A Low Power High Power Supply Rejection Ratio

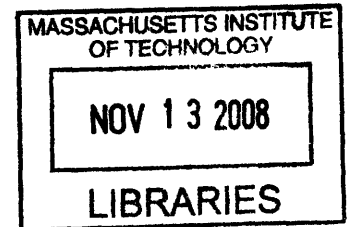# Bandgap Reference for Portable Applications

by

Siddharth Sundar

S.B. EE, M.I.T., 2007

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

January, 2008

Author_____
Department of Electrical Engineering and Computer Science
January 7, 2008

Certified by_____   _____
Eric Kimball
Design Manager, NXP Semiconductors
VI-A Company Thesis Supervisor

Certified by_____   _____
Michael Perrott
Associate Professor of Electrical Engineering
M.I.T. Thesis Supervisor

Accepted By_____
_____
Arthur C. Smith
Professor of Electrical Engineering
Chairman, Department Committee on Graduate Theses

ARCHIVES

A Low Power High Power Supply Rejection Ratio

Bandgap Reference for Portable Applications

By

Siddharth Sundar

Submitted to the Department of Electrical Engineering and Computer Science
on January 14, 2008 in Partial Fulfillment of the Requirements for the Degree of
Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

A multistage bandgap circuit with very high power supply rejection ratio was designed
and simulated. The key features of this bandgap include multiple power modes, low
power consumption and a novel resistor trimming strategy. This design was completed in
deep submicron CMOS technology, and is especially suited for portable applications.

The bandgap designed achieves over 90 dB of power supply rejection and less than 17
microvolts of noise without any external filtering. With an external filtering capacitor,
this performance is significantly enhanced. In addition, the design includes an efficient
voltage-to-current converter and a fast-charge circuit for charging the external capacitor.

Thesis Supervisor: Michael Perrott
Title: Associate Professor of Electrical Engineering

Acknowledgments

To my family - for always being loving, caring and supportive of me. None of this would have been possible without your constant encouragement and support.

To Mike Mills, Riad Wahby, and Marius Goldberg - for introducing me to analog design and making sure that I had the greatest possible internship experiences at Silicon Labs (in addition to being some of the coolest people I know). Without your help, I'd probably never have got into analog design.

To Prof. H.S.Lee, J.K Roberge, J. Alamo, D. Perrault and other professors at MIT: for teaching me what little I know about analog circuit design and its related fields.I'm also greatly indebted to Prof. Mike Perrott, who kindly agreed to supervise my thesis ,and provided valuable help and feedback.

To all of the people in the wireless division of NXP semiconductors - for picking out a great project for me to work on, and providing help and support whenever I needed it. It's been a pleasure to work in such an open, accepting and helpful atmosphere. In particular, thanks go out to my company supervisor, Eric Kimball. I know you're a busy person, and I really appreciate all the time and effort you took to help me all the way through this thesis - from start to finish.

Contents

List of Figures

Chapter 1

Introduction

Reference circuits are ubiquitous in modern circuit design. Ranging from regulators to analog-to digital converters, it is hard to find a block that does not use a reference voltage or current to set up its operating point. With multiple circuit blocks on a modern System on Chip(SoC) circuit, it becomes valuable to build a separate reference block that supplies references to all the other blocks on the chip. This way, the power and area cost of generating an accurate reference voltage is restricted to a single block.

However, with the recent advent of massive SoC designs, this has become harder and harder to do. The largest systems have many different circuit blocks, each with some combination of specifications, which may be mutually incompatible. There can be multiple modes of operation, each with a different set of requirements. Therefore, it has become valuable to develop a reference block that can accommodate many different sets of specifications, and has multiple power/operation modes.

In this thesis, we propose a structure for a novel bandgap reference block, designed to be used in a large SoC intended for use as a single chip cellphone. Although the design is intended and optimized for a particular implementation, the topology and

overall structure is flexible enough to enable similar designs to be implemented in other systems.

The proposed reference uses a unique two stage architecture, with the first stage acting as a power supply for the second stage in addition to supplying a reference. This allows us to efficiently have low and high power settings, and helps achieve high PSRR. In the following chapters, we describe both system-level and implementation details for this reference design.

Chapter 2 endeavors to give some background on reference circuits. We describe two major types of reference circuits - zener and bandgap based. Following this, we discuss the theory behind bandgap references. Common topologies are presented with a discussion of their advantages and disadvantages. We conclude this section by outlining the reasons for choosing a bandgap-based reference structure for this design.

Chapter 3 lays out the key specifications for the block, and how they are derived. In particular, emphasis is laid on the PSRR specification, which is critical for this block. We also discuss the various modes of operation. This is followed by a system level description of the reference block, highlighting the key innovations in the architecture that make the PSRR requirement manageable, and handle the low power and high power modes effectively.

Chapter 4 discusses the coarse bandgap, which is used to provide the reference in low power mode. We discuss the key features of this block that separate it from regular bandgap designs. This includes a biasing scheme that is unaffected by the low beta of the transistors, and an output stage that also functions as a regulator in the high power mode.

Chapter 5 talks about the fine bandgap, which is used to provide the reference voltage in high power mode. The main features of interest include the trimming strategy and

the biasing scheme for the design.

Chapter 6 discusses the vbg_out block, which is used to decide which bandgap voltage - the coarse or the fine to use. It is also responsible for sending out an OK signal to indicate when the coarse bandgap has powered up. In addition, it includes some fast charging circuitry to charge the output node rapidly at startup.

Chapter 7 describes the refgen_i block, which is used to convert the reference voltage to a current. Particular emphasis is laid on the trimming strategy, which saves an extra trimming step.

Chapter 8 describes some limitations of this implementation, and suggestions for future explorations in this area. It concluded with some global simulation results that were not covered in the previous chapters.

Chapter 2

Background

2.1 Introduction

The major goal of any reference circuit is to provide as constant a voltage and/or current as possible, irrespective of changes in the process or environment. Most metrics, such as PSRR (Power Supply Rejection Ratio) and variation with process depend strongly on circuit implementation, and thus have to be discussed in the context of an particular implementation. However, temperature variation is much more fundamentally related to the choice of topology, and less so with the exact implementation. For this reason, this chapter will focus on the temperature performance of various types of reference topologies.

There are two major types of reference circuits currently used - zener based references, and bandgap based references. Both of them attempt to use fundamental properties of silicon to set a fixed voltage even in the presence of large temperature variation. However, there are major differences between the two, and each comes with its own advantages and disadvantages. In section 2.2 we discuss zener based references,

14

while in section 2.3 we discuss bandgap based references. In section 2.4, we explain the reasons for choosing a bandgap-based reference over a zener diode for this design.

2.2 Zener based References

These references are based on using the reverse breakdown region of a zener diode. Because the I-V characteristics of a diode are steep in this region, the output voltage does not drift significantly with changes in operating conditions, as long as a reasonable bias current is maintained.
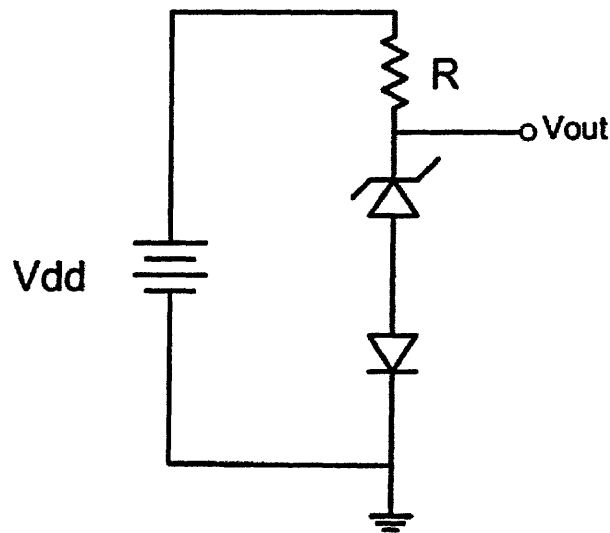
Figure 2.1: Zener Diode Based Reference

The breakdown voltage is a strong function of the doping levels of the diode, and can be controlled quite precisely. The large amount of current at the breakdown voltage is generated by two mechanisms - avalanche breakdown and tunneling from the conduction band to the valence band. Avalanche breakdown has a positive temperature coefficient (i.e., as temperature increases, the amount of current due to avalanche breakdown in-

creases), while the tunneling phenomenon has a negative temperature coefficient. For a doping corresponding to a breakdown voltage of about 5.6V, these temperature coefficients nearly cancel out around room temperature, which implies that the output voltage does not change much with temperature.

There are considerable issues in fabricating such a diode on a modern IC. Surface imperfections and crystal defects cause diodes built at the surface to be noisy and vary with temperature[9]. To avoid these issues, the buried zener reference was invented. These references are built under the silicon surface, and exhibit very good temperature and noise performance. Temperature coefficients of as low as 1 ppm/°C have been reported in commercial designs, along with a noise floor of less than $100nV/\sqrt{Hz}$[10]. Therefore, in terms of temperature and noise performance, they are the best available solution for reference circuits.

Unfortunately, this high performance comes at a steep cost. Most Zener diodes have a breakdown voltage above 5V, thus entailing the need for a power supply well above 5V. In addition, the diodes need to be biased at a fairly high current (typically a few mA) to ensure that the diode operates in the high slope region of the I-V curve (to reduce voltage variation). This, unfortunately leads to high power consumption for these references.

2.3   Bandgap Based References

The fundamental premise of a bandgap reference is to generate a temperature independent voltage by adding two (or more) voltages which have opposite temperature dependence. Thus, when added in the right proportion,the sum is independent of temperature variations (around a particular reference temperature).

16

V_{dd}

V_{BE}

Σ

$V_{OUT} = V_{BE} + KV_t$

KV_t

$V_t = \dfrac{kT}{q}$

V_t

K
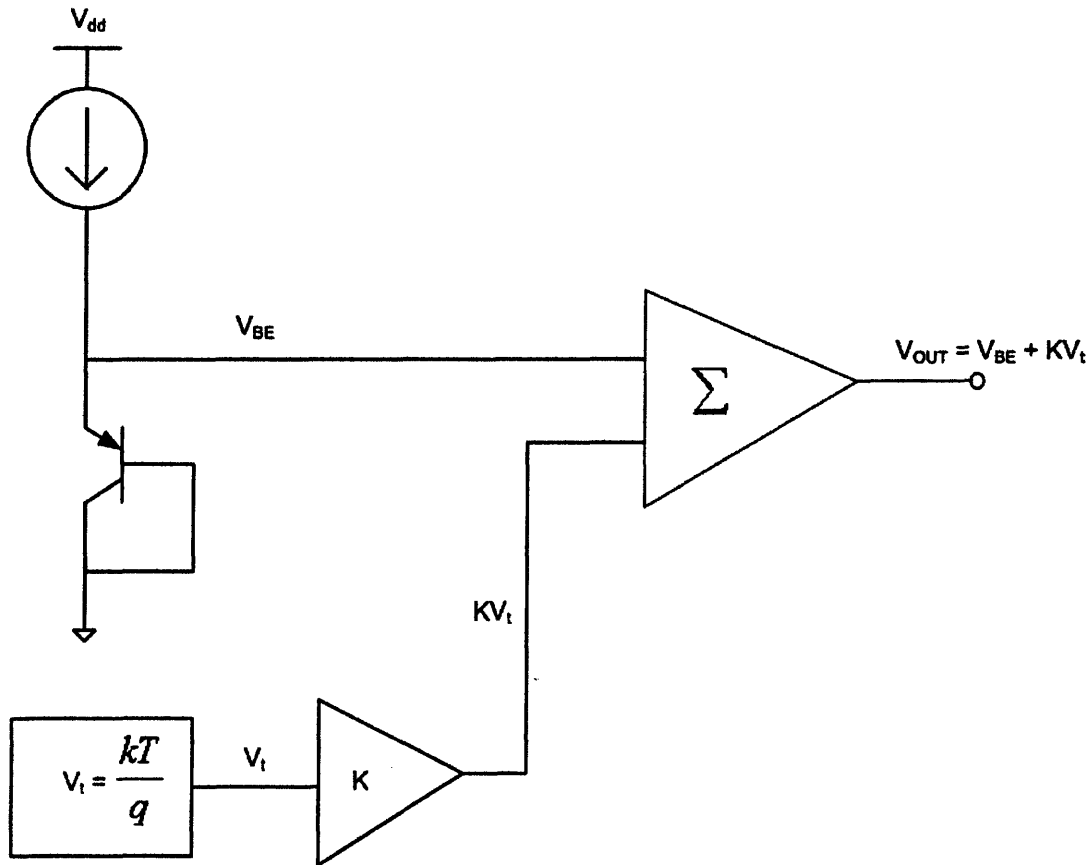
Figure 2.2: Conceptual Idea of a Bandgap Reference

Figure 2.2 shows a conceptual diagram for a first order bandgap reference. The base-emmiter voltage of a bipolar transistor has a complex dependence on temperature, which is typically expressed using a Taylor series expansion. In a first order bandgap circuit, one generates a voltage directly proportional to temperature (normally known as a PTAT voltage). This voltage (V_t in the figure) is then added (after proper scaling by the coefficient K) to the base-emitter voltage. With proper choice of the coefficient K, the KV_t term exactly cancels out the first order term in the Taylor series expansion of $V_{BE}(T)$. This gives us a voltage V_{REF}, which is temperature independent to the first

17

order.

One of the most popular ways to implement a first order bandgap reference is the Brokaw Cell, first published by Paul Brokaw in 1984[1].
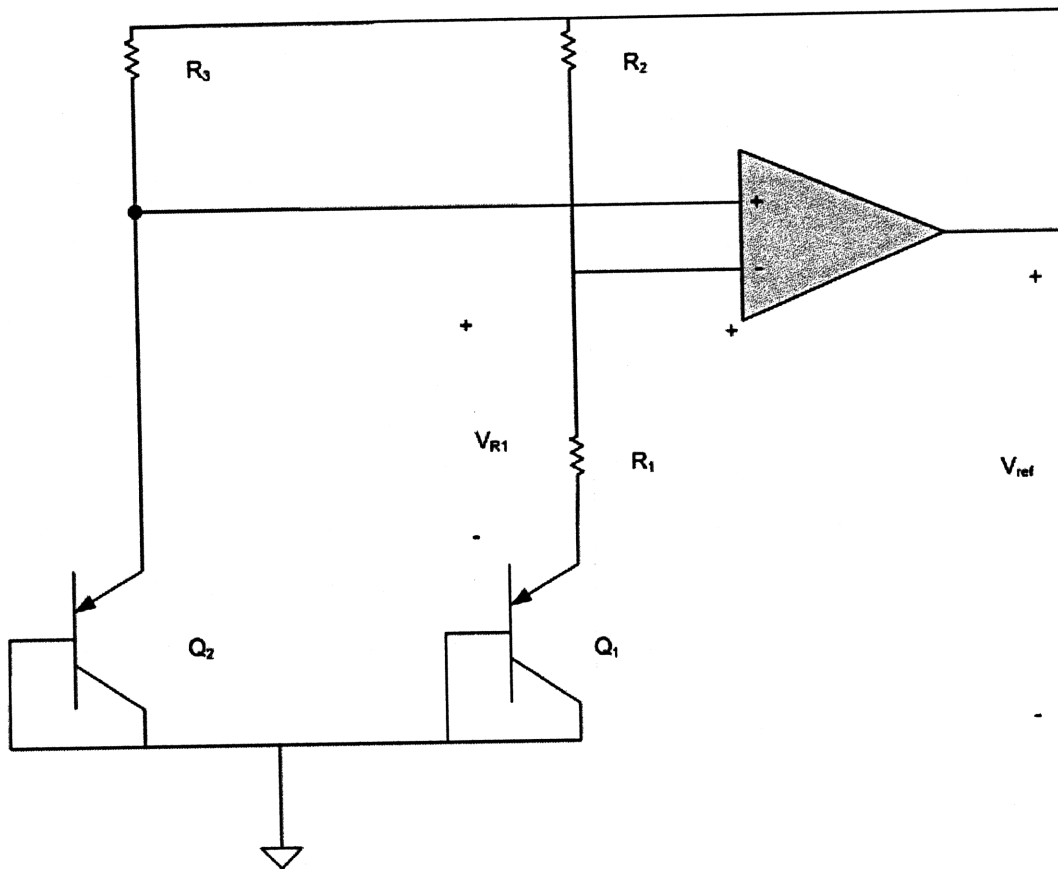
Figure 2.3: An Implementation of the Brokaw Cell

In this circuit (Figure 3.3), V_{R1} generates V_t, while the resistor ratio R_2/R_1 takes the place of the scaling coefficient K. In order to figure out the right value of the scaling coefficient, it is important to understand the temperature dependence of V_{BE}. In his 1980 paper[2],Tsividis outlined the following temperature relation for the base-emitter voltage of a BJT:

$$V_{BE}(T) = V_r + \gamma_r T + f_r(T), \text{ where } V_r = V_{GOr} + (\eta - \theta_r)\left(\frac{kT_r}{q}\right), \text{ and } \gamma_r = -\left[\frac{V_r - V_{BE}(T_r)}{T_r}\right]$$

Here $f_r(T)$ represents higher order terms that are not canceled by the V_t term, and T_r represents the reference temperature for the analysis. A detailed analysis of the components of $f_r(T)$ is done in [2], and is not of critical importance in the present discussion. By setting $KV_t = -\gamma_r(T)$, we can make V_{BE} independent of T to the first order.

