integrated Circuit Design*, 2nd ed. Oxford: 2002.

[11] M. Pelfrom, C.J. Duimaijer, and P.G. Welbers, ”Matching Properties of MOS Transistors,” *IEEE Journal of Solid-State Circuits*, vol. 24, pp. 1433-1440, October, 1989.

[12] P.R. Gray, P.J. Hurst, S.H. Lewis, and R.G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed. Wiley: 2001.

[13] L. Najafizadeh and I.M. Filanovsky, ”Towards a Sub-1 V CMOS Voltage Reference,” *Proceedings of the 2004 International Symposium on Circuits and Systems*, pp. 53-56, May 2004.

[14] B. Verzi and P. Aum, ”Automated Gated Diode Measurements for Device Characterization ,” *Proceedings of the 1994 International Conference on Microelectronic Test Structures*, pp. 141-146, March, 1994.

[15] K. Bult and G.J. Geehan, ”A Fast-Settling CMOS op-amp for SC circuits with 90 dB DC Gain,” *IEEE Journal of Solid-State Circuits*, vol. 25, pp. 1379-1384, December, 1990.

[16] J. Yuan and N. Farhart, ”A Compensation based Optimazation Methodology for Gain-Boosted Opamp,” *Proceedings of the 2004 International Symposium on Circuits and Systems*, pp. 669-672, May 2004.

Appendix A

```
// VerilogA for Verilog, Gateddiode, veriloga

‘include "constants.h" ‘include "discipline.h"



module gateddiode(anode,cathode);
 inout anode,cathode;
 electrical anode,cathode;

 parameter real area = 1 from (0:inf);
 parameter real is=2.0019079135444e-33 from (0:inf);
 parameter real n=1 from (0:inf);
 parameter real cjo=0 from [0:inf);
 parameter real m=0.5 from [0:inf);
 parameter real phi=0.7 exclude 0;
 parameter real fc=0.5 from (0:1];
 parameter real tt=1p from [0:inf);
```

```
 parameter real rs=0 from [0:inf);
 parameter real q=1.60217733e-19 from [0:inf);
 parameter real k=1.380658e-23 from [0:inf);

 real Vd, Id, Qd;
 real f1, f2, f3, fcp;
 real ibv;

 analog
   begin
     @(initial_step)
     begin
       f1 = (phi/(1 - m))*(1 - pow((1 - fc), m));
       f2 = pow((1 - fc), (1 + m));
       f3 = 1 - fc*(1 + m);
       fcp = fc*phi;

     end

     Vd = V(anode,cathode);
     Id = I(anode);

// intrinsic diode.
     I(anode) <+ is*exp(Vd/((5.3199239e-6*$vt*q/k) +    // exponential terms
                (9.481423e-8*$vt*$vt*q*q/(k*k)) + // exponential terms
                (1.97479233513e-2)))*  // exponential terms
                exp(1.6938487052531e-1*$vt*q/k)*  // saturation terms
```

```
                exp(-1.11860604344e-4*$vt*$vt*q*q/(k*k)); // saturation terms


   end
endmodule
```

VITA

Jerry Brewer

Candidate for the Degree of

Master of Science

Thesis: High Temperature SOI CMOS Band-Gap Voltage Reference

Major Field: Electrical Engineering

Education: Graduated from Webbers Falls High School, Webbers Falls, Oklahoma, in May 1998; received Associate of Science from Connors State College, Warner, Oklahoma, in May 2000; received Bachelor of Science in Electrical Engineering from Oklahoma State University, Stillwater, Oklahoma, in December 2002. Completed the Requirements for the Masters of Science in Electrical Engineering at Oklahoma State University in Decemeber 2004.

Experience: Completed high school in small, rural area. Attended junior college for two years before transferring to Oklahoma State University to complete bachelor's and master's work. Worked as a teaching assistant 2000-2001 and as a research assistant 2001-2004. Employed by SPAWAR Systems Center May 2004 to present.

Professional Memberships: Institute of Electrical and Electronic Engineers (IEEE).

Name: Jerry Brewer Date of Degree: December, 2004

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: High Temperature SOI CMOS Band-Gap Voltage Reference

Pages in Study: 79 Candidate for the Degree of Master of Science

Scope and Method of Study

The purpose of this study was to design a Silicon-On-Insulator (SOI) CMOS band-gap voltage reference that achieves a temperature dependence of 300 ppm/ °C. The procedure consisted of three major portions: test, design and simulation, and layout for fabrication. Measurements were completed on two SOI structures, the gated-diode and parasitic bipolar junction transistor (BJT), which determined the gated-diode was more suitable for the voltage reference. The band-gap voltage reference was designed and simulated using the Cadence Design Suite. Finally, the design was prepared for fabrication using the Cadence Design Suite. The circuit was fabricated and tested for comparison against design and simulation results.

Findings and Conclusions

The measurements conducted on the gated-diode and BJT indicated the gated-diode is the preferred solution in the voltage reference. A new model was developed to simulate the measured temperature behavior of the device. The gated-diode exhibited temperature dependence undocumented in current literature. Several small additions to recently published efforts provided improvements in the simulated performance of the circuit, particularly in extending the range of operating temperatures. The design's simulation results exhibited a temperature dependence of 182 ppm/ °C, voltage supply dependence of 123 ppm / V, consumes 180 μW of power, and occupies 460 μm x 3100 μm area. The fabricated circuit achieved a voltage supply dependence of 975 ppm/ V, a minimum supply of 2.5 V, and consumes 310 μW of power. The temperature dependence was measured as 120 ppm / °C at temperatures up to 120 °C, and 202 ppm/ °C at 180 °C.

Advisor's Approval _____