Now, we compute the reference voltage generated by this Brokaw cell. V_{R1}, assuming an ideal opamp with large gain and low offset is given by:

$$V_{R1} = V_{BE2} - V_{BE1} = V_t ln\left[\frac{I_2 A_{E1}}{I_1 A_{E2}}\right]$$

Notice here that the voltage difference is created by sizing the emitter areas of the two transistors Q_1 and Q_2 differently. The opamp also forces the following relation between the currents in the two legs : $I_1 R_2 = I_2 R_3$.

Combining the last two relations, we can compute V_{R1}. After some algebra, we get:

$$V_{REF} = V_r + \gamma_r T + f_r T + \left(\frac{R_2}{R_1}\right)\left(\frac{kT}{q}\right) ln\left(\frac{R_2 A_{E1}}{R_1 A_{E2}}\right)$$

Here the coefficient $\left(\frac{R_2}{R_1}\right)\left(\frac{kT}{q}\right) ln\left(\frac{R_2 A_{E1}}{R_1 A_{E2}}\right)$ can be tweaked by changing $\left(\frac{R_2}{R_1}\right)$, to make the coefficient match γ_r.

It is important to remember that this entire analysis depends upon the Taylor series expansion of V_{BE} about a reference temperature. Thus, this cancellation is most effective near the reference temperature T_r used for the Taylor series expansion. For the lowest temperature dependence, we would like T_r to be in the temperature range of interest in the application, typically 0-125°C. However, process variations can lead to a change in γ_r and $f_r(T)$. This means that the effective coefficient needed to achieve perfect cancellation around T_r will be different for different chips. Therefore, for high accuracy first order bandgap references, it is important to be able to change the resistor ratio $\left(\frac{R_2}{R_1}\right)$ individually for every chip. Using this technique (and calibrating $\left(\frac{R_2}{R_1}\right)$ for every

chip individually), one can achieve temperature coefficients in the tens of ppm/°C.

Even better performance can be achieved by using higher-order bandgap references[3],[4],[5].
These reference topologies attempt to cancel out some of the terms in $f_r(T)$, in addition
to the $\gamma_r T$ term. A common technique is to cancel out the $\gamma_r T$ term as in the first order
reference design, and then modulate the transistor bias current $I_c(T)$ in a fashion as to
cancel the higher-order terms in $f_r(T)$[2],[3]. The actual shape of $f_r(T)$ is complicated,
with many polynomial and exponential terms. In practice however, it is found that a
satisfactory form of $I_c(T)$ for the purposes of minimizing temperature variation is:

$I_c(T) = I_c(T_r) \left(\frac{T}{T_r}\right)^b$, where b is an appropriate constant[2].
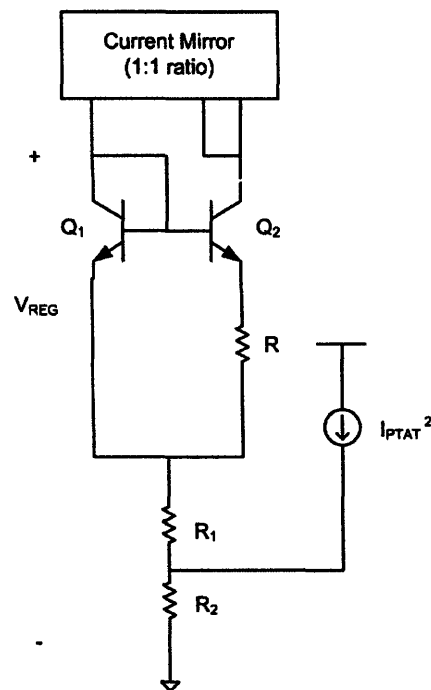
Figure 2.4: A Second Order Bandgap Circuit using I_{PTAT^2} currents

One simple method of implementing a higher order reference is discussed by Rincon-
Mora in[3], and shown in figure 2.4. It involves generating a current proportional to

T^2, also called a $(PTAT)^2$ current. The reference [3] describes some simple ways of generating such a current. This current is then used to bias the emitter degeneration resistors, giving rise to a reference voltage:

$V_{REF}(T) = V_{BE1}(T) + \left(\frac{V_t}{R}\right) ln\left(\frac{A_{E1}}{A_{E2}}\right)(R_1 + R_2) + I_{PTAT^2}R_2$. By appropriate choice of resistor values R,R_1 and R_2, we can eliminate both first and second order temperature dependence from V_{REF}. As I^2_{PTAT} is typically generated using a resistor, all the voltage terms added to $V_{REF}(T)$ are dependent on ratios of resistors, which track very well with process and temperature on IC's.

An important fact to remember is that most bandgap circuits have two or more stable operating points. Therefore, special circuitry (called the startup circuit) has to be used to ensure that the bandgap reference reaches the proper operating point. Startup circuits are usually customized to the particular bandgap circuit they are meant to work with, but most of them follow a similar pattern. In general, the circuit attempts to inject some current into the bandgap cell to pull it out of the (stable) zero current operating point. Once the bandgap reaches proper operation, it turns on another device, which shuts off the startup circuit. The major aim of the startup circuit is to turn the circuit on reliably in all possible conditions, and consume as little current as possible during normal operation of the bandgap circuit. The references [3],[6] describe many possible topologies for consideration. In chapters 4 and 5, I discuss the operation of the startup circuits that were used in this block.

There are many other techniques that have been developed to improve the temperature performance of bandgap references. These are too numerous to be discussed here, but include exponential curvature compensation[4], beta compensation[7], exploiting the resistors' temperature coefficient[9]. The reference section lists works that discuss these techniques in detail.

2.4 Reason for Design Choice

As mentioned earlier, this design uses a bandgap based reference rather than a zener based reference. In section 2.2, we mentioned that zener based references had very good temperature and noise performance, with the best ones going well below 1 ppm/$^\circ$C. Compared to zener references, the temperature and noise performance of basic bandgap circuits is much worse. However, zener references have some major disadvantages, that have led us to use bandgap-based references instead.

The biggest reason for this decision was power consumption. Zener based circuits cannot operate with good performance under a few milliamps of current. As the current budget for the bandgap is well under a milliamp, this would pose significant problems in using a zener-based circuit. The high power supply requirements for most zeners (about 5.6V) also creates problems, as modern Li-Ion batteries can drop as low as 2.9V before shutting off(this is done to extract the maximum possible energy from the battery).Finally, standard CMOS processes are typically incompatible with building an high quality buried zener diode. Therefore, using a zener-based reference would force us to use an external component, which increases cost and board area. Due to all these reasons, we decided to use bandgap-based references in the current design.

Chapter 3

System-Level Details

3.1　Introduction

The previous chapters have laid out the motivation for this thesis, and background information about reference designs. In this chapter, we discuss the important specifications for this particular implementation, and give a system-level overview of the entire reference block.

Section 3.2 starts off with an overview of the major sub-systems on the chip that use the bandgap voltage. It then describes the critical specifications for the overall block, including the power supply rejection(PSRR) and noise requirements.

Section 3.3 contains a block-diagram description of the overall system, and describes how the overall specifications from section 3.2 are split across each of the sub-blocks.

3.2   Key Specifications

3.2.1   Available Process and Supply Options

The reference block makes use of various supplies that are available to the chip. The primary supply is directly taken from the cellphone battery. This supply can vary from about 3.1 V to 5.5 V depending on the type and condition of the battery. Using this supply poses some issues, as the supply goes higher than the maximum voltage limits of the transistors in the process. The transistors cannot handle more than 3.63 V across any two junctions, necessitating the need for protection circuitry for any sub-blocks using this supply.

To facilitate the running of logic, there is a 3V supply called Vperm. This supply is turned on all the time, therefore it is available during both the low-power and high power mode. However, it cannot supply too much current. Therefore only essential logic that cannot be powered off other supplies easily are connected to it. During the high power mode, there is another 2.85V supply called V_{ana} available. This supply can deliver a lot more current than Vperm, and therefore it can be used for sub-blocks that only need to be on during high power mode.

The process used is a 65 nm low power process, with high voltage transistors capable of supporting up to 3.63 V available.

3.2.2   Chip-Level Overview

The reference block supplies a reference to various sub systems in the chip. There are five major subsystems that depend on the reference block for reference voltages and bias currents. These are, the DC/DC converter, the LDO(Load Dropout Regulator), the AuxADC(Auxiliary ADC), Battery Charger and the Audio sub system. The DC/DC

converter is used to generate the core voltage for the digital section of the chip from V_{BATT}. The LDO module supplies a regulated power supply to various parts of the chip. The AuxADC is used to interface with any externally connected auxiliary devices. The battery charger is used to charge the Li-Ion batteries for regular use. Finally, the Audio sub-system is used during calls, and to play songs, ring tones,etc. They can broadly be grouped into two categories. The first category is "alltime" subsystems, which means that these sub-systems have to be on most the time, including the standby modes. The DC/DC Converter and the LDO modulator fall under this category. The other three blocks are only switched on when needed, and are powered off otherwise.

3.2.3 Modes of Operation

The reference block needs to operate in two possible modes - low power and high power. The low-power mode is on when the cellphone is in standby. This mode is used to supply a reference voltage to blocks such as the DC/DC converter and dropout regulators which are on most of the time the cellphone is powered on. As the low-power mode contributes to standby power, emphasis is laid on having low power consumption. To accommodate this, the specifications for the low power mode are not as stringent as the high power mode.

The lower power levels do not permit us to run certain blocks, like the audio subsystem, auxiliary ADC and the battery charger. In order to turn these blocks on, we need to enable the high power mode in the reference block. In this mode, the block provides a higher-quality reference voltage (where higher quality implies that certain performance metrics like output noise and PSRR are better). On average, the time the high-power mode is enabled is very small compared to the time the phone is on standby. In addition, both the audio sub-system and the battery charger consume significantly more power

than the reference in high power mode. Therefore, the larger power consumption of the bandgap in this mode does not affect the standby time of the cellphone significantly.

3.2.4 Power Supply Rejection Ratio

This is one of the most important specifications for this block. The reference block runs off the same power supply as a power amplifier (PA) used for transmission in a cellular phone. As the power amplifier consumes a significant portion of the power budget of a cellphone, it is only turned on when required to save power.

Unfortunately, this causes the power supply of the system to droop when the PA is switched on, and to recover when the PA is switched off. This causes large transients (as high as 300 mV) in the power supply that need to be suppressed by the reference generator. As the PA turns on and off in a periodic manner, it adds ripple on the power supply at the switching frequency and its harmonics. The reference has to significantly attenuate the effect of this ripple on its output voltage.

The audio sub-circuit has the most critical requirement in this aspect. The audio sub-block uses the reference voltage to convert digital bits to audio, and any non-ideality in the reference voltage shows up directly in the audio stream. The human ear is significantly more sensitive to tones compared to white noise, therefore suppressing this periodic ripple on the power supply is very important.

The audio sub-block aims for an SNR of about 90 dB. To ensure that the tonal components of noise (from the power supply) are suppressed much further, we need to have a PSRR > 90dB. For a further suppression by a factor of 10, we need at least 20 dB of additional PSRR. This results in an overall PSRR figure of 110 dB at the audio output. The frame rate of a GSM PA (the rate at which it switches on and off) is 217 Hz. Therefore, maximum ripple is seen at this frequency. The ripple sidebands fall off

as 1/f with frequency, reducing the PSRR requirement at higher frequencies. As we care about the audio band (20-20kHz), this means we need >110 dB of rejection at 217 Hz, and > 70 dB of rejection at 20k. As there are always coupling mechanisms that cannot be simulated, the specification conservatively aims at 120 dB of PSRR at 217 Hz, and 80 dB of PSRR at 20 kHz. These numbers are much higher than normally required, and are a critical challenge in this design.

3.2.5   Output Variation

The DC variation in the bandgap voltage level with process and supply changes is another important specification. The block most sensitive to DC variation is the battery charger. It cannot tolerate more than +/- 1% variation in its output voltage (including DC and temperature variation) . Therefore, it was decided to limit the DC variation for the reference voltage going to the battery charger to less than +/-0.4%, which leaves plenty of margin for temperature variation and offsets internal to the battery charger. To attain this level of DC precision with changing process corners, it was necessary to trim the voltage reference for the battery charger. The details of this trimming process are discussed in section 6.1.1.In low power mode, the DC precision required is much lower; the DC accuracy level required is about +/- 1%.

In addition, after the initial transient, the bandgap voltage must not vary by more than +/- 5%, for any steps, transients or glitches on the power supply.

3.2.6   Noise

The audio sub-block is the primary determiner of this specification. As the audio bandwidth of interest is 20-20kHz, the tightest specification is to have less than 3 μV of noise, integrated in the 20-20k bandwidth. In addition, some of the noise from higher

frequencies is downconverted into the audio range during audio processing. For this reason, the noise in the 20-1GHz range is limited to about 100 μV.

3.2.7 Power Consumption

As mentioned previously, the block needs to have two power modes - low power and high power. In the low power mode, the block is expected to consume less than 30 μA of power, drawn from the V_{BATT} supply. This does not include the currents supplied to other blocks. During high power mode, the power budget goes up to around 300 μA drawn from the V_{BATT} supply for the fine bandgap. This does not include the power drawn for refgen_i, that supplies currents to other blocks, and can be turned off independently.

3.2.8 Temperature Variation

Again, as mentioned in section 3.2.5, the battery charger sets this specification. The target is a temperature variation of less than +/-0.4% for a temperature range of -40°C to 125°C. Combined with the the +/-0.4% specification for the DC variation, it leaves us enough margin for the overall +/- 1% requirement. In addition, the battery charging circuit is only required to function in the 0°C to 85°C range, giving us additional margin.

3.3 Overall Architecture

Figure 3.1 gives a block-level description of the structure of the bandgap block. It is a two-stage design, with the coarse bandgap being used to supply the reference in the low power mode, and the fine bandgap being used in the high power mode. In addition to the two main bandgap circuits, there are two other important blocks - the refgen_i
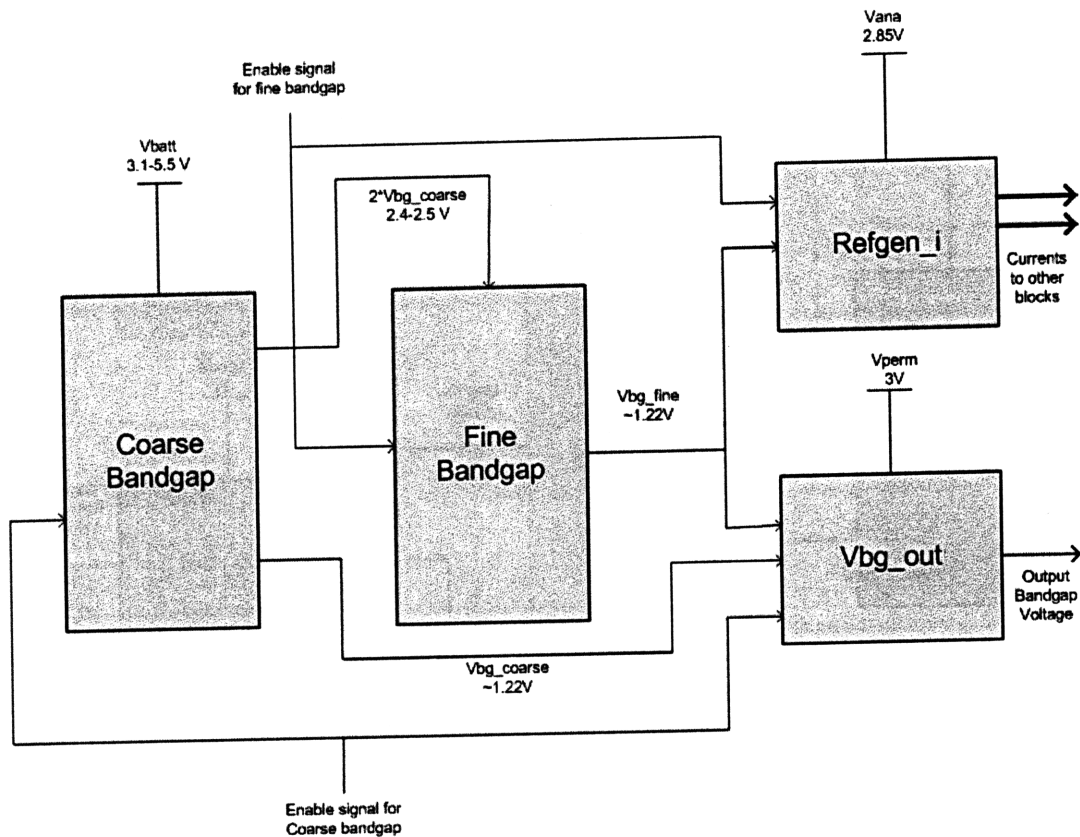
Figure 3.1: Block Diagram of Bandgap Circuit

and the vbg_out block. The following subsections analyze the operation of the circuit in both low and high power mode.

3.3.1 Operation in low power mode

Figure 3.2 shows the reference block in low power mode. When the reference block receives an enable signal to power itself on, it turns on the coarse bandgap and vbg_out sub blocks. The fine bandgap, and the refgen_i blocks are kept powered down to save power. The coarse bandgap then supplies the reference voltage to the vbg_out block.
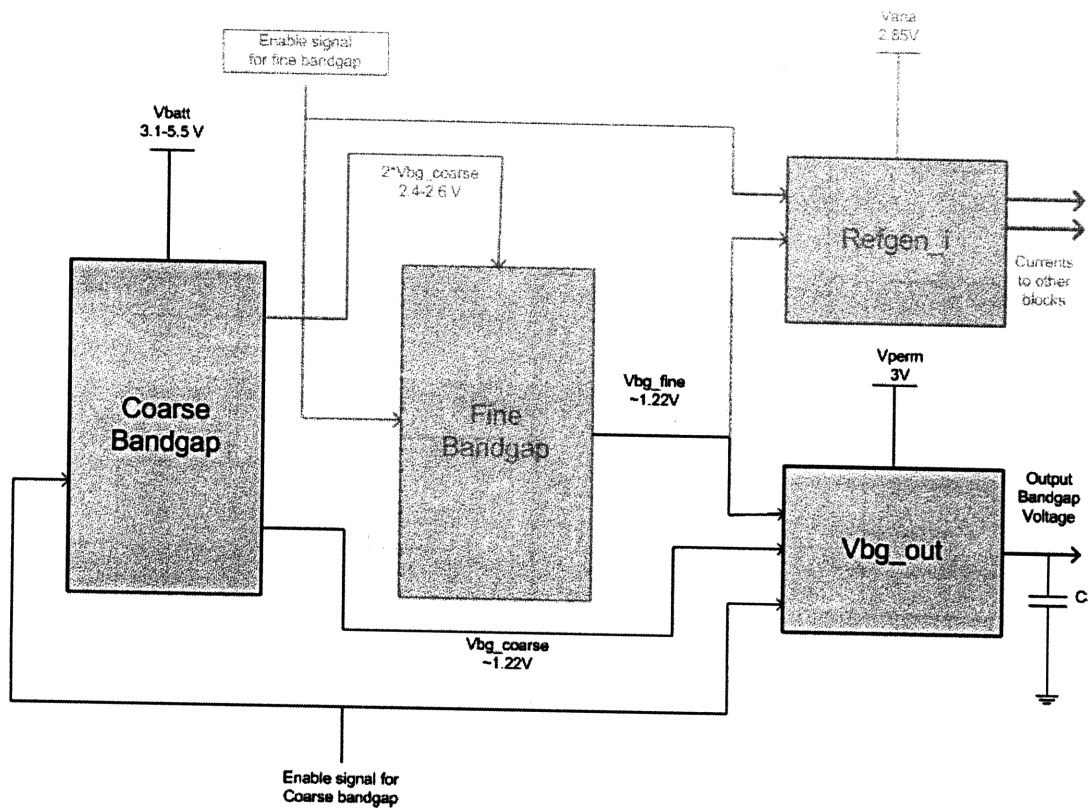
Figure 3.2: Block Diagram of Bandgap Circuit in Low Power Mode

The coarse bandgap block also has a small voltage-to-current conversion circuit built in. This current is then mirrored and set out to some other blocks on in low power mode, including the DC/DC converter and the regulators. This current is drawn from another supply (Vperm) to prevent other blocks from potentially being exposed to voltages > 3.63V.

The vbg_out block has two major roles in the low power mode. It includes an hysteretic comparator that compares the bandgap voltage to Vperm/3. Once the coarse bandgap voltage rises above this threshold, it sends out an OK signal to the rest of the chip, which allows other blocks to begin using this voltage.

30

The other major role of the vbg_out block is to act as a fast charging circuit. The output of the vbg_out block is connected to a large external cap, to provide external filtering as well as reduce the amount of thermal noise on the reference voltage. Therefore, the turn-on time of the bandgap block is slew-rate limited by the coarse bandgap. As the current levels in the coarse bandgap circuit are low, this caused the output node to take a long time in charging up. Therefore, vbg_out has a fast charging circuit that provides extra charging current while the output of vbg_out is below the output of the coarse bandgap, and turns itself off automatically.

3.3.2 Operation in High Power Mode

When the bandgap is told to switch to the high power mode, both the blocks that were switched off in the low power mode (the fine bandgap and the refgen_i) are switched on.

When the fine bandgap is turned on, it draws current through the output stage of the coarse bandgap. The output stage of the coarse bandgap is built to act as an rough regulator for the fine bandgap. The output voltage is regulated to be around twice the bandgap voltage. Therefore, the fine bandgap sees a power supply of about 2.5V.

By using the output stage of the coarse bandgap as a supply for the fine bandgap, we achieve two major gains. By limiting the supply voltage for the fine bandgap to 2.5V, we do not have to worry about any protection issues in the fine bandgap. If we had instead used V_{batt} directly, we would have had to add protection circuitry to the fine bandgap as well. This would have reduced performance as well as increased the complexity of the circuit.

However, the most important reason for the two stage design was that it made the PSRR requirement manageable. In section 3.2.4, we saw the high PSRR requirement

(120 dB at 217 Hz, and 80 dB at 20k) for the system. The two stage design allows us to split this requirement across 2 circuits, rather than try and achieve it in a single stage. The different power and accuracy requirements of the low power and high power almost necessitates using different circuits for each mode. By isolating the fine bandgap from the power supply using the coarse bandgap, we can harness this requirement to help us achieve high PSRR. The major disadvantage of this technique is that it increases the complexity and the power consumption of the coarse bandgap somewhat, as now it has to act as an regulator in addition to a bandgap circuit.

The implementation of the fine bandgap is very close to a standard Brokaw cell, with some changes in the biasing structure. In addition, the bandgap voltage is trimmed to ensure better accuracy. By trimming, we can calibrate the chip to operate at the voltage at which it has the least temperature variation.

The refgen_i block is used to convert the bandgap voltage into a current. This current is used both as a bias current and used to recreate the bandgap voltage in different blocks (by running this current through a resistor identical to the one in the refgen_i block). This allows us to transfer the bandgap voltage to different blocks without having to worry about picking up interference on the line. The refgen_i block has a small-regulator type circuit, using the bandgap voltage as its reference. The regulator is loaded with a resistor (that is trimmable to change the effective resistance). This current is then mirrored out to the battery charger block, which has the tightest accuracy requirements. The trimmable resistor is controlled by trying to match the output of a battery charger to a golden reference.

This two stage trimming process allows us to get both good temperature performance and DC accuracy. Many applications (like the battery charger) want a constant voltage with small temperature and process variation. However, the DC voltage for the best

32

temperature performance is not necessarily the exact DC voltage required by the charger, and changes with process corners. With this two step trimming process, the first step reduces temperature variation by trimming to achieve the best temperature coefficient. The second trimming step takes this voltage, and adjusts the DC level independent of temperature variation (this approximation is valid to the extent that resistor ratios track well across temperature). This allows us to get an accurate DC voltage as well as low temperature variation.

When the high power mode is enabled, the vbg_out block switches off the connection to the coarse bandgap, so the output cap is connected to the fine bandgap. Some filtering is employed to make the transition gradual. The coarse bandgap is still used to provide reference currents to blocks like the LDO and the DC/DC converter.

Chapter 4

The Coarse Bandgap

In this chapter, we discuss the implementation details of the coarse bandgap.

Section 4.1 describes the overall operation of the coarse bandgap, and identifies the various sub-circuits that are discussed at length in the latter sections.

Section 4.2 talks about the core bandgap-generating section of the circuit, along with the output stage that biases the bipolar transistors and acts as an regulator in high power mode. We also talk about the techniques used to protect the transistors from seeing excessive voltage across their terminals.

Section 4.3 discusses the operation of the startup circuit in detail. It also describes the voltage stress issues associated with the power-down devices during startup, and how those issues were managed.

Section 4.4 describes how the bandgap voltage is used to generate a bias current that is then mirrored to other sections of the chip.
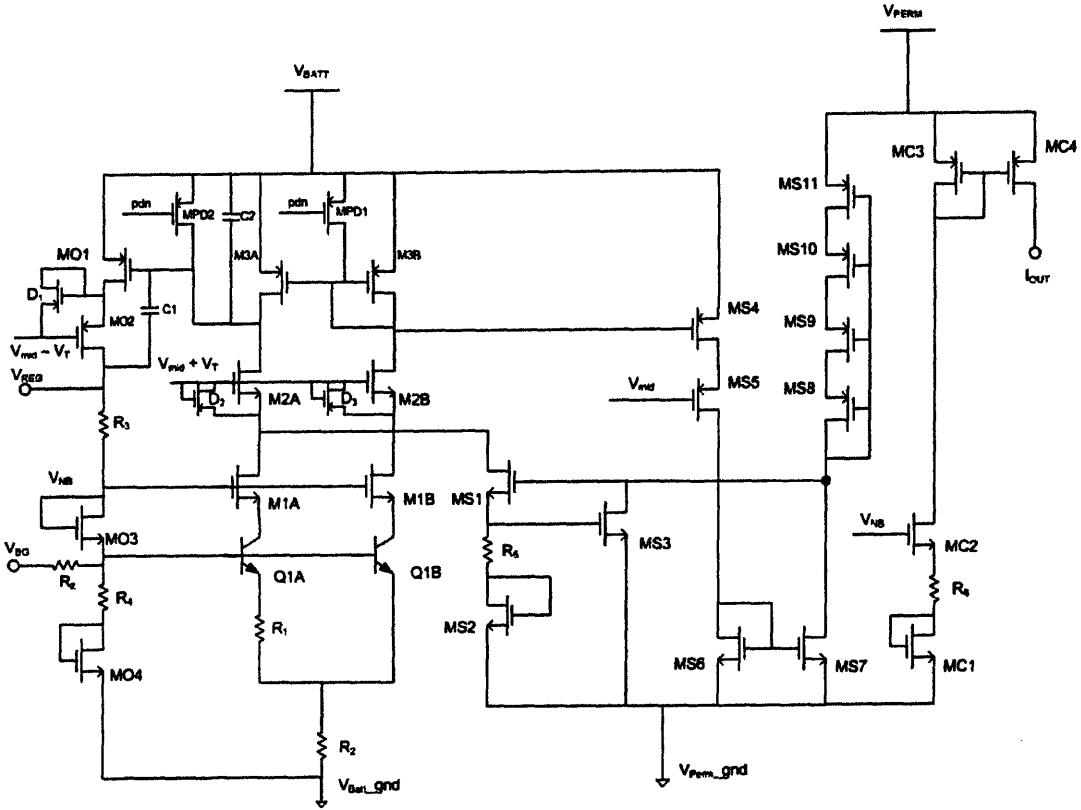
Figure 4.1: The Circuit Used to Implement the Coarse Bandgap

4.1 Coarse Bandgap Overview

Figure 4.1 shows the schematic implementation of the coarse bandgap. It uses V_{BATT} as the primary supply, using it to run the bandgap core, output stage and the startup circuit. V_{PERM} is used to draw the output currents sent to other blocks.

The core section uses a pair of differently sized BJT's (Q1A and Q1B) and NMOS devices (M3A and M3B) to force a voltage across R_1 that is proportional to V_{th}, the thermal voltage. The resistor R_1 then sets the current that flows through M3A and M3B. R_2 carries the sum of both those currents, and can be trimmed to adjust its resistance.

Therefore, the output voltage is set by $V_{BE(Q1B)} + V_{R_2}$. As V_{R_2} can be adjusted, and is directly proportional to temperature, the resulting voltage is a first order temperature compensated bandgap.

The output stage (defined by MO1,MO2,MO3,MO4,R_3 and R_4) helps supply the base currents for Q1A and Q1B. In addition, MO1 and MO2 are sized up to ensure that they can support the current load of the fine bandgap stage. The capacitor C1 is a miller compensation capacitor that ensures stability in the system. The transistors M1A and M1B improve stability in the system by removing the miller effect contributed by the C_{bc} of the bipolar transistors. C2 is used to prevent stress during the coarse bandgap startup; this will be discussed in more detail in section 4.2.

There are also three voltage levels that have not been previously mentioned - V_{mid}, $V_{mid} - V_T$ and $V_{mid} + V_T$. These voltages are generated by another block on the chip, and are based of V_{BATT}. Section 4.2 discusses how these voltages are used to prevent device stress in the circuit. The unique stress requirements of the design also affect the way powerdown devices are used (and the pdn signal), which is covered in section 4.3. The diodes D1-D3 are used for protection during the power-down mode.

The devices MS1-MS11 are used to start up the bandgap circuit out of a stable, zero current state. The circuit consumes very little static current, and disconnects itself automatically once the circuit is near proper operation. The operation of this startup circuit is explained in section 4.3.

Finally, the bandgap voltage is then placed across a scaled resistor/diode combination (R_5 and MS2), which are a scaled version of the devices in the output stage (R_4 and MO4). This generates a current $= V_{BG}/R_5$, which is then mirrored to other blocks. By using a resistor/diode connected transistor combination, we can save area (because we no longer have to use a larger resistor for the same current consumption). To recreate the

36

bandgap voltage, the current is simply passed through a resistor/diode stack identical to R_6 and MC2.Ignoring mismatch, the voltage reproduced is exactly the bandgap voltage.

4.2   Bandgap Core and Output Stage
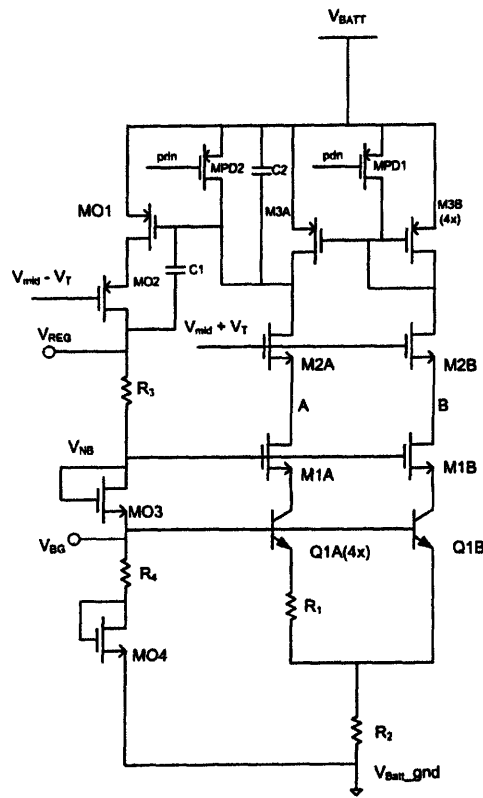
Figure 4.2: The Circuit Used to Implement the Coarse Bandgap

Figure 4.2 shows the bandgap core and the output stage. The BJT Q1A is 4 times larger than Q1B, while M3B is also 4 times larger than M4A. To determine the voltage V_{R1}, we ignore the protection transistors for now. Then: $V_{R1} = V_{BE,Q1A} - V_{BE,Q1B} = \left(\frac{kT}{q}\right) ln \left(\frac{I_{C,Q1A}}{I_{S,Q1A}}\right) - \left(\frac{kT}{q}\right) ln \left(\frac{I_{C,Q1B}}{I_{S,Q1B}}\right)$, where I_C indicates collector current and I_S is the

bipolar saturation current when $V_{BE} = 0$.

Now, $I_{C,Q1B} = 4I_{C,Q1A}$ (because the current mirror M3B is 4 times larger than M3A). Also, $I_{S,Q1A)} = 4I_{S,Q1B}$ (as Q1A is 4 times larger than Q1B). Simplifying the expression for V_{R_1}, and plugging in the previous two relations, we get $V_{R_1} = \left(\frac{kT}{q}\right) ln(16)$. Therefore, using this structure, we have generated a voltage that is only dependent on temperature (and the ratios of the transistors). However, to generate a bandgap voltage, we cannot simply add this voltage to the V_{BE} of a BJT. We need a way to have a multiplicative constant that can be changed, to ensure we can always reach the optimum temperature coefficient even with process variation. This is the role played by R_2. As the current flowing through R_2 is simply the sum of currents flowing through Q1A and Q1B, which is fixed by the resistor R_1, we have:

$I_{R2} = I_{C,Q1A} + I_{C,Q1B} = 5I_{C,Q1A} = 5\left(\frac{V_{R1}}{R1}\right)$. Therefore, $V_{Bandgap} = V_{BE,Q2A} + V_{R1} + V_{R2} = V_{BE,Q2A} + 5\left(\frac{R_2}{R_1}\right)\left(\frac{kT}{q}\right) ln(16)$. By making R_2 trimmable (i.e its value can be changed), we can trim the overall bandgap voltage to achieve the optimum temperature coefficient. Figure 4.3 shows the temperature variation of a properly trimmed coarse bandgap output voltage.

This bandgap voltage is then placed across a resistor/diode combination (R4/MO4). Then, another (almost) identical stack is placed on top of it(M03/R3). Therefore, the voltage at the V_{REG} is approximately $2V_{BG}$. The node V_{BG} is connected in negative feedback with the base of the bipolar transistors. In order to support the bandgap voltage, and the current flowing through the bipolar transistors, the output leg supplies current to the base of the transistors. Because of the low β of the transistors, the base current is comparable to the collector current. Therefore, there is a difference between the current flowing through R_3 and R_4. To account for this, R_3 is sized to be smaller than R_4.
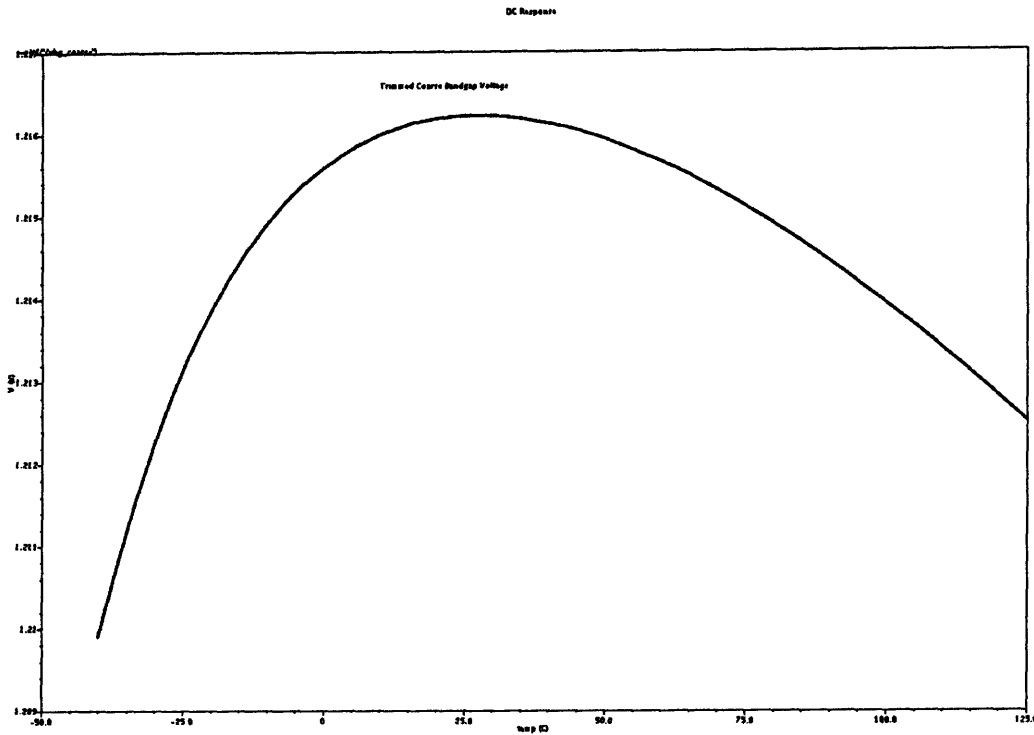
Figure 4.3: Temperature variation of the Coarse Bandgap Circuit

By supplying the base current from the output stage(rather than diode connecting the BJT's), we reduce the impact of the low beta of the transistors on V_{R1}. Had we diode connected the transistors to one particular leg, there would have been a significant deviation of the currents through M3A and M3B (from the ideal ratio of 4:1), because M3A would have to supply all of the base current. In the present architecture, the base current also splits in the correct ratio between the two legs, thereby keeping the current ratios near 4:1. This also allows us to tolerate lower power supply voltages (as $V_{CE,Sat}$ is less than V_{BE} for a bipolar transistor).
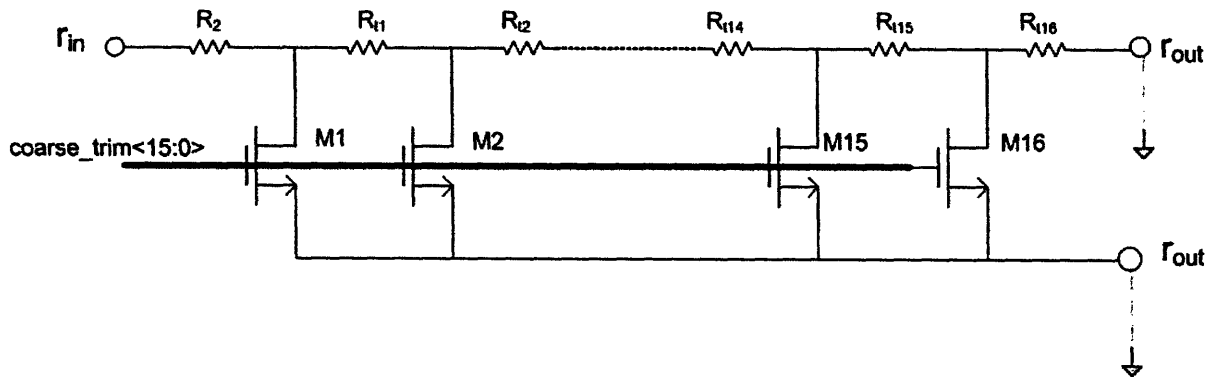
39

Figure 4.4: Schematic of the Trimmable resistor

4.2.1 Resistor Trimming

Figure 4.4 shows the implementation of the trimmable resistor R_2 from Figure 4.2. There are 16 small resistors in series, with a switch connected after each resistor. The sources of all these switches are shorted. This shorted node is connected to ground. In order to set the effective resistance, we turn one switch on, while leaving the rest of the switches off. This lets us bypass the resistors that follow the switch. To change the resistance, we simply turn on an different switch. For example, turning M1 on gives us a resistance of R_2, while turning on M2 instead gives us $R_2 + R_{t1}$. The switch control signals are supplied as a 4 bit binary number. To control the 16 switches, we use a binary to thermometer code converter, which is not shown in the figure. Figure 4.5 shows how we can vary the output voltage (and bring it to the optimum value for low temperature variation) by changing the trim code.

4.2.2 Transistor Protection Strategy

The voltages V_{mid}, $V_{mid} + V_T$ and $V_{mid} - V_T$ are used to protect transistors from seeing excessive voltage across any two terminals. V_{mid} is $\left(\frac{V_{BATT}}{2}\right)$, and V_T is the thresh-
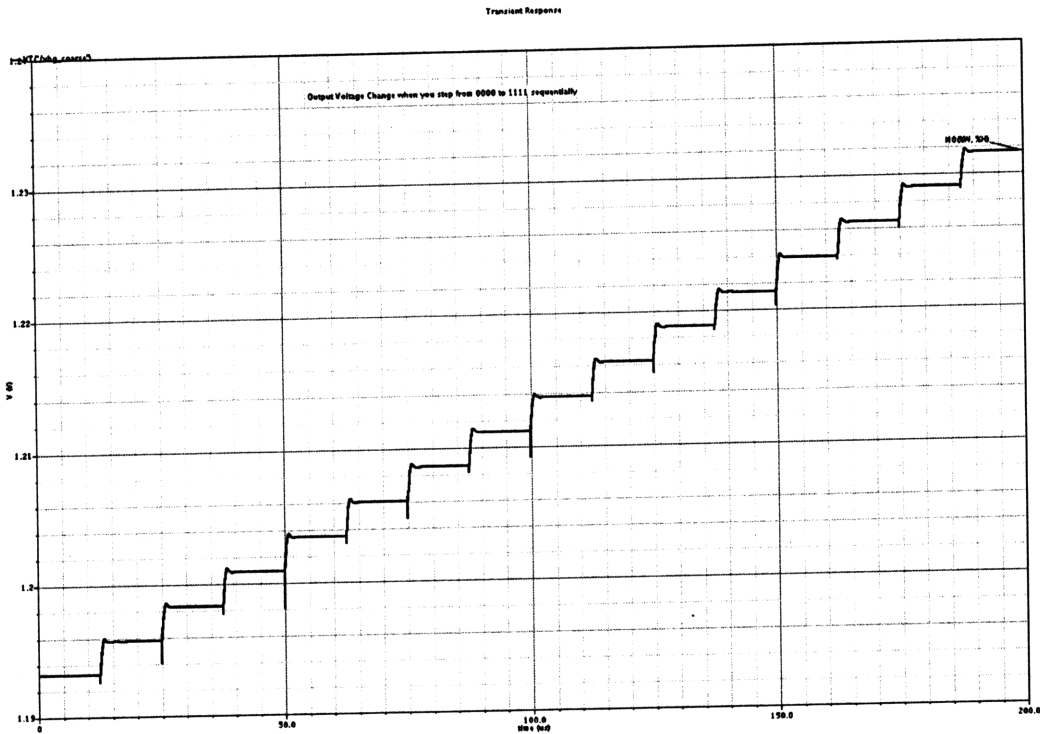
40

Figure 4.5: Output Change with Trim Code stepping from 0000 to 1111 sequentially

old voltage drop of a representative transistor (it is designed to be around 0.7V). The protection strategy on this chip relies on using this voltage to bias protection transistors. By biasing PMOS transistors using $V_{mid} - V_T$ and NMOS transistors using $V_{mid} + V_T$, we can force the source nodes of these transistors to be around V_{mid}. As the maximum value of V_{batt} is around 5.5V, both V_{mid} and $V_{BATT} - V_{mid}$ are below the voltage stress limits of the transistors. Therefore, all the transistors on that particular leg are protected.

For example, take M2A and M2B from figure 4.1. When the circuit is fully operational, the source of the two devices (nodes A and B respectively) are at V_{mid}. Thus all the devices below A and B are protected, as V_{mid} is lower than the breakdown voltage for the process. Similarly, all the devices above A and B never see more than $V_{BATT} - V_{mid}$,

which again is safe for the process.

When the block is turned off, the only current flowing through these transistors is leakage current because of parasitic diodes. Therefore, floating nodes either float up to V_{BATT} or down to ground, depending on relative strength of the parasitic diodes. This may cause stress issues on some of the transistors. To prevent this from happening, we have diode connected transistors that turn on and clamp a node when it falls too low or too high(D1,D2,D3).

4.2.3 Loop Stability

For stability analysis, the coarse bandgap looks like a two stage amplifier, with the core making up the first stage, and the output leg making up the second stage.

The primary method of compensation is miller compensation, using C_m. As the output transistor is sized to carry a lot of current, the gain across the stage is significant. Therefore, the dominant pole of the system is at $\frac{1}{(g_m R_1 C_m + C_{gs})R_o}$, where R_o is the output impedance of the bandgap core, and C_{gs} is the gate to source capacitance of the output transistor.

The largest parasitic pole is at the input of the bandgap core, and the capacitance is contributed by the bipolar pair. The frequency of this pole is $C_p R_1$. Therefore, in order to have good phase margin, we need to either reduce C_p or R_1. Reducing C_p involves using a smaller bipolar device. As the bipolar devices in a CMOS process have poor performance, this is not always a practical solution. Therefore, improving stability in the system involves lowering R_1, which increases current. For this reason, the current in the output stage is set by the stability requirements of the circuit.

One of the reasons to cascode the bandgap core using M1A and M1B is to improve stability(Figure 4.2). By adding that cascode pair, we eliminate the parasitic pole
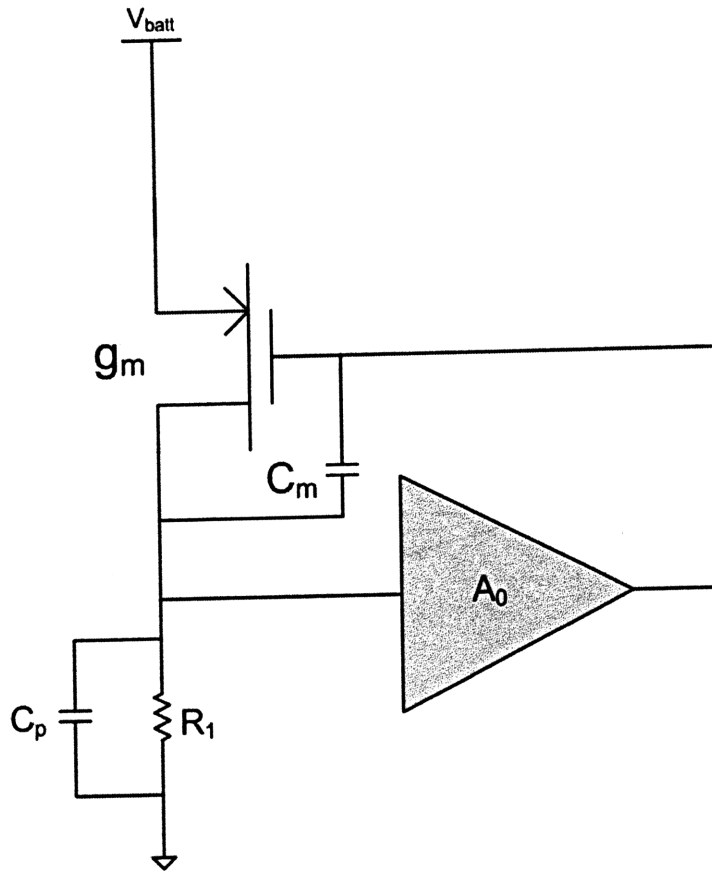
42

Figure 4.6: Stability model of the Coarse Bandgap

associated with the C_{bc} of the bipolar devices (which would have shown up as additional miller capacitance in parallel with C_p in Figure 4.6). It also helps improve PSRR, which is discussed in another section.

Finally, there is also a large impedance R_z in series with the output of the bandgap. This resistor comes into play when the bandgap is connected to a large external capacitor (to improve noise and PSRR performance). Without the resistor, this simply adds a low frequency pole at $\frac{1}{C_{out}R_1}$ that seriously compromises the stability of the system. Adding the large resistor in series adds a low frequency zero (at $\frac{1}{C_{out}R_z}$) that changes in

proportion with the additional pole at $\frac{1}{C_{out}(R_1+R_z)}$, and prevents loss of phase margin.
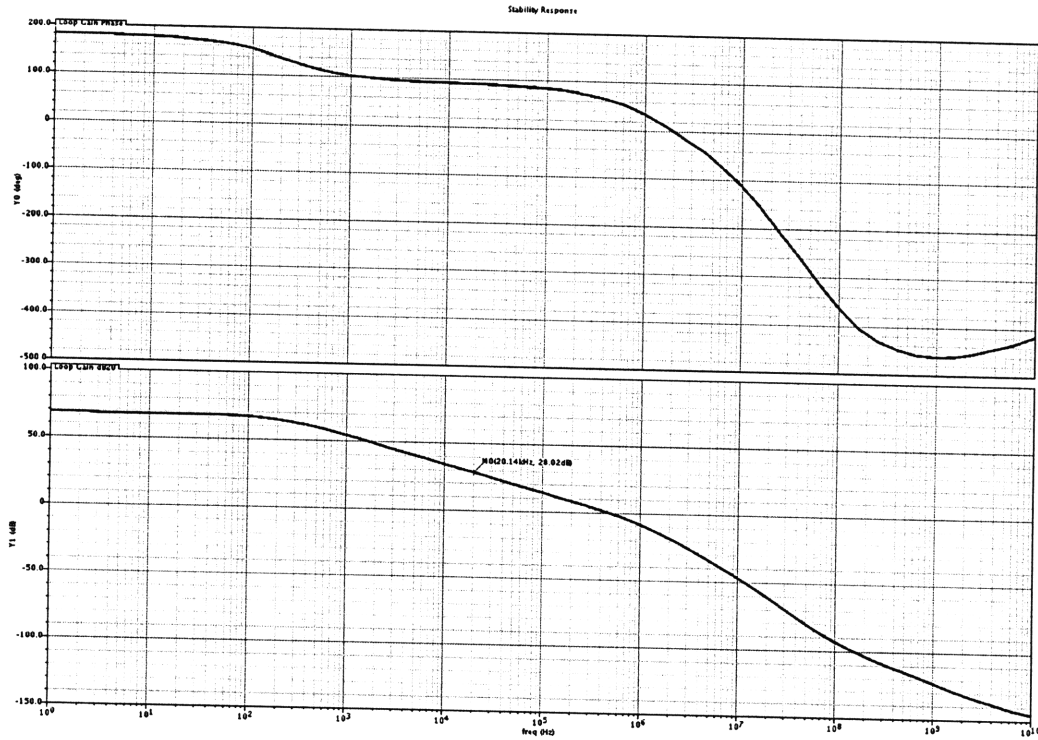
Figure 4.7: Gain and Phase Stability Plots for the Coarse Bandgap

4.2.4 PSRR in Coarse Bandgap

The primary method of suppressing power supply variation is the gain of the feedback loop. DC suppression is approximately equal to $A_o g_m R_1$, with PSRR at 20 kHz (the highest frequency of interest) being around $\frac{A_o g_m R_1}{m(g_m R_1+1)r_o}$. Therefore, the PSRR performance of the bandgap is essentially just the inverse of the loop gain, which can be seen in corresponding plots Figure 4.7 and 4.8.

Figure 4.8: Power Supply Rejection of the Coarse Bandgap Voltage

44

4.3 Startup Circuit and Power Up
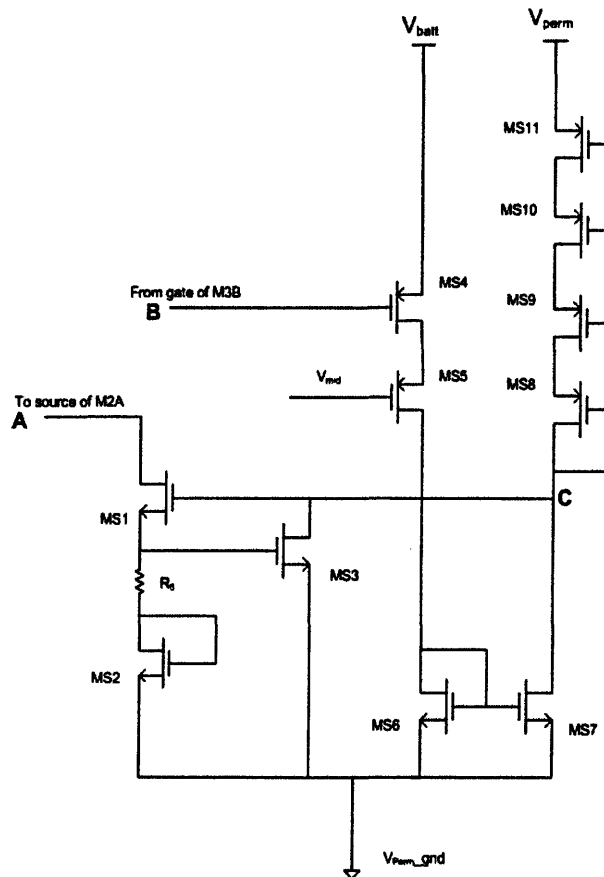
Figure 4.9: Startup Circuit for Coarse bandgap

Figure 4.9 shows the startup circuit used in the coarse bandgap. To see how it works , let us assume the bandgap is initially off. This implies that node B is high (at V_{batt}), and node A is low. Therefore, there is no current in the leg containing MS6, and keeps the V_{gs} on MS7 close to zero. The transistors MS8-MS11 form one (long) PMOS transistor. Therefore, with MS7 off, MS8-MS11 pull the gate of MS1 high. This turns MS1 on, and forces current through the bandgap core, forcing the bandgap circuit on. Once the circuit reaches normal operation, the gate of M3B (node A in the figure) starts

45

to go below V_{batt}. As MS4 is matched to M3B, an scaled amount of the current flowing in M3B starts to flow in MS4, and consequently MS6. This raises the gate voltage of MS7, and turns it on. At this point, MS7 and MS8-MS11 are competing for control of the gate of MS3 (node C). Now, the circuit is designed so that MS7 is much stronger than MS8-MS11. Therefore, when the gate of MS7 goes high, it forces node C low. This turns off MS1, which was supplying the startup current to the bandgap core. As long as the bandgap circuit is turned on, with proper current flowing through it, MS7 keeps the startup current off. However, if the bandgap circuit is not on, MS7 is off, and the transistors MS8-MS11 turn on a startup current. The resistor R_5 and MS2 control the amount of startup current provided to the circuit; their sizing is empirical. MS5 is used to protect the transistors from stress. The transistors MS7, and MS8-MS11 are powered from V_{Perm} rather than V_{batt}. This is to reduce the need to manage stress concerns. As these transistors will not draw current during normal operation, there is not much loading on the supply. All the DC current is drawn directly from V_{batt}. Figure 4.10 shows the bandgap voltage reaching its steady state value after the block gets its enable signal.

In addition to normal startup, the power-down devices require some attention. In Figure 4.1, MPD1 and MPD2 are the power-down devices, designed to pull the gates of M3B and MO1 high when the circuit is switched off. They are controlled by the power-down signal pdn. Now, pdn is controlled by a digital signal based off $Vperm$. However, MPD1 and MPD2 cannot be pulled down to ground because that would cause V_{gs} stress on the transistors. Also, to turn the devices off, pdn needs to go to V_{batt}, not just V_{PERM}. This necessitates the use of a voltage level shifter. The level shifter takes in a V_{PERM} domain input, and translates it into the V_{BATT} voltage domain. It also has protection circuitry (because it sees V_{BATT}) that prevents the output from going
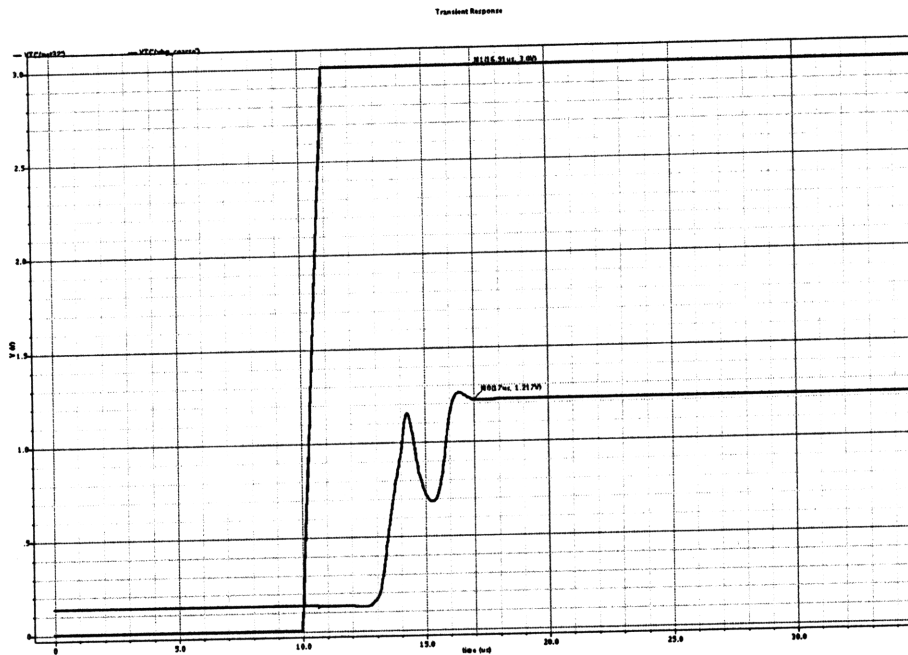
Figure 4.10: Start-up of the Coarse Bandgap voltage. The enable signal steps from 0-3V in 1μs. The bandgap voltage starts up in 5-7 μs

too low when the input is ground. Therefore, it converts V_{PERM} to V_{BATT}, and ground to V_{MID}. Figure 4.12 shows the voltage on the pdn node when the circuit is switched on and off. Figure 4.11 shows the level shifter topology used. The topology is fairly standard, and described in literature[11]. The major difference is the use of the diode-connected transistors D_1-D_4. These transistors prevent the nodes they are connected to from moving too low or too high by clamping them to V_{mid}. In addition, the transistors M1-M4 are used for high voltage protection, as outlined in the transistor protection strategy section(Section 4.2.2).

Figure 4.11: Circuit Diagram of the Level Shifter

4.4 Reference Current Generation

To generate reference currents for blocks like the DC/DC, we simply convert the bandgap voltage to a current using a resistor. Figure 4.13 illustrates this circuit. MC2,R6 and MC1 are scaled replicas of MO3, R4 and MO4 respectively in the output stage. Therefore, when the gate of MC2 is tied at V_{NB} (which is the voltage on the gate of MO4), the voltage at node A is V_{BG} (up to matching limits). Therefore, the
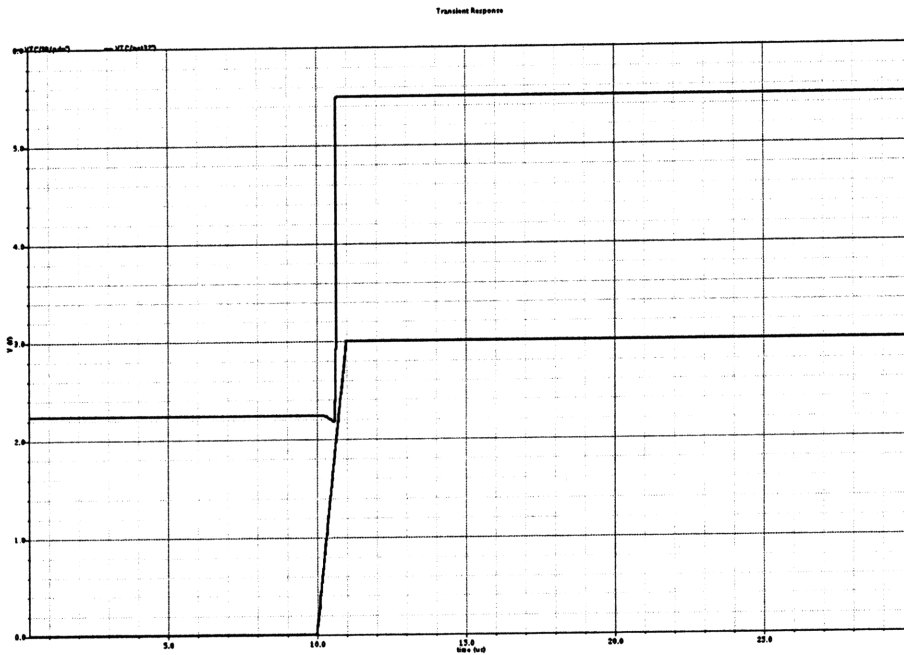
Figure 4.12: Voltage on the pdn node of the Coarse Bandgap. When the Bandgap is disabled, the pdn node is low. When it is enabled, it moves up to V_{batt}

current is approximately given by $\frac{V_{BG}}{R_6}$. This current is then mirrored using the current mirror MC3 to outside blocks that are connected to the pin I_{OUT}.
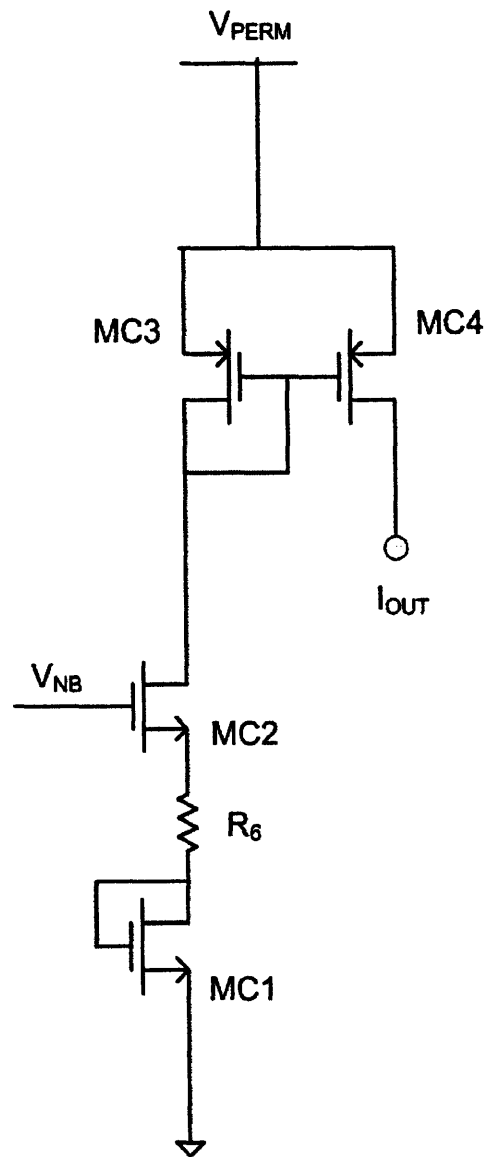
Figure 4.13: Voltage to Current Conversion Circuit for Coarse bandgap

Chapter 5

The Fine Bandgap

In this chapter, we discuss the implementation details of the fine bandgap.

Section 5.1 describes the overall operation of the fine bandgap, and identifies the various sub-circuits that are discussed at length in the latter sections.

Section 5.2 talks about the core bandgap-generating section of the circuit. We discuss the biasing structure, which enables us to save a bias leg of current. It also describes how the PSRR and noise specifications shaped the design of the opamp used in this structure.

Section 5.3 discusses the operation of the startup circuit.

Section 5.4 describes the control circuitry used to enable the bandgap circuit, and to generate an OK signal when the fine bandgap is ready for use.

5.1   Fine Bandgap Overview

The fine bandgap is based on the Brokaw architecture[1], but with a few changes. Figure 5.1 shows the overall circuit. As in the coarse bandgap, Q1 and Q2 are sized
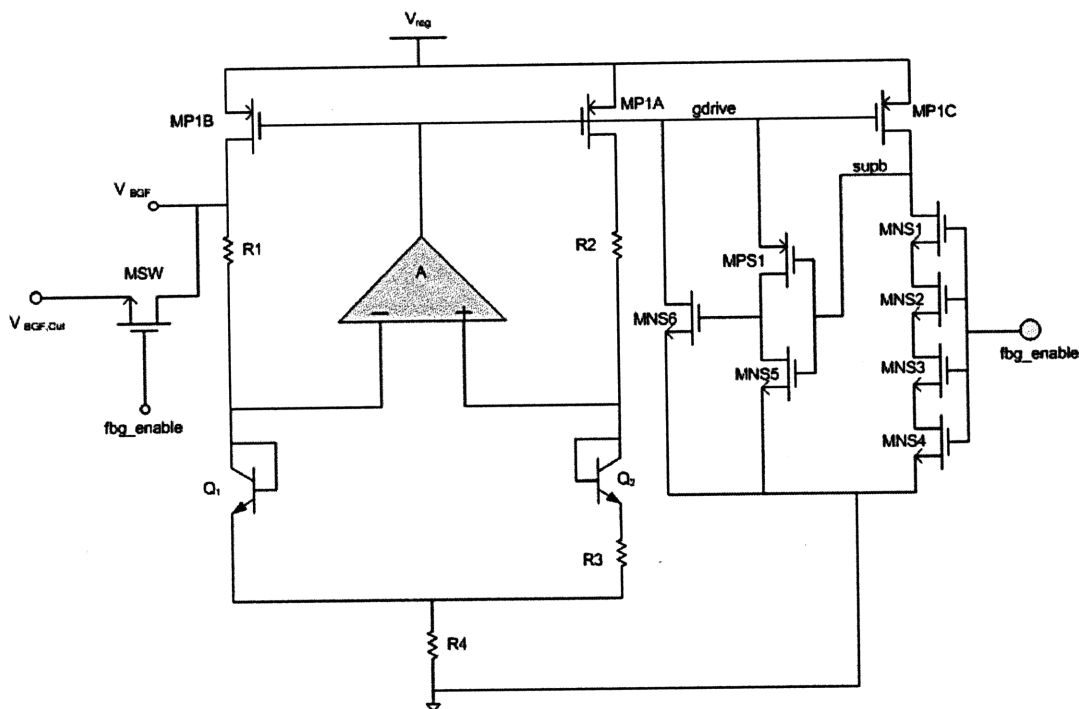
Figure 5.1: Circuit Schematic for the Fine Bandgap

differently, as are MP1A and MP1B. This forces a constant voltage of $V_{th}ln(16)$ across R_3. The reader is referred to chapter four for details of the derivation. This voltage causes a current of magnitude $\frac{V_{th}ln(16)}{R_3}$ to flow through Q1 and $\frac{V_{th}ln(16)}{4R_3}$ to flow through Q2.

In order to tune this voltage to the right bandgap voltage, it is necessary to be able to adjust one of the resistors R_1, R_3 or R_4. In this design, we have chosen to make R_1 the tunable resistor by adding trimming bits in it. The startup circuit again is controlled by an external signal, which turns it on.

As mentioned earlier, the fine bandgap runs off a supply (V_{reg}), which is approximately $2V_{BG,Coarse}$. This voltage is around 2.4V, therefore, there is no need to add any protection circuitry to this sub-block.

We have a control sub-circuit (fbg_control). It takes in control signals from other blocks, and uses it to power up and power down the circuit. In addition, it turns on/off MSW, which is the main path for the bandgap voltage to leave the sub-block. Finally, it also produces an OK signal when the fine bandgap voltage is ready for use.
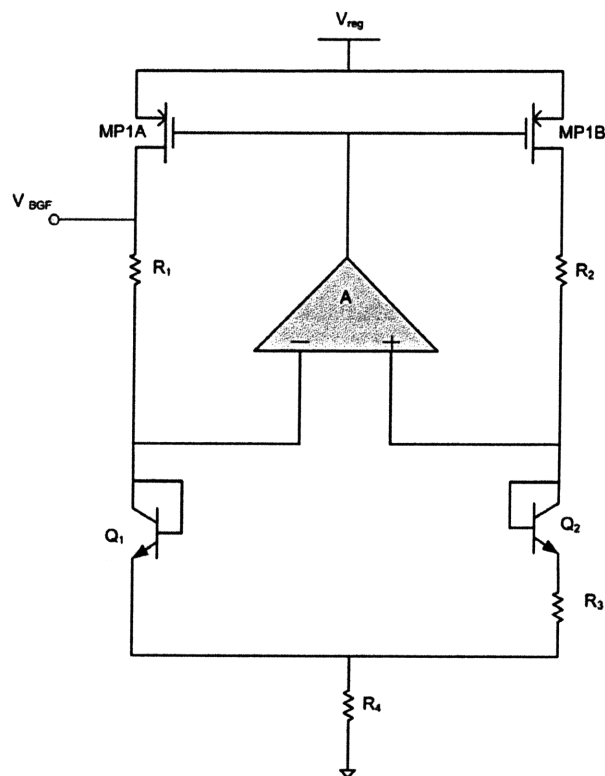
5.2 Fine Bandgap Core

Figure 5.2: Bandgap Core for the Fine Bandgap

Figure 5.2 shows the core bandgap circuit. This circuit is somewhat different from a classical implementation of a Brokaw cell. Adding the current mirrors M1A and M1B helps us in two ways. First, it prevents us from having to design a high gain opamp

with a output stage that can supply significant amounts of current. Instead, this design uses an extremely simple opamp (an actively loaded differential pair with a resistor tail source), which results in significant design time, area and power savings. In addition, by having scaled current mirrors(4:1), we are able to increase the voltage across R_3 from $V_{TH}ln(4)$ to $V_{TH}ln(16)$, without having to scale the bipolar transistors massively (4:1 instead of 16:1).

Also, in a traditional implementation like the one shown in figure 2.3, the opamp has to be able to run off a input common mode voltage of about V_{BE}, which is the base-emitter voltage drop of a bipolar transistor. Generally, this ranges between 0.6-0.8 V, depending on the current levels and the process corner. A voltage of around 0.6V is not enough to turn on an NMOS differential pair (and drop the required V_{DS} across the tail current source). This would force us to use an PMOS input pair. However, the output of the opamp needs to drive a PMOS transistor based off V_{reg}. Therefore, the output common mode voltage needs to be fairly high. This set of requirements complicates the opamp design, adding power and noise into the system.

One way of solving this problem is to increase the input common mode voltage of the opamp. Figure 5.3 shows a basic schematic of this implementation.This implementation gives a common mode voltage of $2V_{BE}$ for the opamp by using a pair of stacked bipolar transistors. This allows us to use an NMOS input differential pair for the input of the opamp. However, the bandgap voltage for lowest temperature variation is given by $V_{BG} = V_{BE} + K_1V_T$, for some K_1 dependent on process conditions. Now, with 2 stacked bipolar devices, if we wanted to generate a bandgap voltage by adding a resistor on top of Q2A(R_3 in Fig. 5.2), we would have $V_{out} = 2V_{BE} + K_2V_T$. Now, to get the best temperature performance, $K_2 = 2K_1$, which gives $V_{out} = 2(V_{BE} + K_1V_T)$. Now, as $V_{BE} + K_1V_T \approx 1.2V$, $V_{out} \approx 2.4V$. This is almost equal to the power supply voltage for
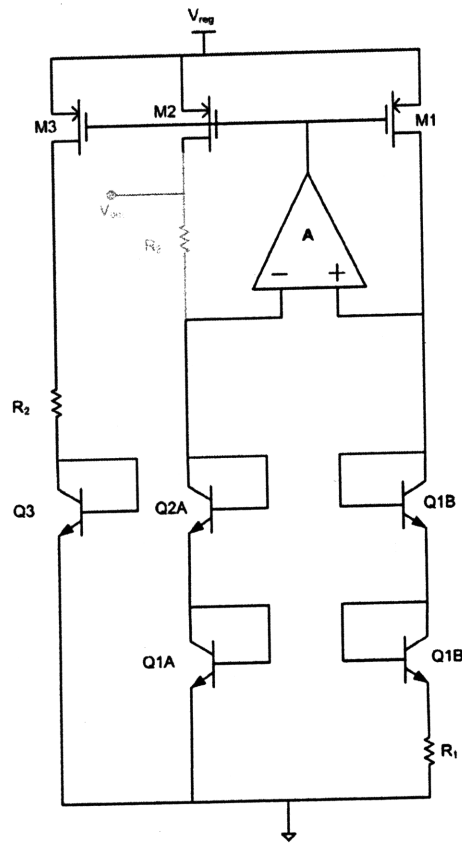
Figure 5.3: Alternative Implementation for the Core

this block, which leaves little to no headroom for the current mirrors. Also, the output bandgap voltage cannot be switched easily between the coarse and the fine bandgap voltage. The external capacitor used for filtering would need to charge for a while, and all the circuits connected to that node would have to be made aware of when the bandgap reference switches from a 1.2V reference to a 2.4 V reference. Due to all these issues, the implementation shown in Fig 5.2 does not generate the bandgap voltage on the same leg as the PTAT current. Instead, it mirrors the PTAT current, into a separate resistor/BJT leg, generating a bandgap voltage close to 1.2 V. This method

avoids the issues outlined above; however, it comes at the cost of additional power and area consumption.

Our solution to this problem was to use an architecture similar to the coarse bandgap. Instead of using two bipolar transistors to raise the input bias voltage of the opamp, we use only one, with a common resistor(R_4) to raise the input common mode voltage. This architecture is shown in Fig 5.2. In this case, the bandgap voltage is given by :
$V_{BGF} = \left(\frac{V_{TH}ln(16)}{4R_3} + \frac{V_{TH}ln(16)}{R_3}\right) R_4 + V_{BE,Q2} + \left(\frac{V_{TH}ln(16)}{4R_3}\right) R_1$. In order to tune this voltage across process corners, R_4 is built to be a trimmable resistor.

The reason for using both R_1 and R_4 to generate the PTAT voltage (unlike the coarse bandgap where we only use one resistor) is to improve the noise performance. To see the justification for this statement, let us look at the noise added by R_4. As any noise contribution from R_4 is common mode to the opamp, it is not reduced by the loop gain of the system, and contributes directly to output noise. In contrast, any noise added by R_1 causes a differential voltage to be developed across the opamp. This allows the output noise to be reduced by the loop gain of the system. Thus, the noise contribution of R_4 per unit resistance is higher than the noise contribution of R_1. To minimize the impact of noise from R_4, we placed just enough resistance R_4 in order to be able to bias the input NMOS pair of the opamp reliably, and moved the remaining resistance required to generate a trimmable bandgap voltage to R_1.

5.2.1   Loop Stability

From the stability point of view, the system looks like a two stage amplifier. The opamp comprises the first stage, and is simply an actively loaded NMOS differential pair. The second stage is formed by the current mirrors, which are loaded by the diode connected BJT's. The overall loop gain is given by Gain $= Ag_{m,MP1B}R_1$. To derive this

expression, let us first break the feedback loop at the output of the opamp in figure 5.2. If we apply a test small signal v_{in} to the gates of MP1A and MP1B, we get $i_{MP1A} = i_{Q1} = 4g_{m,MP1B}v_{in}$ and $i_{MP2A} = i_{Q2} = g_{m,MP1B}v_{in}$ (Because $g_{m,MP1A} = 4g_{m,MP1B}$). Now, the impedance of a diode connected BJT is equal to $\frac{1}{40I_C}$. Therefore, $V_- = V_{R4} + \frac{i_{Q2}}{40I_{C,Q2}}$, and $V_+ = V_{R4} + \frac{i_{Q1}}{40I_{C,Q1}} + i_{Q1}R_3$. As $\frac{i_{Q1}}{40I_{C,Q1}} = \frac{i_{Q2}}{40I_{C,Q2}}$, we have $V_+ - V_- = i_{Q1}R_3$. Therefore, loop gain is given by $A(V_+ - V_-) = Ag_{m_M P1B}R_3$.

This system is dominant pole compensated. The capacitance is provided by the C_{gs} of MP1A and MP1B. Due to the relatively small loop gain, and the large size of the transistors, there is no need for any extra compensation capacitance.
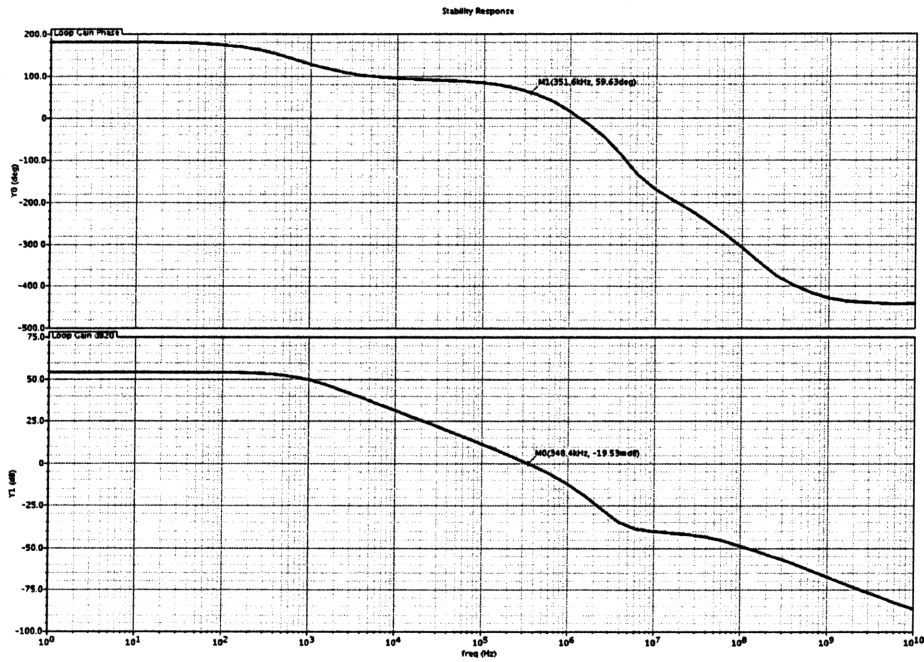

Figure 5.4: Representative Gain and Phase Margin for the Fine Bandgap

5.2.2 PSRR of Fine Bandgap

To a first order approximation, the power supply rejection of the fine bandgap is equal to the loop gain in the core bandgap. Because of the differential nature of the opamp input, the effective loop gain is given by Ag_mR_3, where A is the gain of the opamp, and g_m is the transconductance of MP1A. However, this first order approximation suffers from many faults. It does not take into account the effects of the output impedance of the current mirrors, and the finite PSRR of the opamp itself. Therefore, the actual PSRR performance is significantly different from this first order calculation.
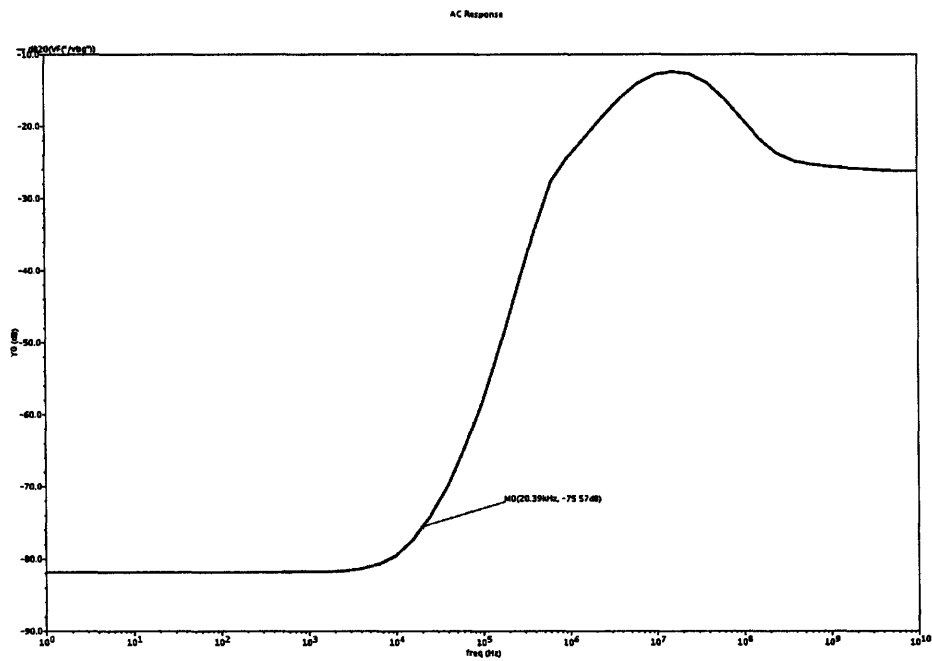

Figure 5.5: Power Supply Rejection of the Fine Bandgap Voltage

5.2.3 Noise Performance

In contrast to the coarse bandgap, the noise performance of the fine bandgap is quite critical. There are three major sources of noise in this circuit - flicker noise from transistors, thermal noise from transistors and thermal noise from resistors. By using an external capacitor for filtering, we are able to reduce the impact of the thermal noise from both resistors and transistors. However, as we are interested in the audio-band noise for this circuit (20-20 kHz), flicker noise plays an important role too. MOS transistors have significantly higher flicker noise than bipolar transistors. Therefore, for the purposes of this discussion, we ignore flicker noise contributions from the bipolar transistors.

There are two major ways of dealing with flicker noise. Chopping is a technique that is commonly used to eliminate flicker noise[12]. By flipping transistors in pairs at a frequency f, we can translate the low frequency flicker noise to that frequency level. By choosing a large enough frequency, we can shift most of the flicker noise away from the audio band. The disadvantage of this technique is the added complexity, as well as the higher power consumption that would be required. In addition, switching introduces tones at multiples of the switching frequency, which have an adverse impact on the RF blocks on this chip. The other way is to increase the size of the devices, as flicker noise goes down as the square root of the area. In this circuit, we chose to increase the size of the devices to reduce flicker noise.

Overall, by using larger transistors and external filtering for thermal noise, we have been able to bring the noise in the audio band down to about 0.6 μV, well within the required specification.
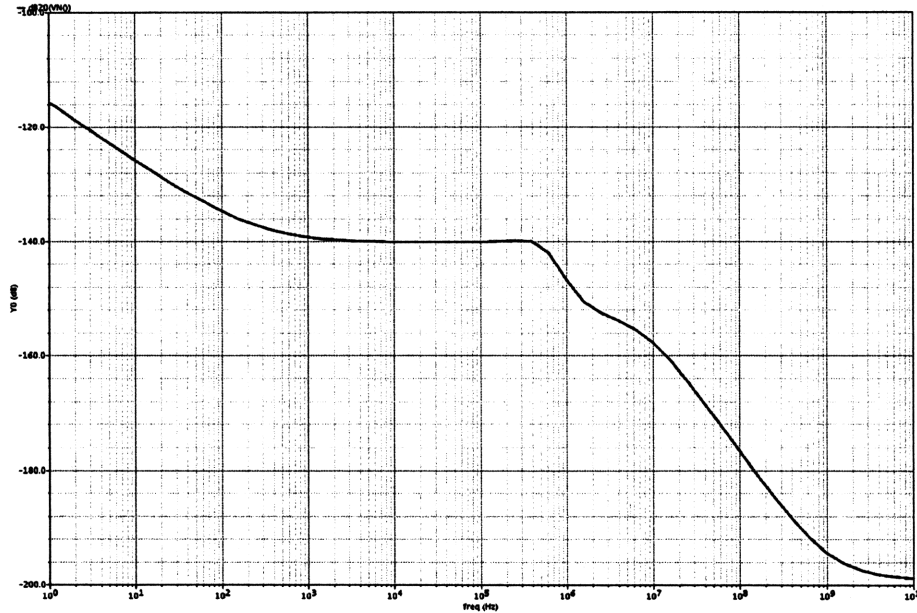
Figure 5.6: Output Noise - Unfiltered

5.2.4   Resistor Trimming

Figure 5.8 shows the trimmable resistor used. Only a small portion of the resistor is trimmable; the rest is fixed. We have 16 settings(4 bits) for the effective resistance value. To change the effective resistance (and the bandgap voltage), we simply short a new switch to the output. The resistance above the shorted switch does not affect the bandgap voltage. The trimming resistor step size is 60 Ohms. With around 100 μA of current flowing (in typical case), we have a step size of 6 mV. In order to trim the voltage, the production tester looks at the output bandgap voltage, and shorts different switches till it finds the setting that is closest to the reference level desired. The desired reference level will be determined by lab measurements of the chip to figure out the optimum voltage for best temperature performance.
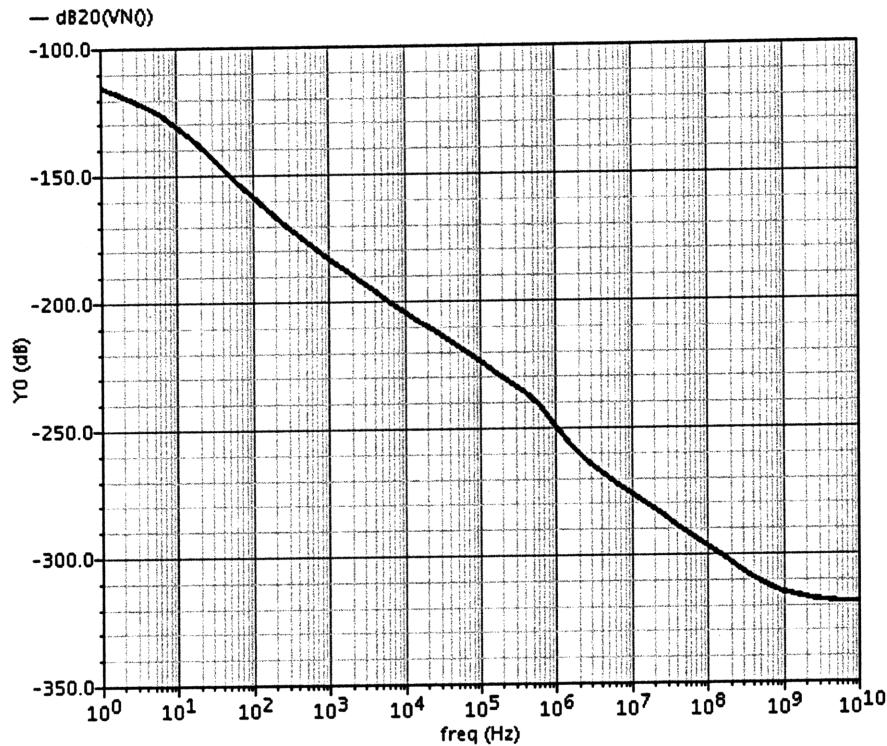
Figure 5.7: Output Noise - Filtered with a 1μF External Capacitor

5.3 Fine Bandgap Startup Circuit

The startup circuit for the fine bandgap is similar to the circuit used in the coarse bandgap. Figure 5.9 shoes the startup circuit used. Again, the aim is to pull down the gates of the devices MP1A and MP1B when the circuit has to be turned on.

When the circuit is off, the node gdrive is pulled up to V_{reg} by power down devices. The node supb is also pulled up by a pullup device(MP1C).This turns the output of the inverter formed by MPS1/MNS5 low, and keeps the transistor MNS6 off. Now, when the bandgap is required to be turned on, the node fbg_enable is pulled up to V_{reg}. This
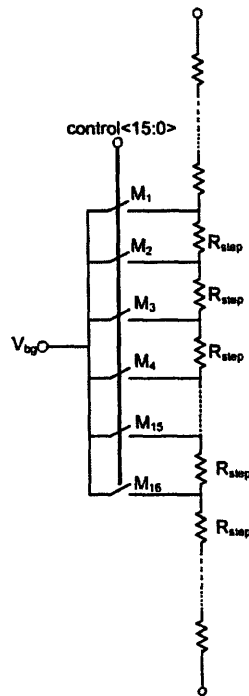
61

Figure 5.8: Trimming resistor

causes the node supb to collapse (as M1D is still off). This turns on MNS6, which forces gdrive to go low, as it can not supply DC current. Once the bandgap is near proper operation, M1D is turned on strong enough to pull supb up near V_{reg} (because M1D is built to be much stronger during normal operation than MNS1-MNS4). This turns off MNS6, and the startup circuit is disconnected from the main circuit.

5.4 Fine Bandgap Control Circuitry

In order to make efficient use of power, the fine bandgap is turned off until its accuracy level is required. This brings about the problem of switching accurately between the coarse and the fine bandgap levels. The block vbg_out achieves this(Figure 7.1).
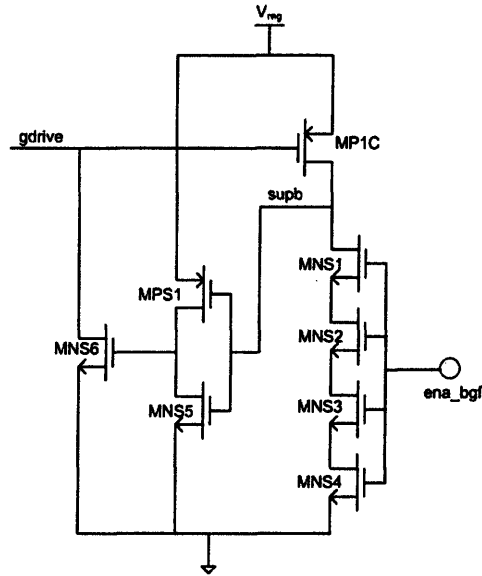
62

Figure 5.9: Schematic for Fine bandgap Startup Circuit

In order for that block to be able to do its job properly, it needs to get a signal that indicates that the fine bandgap is operational and the reference provided by it is ready to use. This signal is generated internally in the fine bandgap.

The control circuitry for the fine bandgap, in addition to controlling the power-down devices, is responsible for generating this signal. To generate this signal, it uses a hysteretic comparator, with a hysteresis of about 125 mV. This comparator is powered using V_{PERM}, and uses bias currents from the coarse bandgap. It compares the fine bandgap voltage to the coarse bandgap voltage, and switches high when $V_{BG,Fine} >$ $V_{BG,Coarse} - 75mV$. The output goes low when the $V_{BG,Fine} < V_{BG,Coarse} + 75mV$. The hysteretic comparator is shown in Figure 5.11.

The hysteresis prevents the power_ok signal from turning on and off repeatedly due to chatter noise on the fine or the coarse bandgap. The main for using hysteresis, however, is to prevent a situation where the power_ok signal does not turn on at all because of a
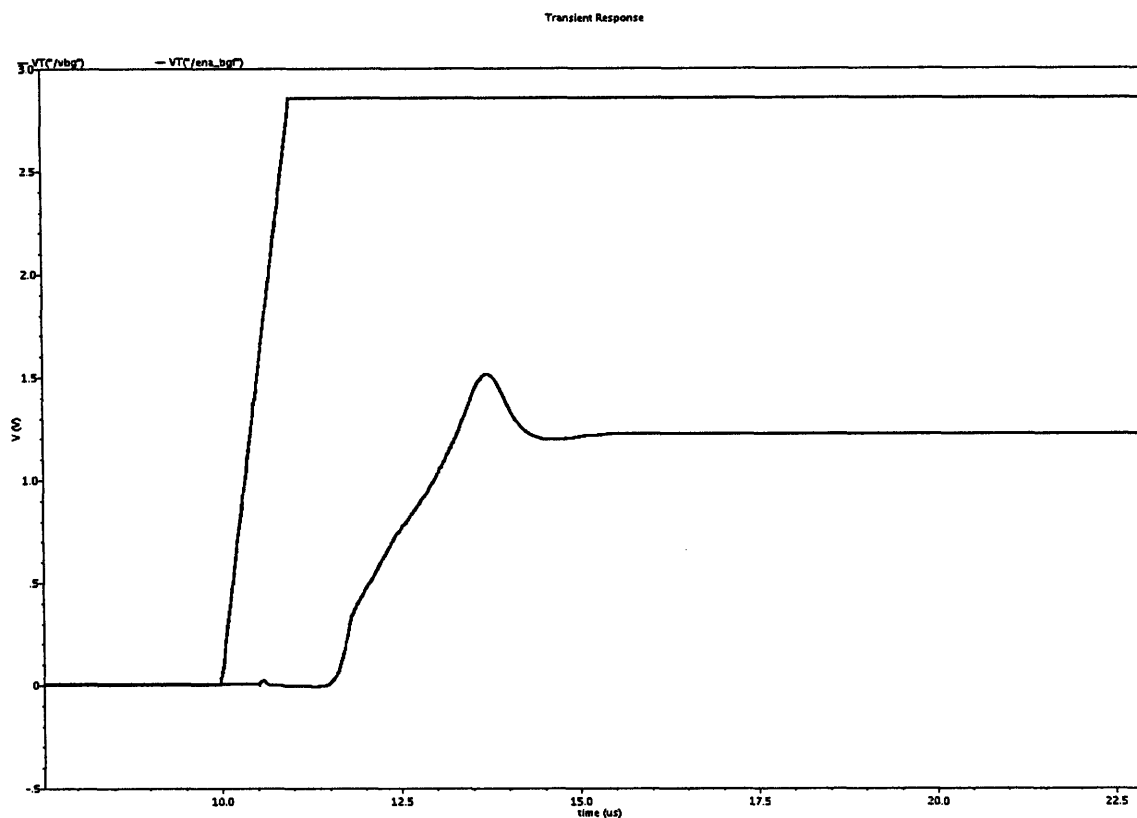
63

Figure 5.10: Fine Bandgap Starting up when Enable Signal is Stepped from 0 to 3 V

small mismatch in the two voltage levels. Because the topologies, current and trimming levels of the fine and the coarse bandgap are different, it is likely that they will have slightly different DC levels. If the fine bandgap level happened to be below the coarse bandgap level, the fine bandgap would never turn on. By switching 75 mV below the coarse bandgap level, we ensure that regardless of small DC differences between the coarse and the fine bandgap, it is always possible for the comparator to trigger.

There are different ways of introducing hysteresis in a comparator. The most commonly used ones are described in[6]. The scheme currently used is described in[13]. Figure 5.12 shows a plot of the hysteresis produced by the circuit in figure 5.11.

Figure 5.11: Circuit for Hysteretic Comparator

Figure 5.12: Plot of Hysteresis

Chapter 6

Voltage to Current Converter for Fine Bandgap

Almost all the blocks in the chip this bandgap was designed to use currents rather than voltages. The reason is twofold. First, the current can be used as a reference to set up bias levels inside the block. Also, regenerating the bandgap voltage inside the target block from a current is typically more accurate than routing the voltage directly. This is because any routing impedance can cause a voltage drop, and any interference from other blocks has a bigger impact.

Section 6.1 describes the operation of the refgen_i circuit. It also discusses the trimming strategy employed, which allows a single trim to be used for two different set of reference currents.

Section 6.2 outlines the selection process for the opamp, and the stability analysis for the system.

6.1 Refgen_I overview and Trimming strategy

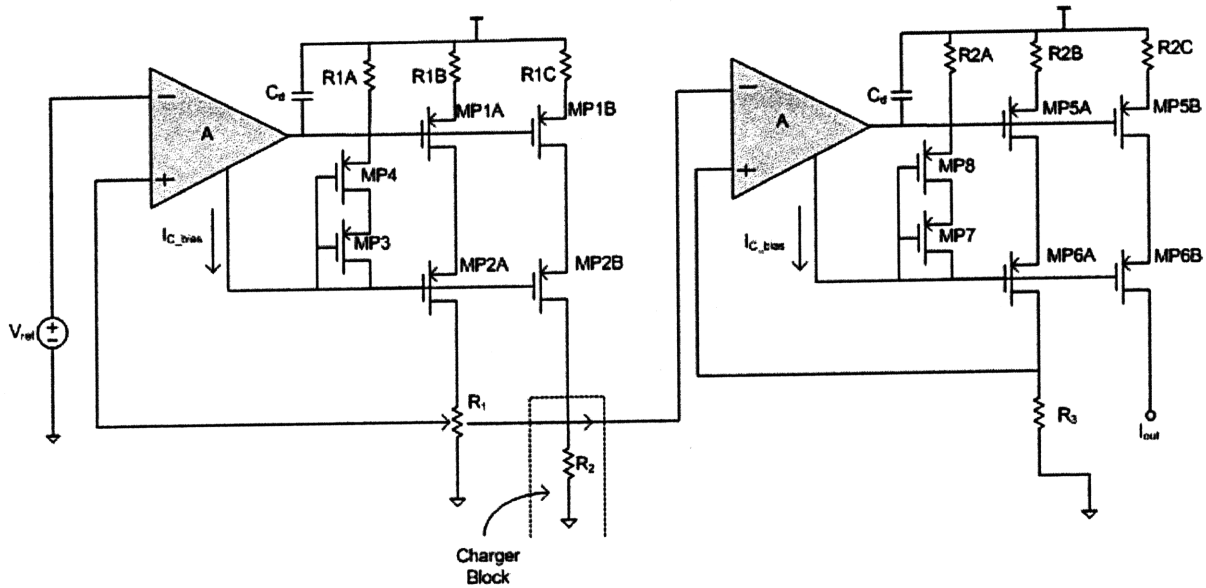Figure 6.1: Voltage to Current Conversion Schematic

Figure 6.1 shows the operation of the refgen_i block. Overall, the block looks similar to a regulator with a trimable resistor. The opamp input compares a reference voltage to the voltage across a trimable resistor. It then drives that voltage close to the reference level. By adjusting the value of the resistance R_1, we can change the amount of current that flows through MP1A. This current is then mirrored through MP2A to the charger block. This current is given by: $I_{MP2A} = \frac{V_{BGF}}{R_1}$. The charger block has a resistance R_2 which converts this current to a scaled bandgap voltage given by $V_{CHG} = \frac{R_2}{R_1} V_{BGF}$.

To trim the output voltage the tester looks at V_{CHG}(Output of the charger block), and changes R_1 till the voltage is close to the intended DC target. The original bandgap voltage was trimmed for an accurate temperature coefficient, which may not necessarily be at the DC level the charger requires. With this second trimming step, the output

voltage on the charger is both DC-accurate and has a low temperature coefficient.

The current mirrors are cascoded for better matching accuracy between the legs. The bias voltage for the cascode transistors is generated by the opamp. In order to further increase matching accuracy, we degenerate the current mirrors using R_{D1} and R_{D2}.

There is another identical regulator, used to supply the other blocks with a reference current. The reference voltage for this regulator is generated by the regulator used for the battery charger. The following subsection discusses the advantage of using this mechanism.
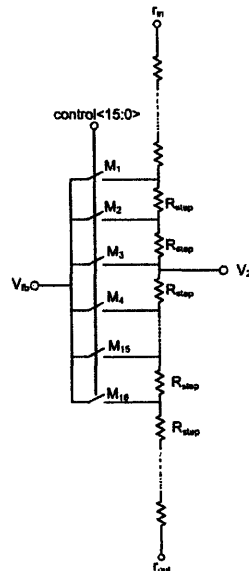
6.1.1 Trimming Strategy

Figure 6.2: Trimming resistor

Figure 6.2 shows the trimming resistor used in the refgen_i circuit. Only a small part of the resistor is actually trimmable, and shown in the circuit. With 4 control bits,

there are 16 possible trim states (there is a binary-to-thermometer decoder not shown in the figure). Each unit resistor (R_{Step}) is around 60 Ohms, and the nominal current in the resistor is about 100μA. This gives a step size of about 6mV. Multiple switches are connected, with one end shorted common and the other end going to individual step resistors. Only one switch is shorted at any time.
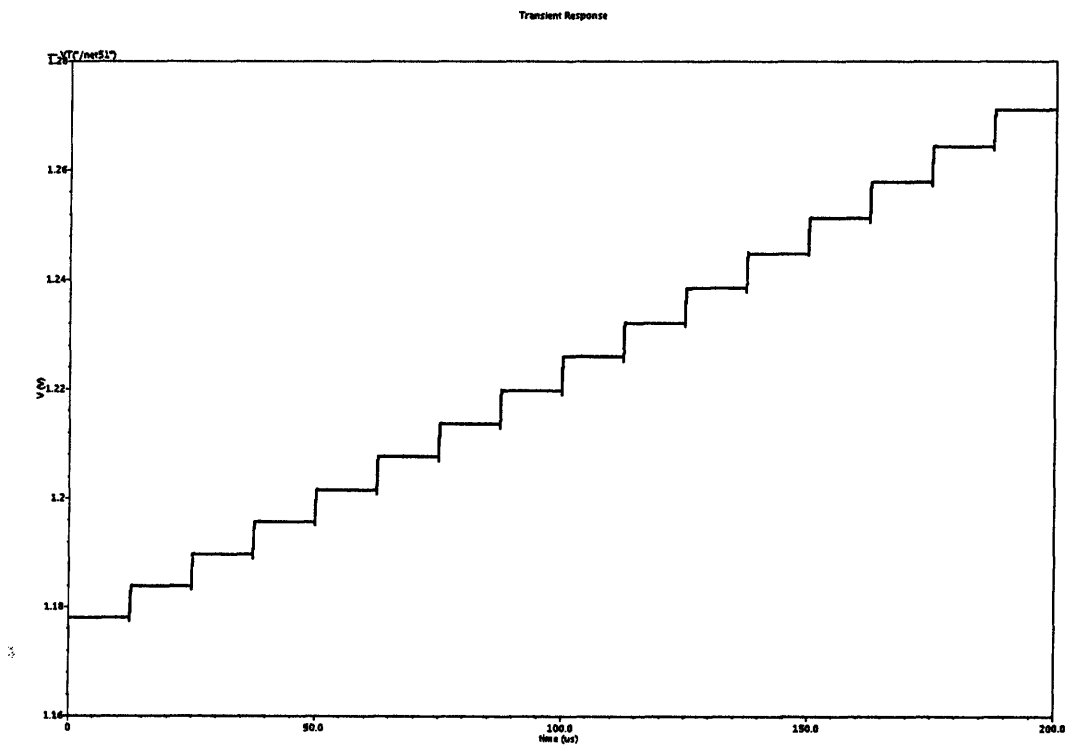

Figure 6.3: Trimming range of the resistor, when the trim code is swept from 0000 to 1111

The "effective" R_1 is determined by looking at the impedance from the shorted switch to r_{out}. The impedance seen above the shorting switch is not in the feedback path, and therefore, not a part of the "effective" R_1 seen by the circuit. To change R_1, one simply needs to change the switch that is shorted. The shorted node is used as the

69

feedback node into the opamp.

There is another connection from a fixed point in this resistor (V_{ref2}). This node is used as the reference voltage by the second regulator in refgen_i. This is used to essentially auto-correct the reference voltage. Let the (fixed) resistance from V_{ref2} to ground be $R_r gen$. Now, $I_{ref} = \frac{V_{BGF}}{R_1}$. Therefore $V_{ref2} = \frac{R_{rgen}}{R_1}$. Comparing to the expression for V_{CHG}, we have, $V_{ref2} = \frac{R_{rgen}}{R_2} V_{CHG}$. Therefore, if $R_{rgen} = R_2$, V_{ref2} automatically gets trimmed when the battery charger is trimmed. The only differences are caused by mismatch between the pairs of current mirrors, or any errors induced in the charger block (which are expected to be small).

6.2   Opamp Selection

The opamp used in this circuit has some specific criterion it must satisfy. It needs to be able to run off a common mode voltage of around 1.2 V on the input, and drive outputs referenced to V_{ana}. It also needs to have relatively high gain to ensure that the feedback voltage is close to the bandgap reference.

There are three major topologies that could have been considered : A two stage opamp design, a telescopic opamp, and a folded cascode opamp. The two stage opamp was abandoned because of the current mirror leg that follows it. With a two stage opamp design, the current mirror/feedback resistor would act like a third gain stage, and add another parasitic pole at the input of the opamp (node A). This, combined with the dominant and non-dominant pole generated in a two-stage opamp design would make such an opamp hard to stabilize. Therefore, we need to use a topology without any internal high impedance nodes. The two major topologies of this kind available are telescopic and folded cascode opamps. Both of them offer similar performance

70

levels. The biggest difference in our application is their swing range and input offset. A telescopic opamp has significantly constrained swing range compared to a folded cascode opamp. However, a folded cascode opamp contributes more offset (because there are three major pairs of transistors that can contribute to input offset compared to two pairs in a telescopic.

Because of the relatively large current levels in the current mirrors (100 μA) and the source degeneration (R_{D1} and R_{D2}), the output node of the opamp has to be able to go low; in fact sometimes close to 1V. Therefore, we decided to choose a folded cascode topology in fact, despite the slightly higher complexity and offset penalty.
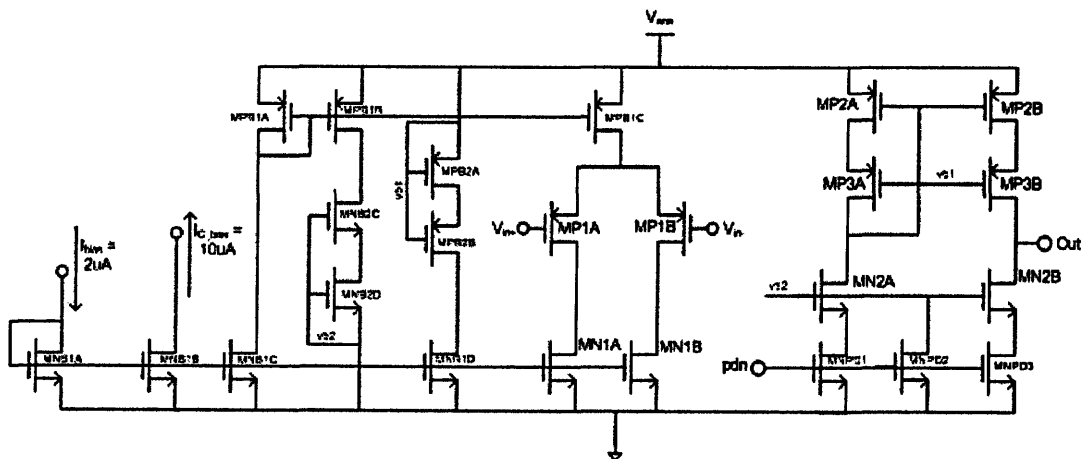

Figure 6.4: Opamp

Figure 6.4 shows the schematic for this opamp. As this is a common topology well described in literature [6], we shall not describe it in detail here.. The only point of interest is current I_{C_bias}. The transistor generating is sized to make it about 1/10 the current of the output legs. This current is then sent through MP3 and MP4 to generate the bias voltages for the cascode transistors.

6.2.1 Stability

The dominant pole of this regulator structure is at the output of the operational amplifier. The primary non-dominant pole is at the feedback input of the opamp. In order to ensure stability at all corners, we have added additional capacitance (C_d) to the output of the opamp. Figure 6.5 shows a representative stability plot for this system.
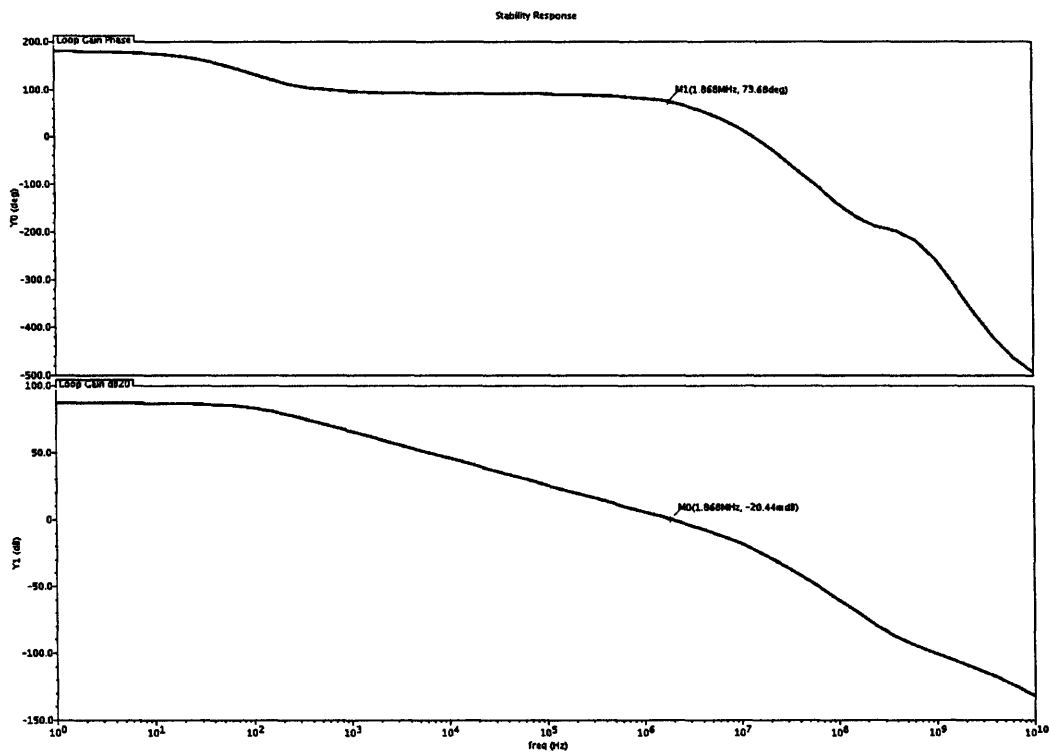

Figure 6.5: Loop gain and phase margin

Chapter 7

Output Block for Bandgap Voltage

Because this block uses two different bandgap voltages, we need a way to switch between the two voltages. The major responsibility of the vbg_out block is to ensure that this happens in a timely, predictable manner. In addition, it serves as a fast charging circuit for any external cap connected to filter the bandgap voltage. The two circuits are independent of each other.

Section 7.1 discusses the circuitry used to switch between the coarse and the fine bandgap voltage. This circuit also outputs an OK signal when the output bandgap voltage is ready to use (similar to the way the fine bandgap sub-block sends out an OK signal).

Section 7.2 describes the fast-charging circuitry that is used to charge the output capacitor when coarse bandgap is turned on. This circuit is self regulated(i.e it turns itself off once the voltage is close enough to the final required voltage).
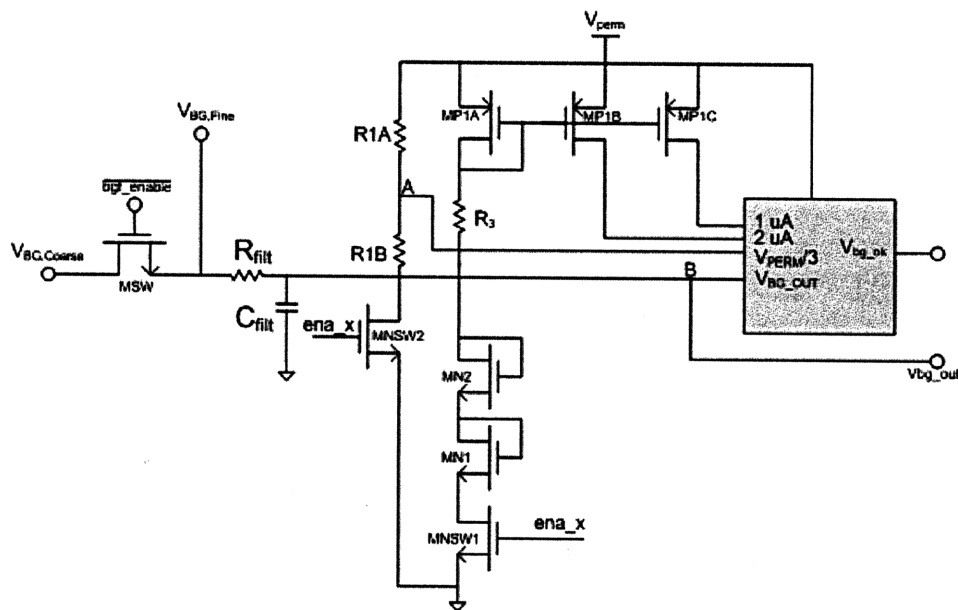
Figure 7.1: Switching from Coarse to fine circuit

7.1 Switching Circuitry

Figure 7.1 shows the circuit used to switch between the coarse and the fine bandgap voltage. The coarse bandgap voltage is separated from the fine bandgap voltage by a switch. When the fine bandgap is disabled, the bgf_enable signal is low and the switch MSW is turned on. The Fine bandgap sub-block has its own isolation switch built into the sub-block (MSW in Figure 5.1). Therefore, the coarse bandgap node is never directly shorted to the fine bandgap node. The voltage is then filtered through a filter R_1 and C_1. This helps ensure a smooth transition from the coarse bandgap voltage to the fine bandgap voltage and vice versa when the voltage is switched.

The resistors R_{1A} and R_{1B} are used to divide the vperm voltage down to about $\frac{V_{PERM}}{3}$. Given a nominal V_{Perm} of about 3V, this makes for a node voltage at node V_{ref} of around 1V. This voltage is then compared to the bandgap voltage on node B

74

to determine if the bandgap is ready to use. To reduce design time, the same circuit used to determine whether the fine bandgap was ready for use (Fig 5.11) was reused here. However, instead of comparing the coarse and the fine bandgap voltages, here we compare the coarse bandgap (or the fine bandgap) voltage to $V_{Perm}/3$. The fact that $V_{Perm}/3$ is always less than 1.2V allows us to ensure that the OK signal always turns on.

The bias currents required by the comp_vbg_ok block are generated by the MP1A/R_3 leg. The resistor and the transistors are sized to approximately generate around 2μA of current. This current is then mirrored and supplied to other blocks. There will be significant variation in this current, but this does not affect the performance of this circuit measurably. All of these current legs have power-down devices to ensure that the sub-block draws no power when it is powered down. Figure 7.2 shows the output voltage switching between the coarse and the fine bandgap at around 100 μs.

7.2   Fast-Charge Circuit

Figure 7.3 depicts the circuit used to fast charge the output capacitor that is connected to the *vbg_f* node. The main inputs to this block are *vnb* , *ip2u* and *vbg_fchg_enb*. The node *vnb* is a voltage generated internally in the coarse bandgap circuit, and is equal to $V_{BGC} + V_T$. *ip2u* is a bias current supplied by the coarse bandgap to this block, while *vbg_fchg_enb* is a power-up signal used to enable the fast charge circuit. This signal is an active low signal.

To see the operation of this circuit, let us see what happens when *vbg_fchg_enb* is switched low. This forces the output of the NOR gate high, and turns on the transistor MNSW3 hard. As the output capacitor is initially discharged, *vbg_f* is around zero.
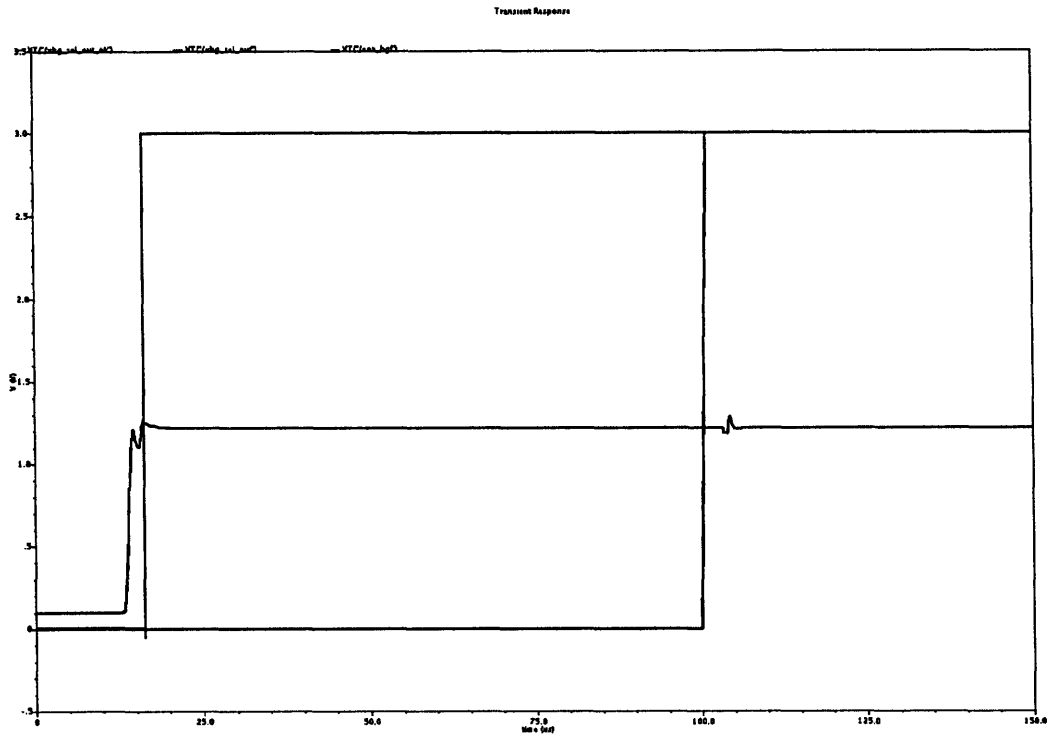
Figure 7.2: Switching from Coarse to Fine Bandgap at 100 μs

Therefore, MNSW3, acting like a switch, pulls down its drain close to zero. This allows a large current to flow to the output capacitor through MNO1C and MPM1A. This current then tries to flow through the current mirror MPM1B. However, the current mirrors MNM1A and MNM1B attempt to force a much lower current of 2 μA at the same time through the same leg. Therefore, MPM1B pulls up node A high (close to V_{Perm}), as long as the current in MPM1A is much higher than 2 μA.

Because of this large charging current from MPM1A, the node vbg_f starts rising, and so does the drain of the transistor MNSW3. This reduces the overdrive on MNO1C, and thus reduces the amount of current available for charging that flows through MPM1A. As the current approaches 2 μA, the node A collapses down, and causes the output

76

Figure 7.3: Fast Charge Circuit

of the NOR gate to go low. There is a hysteretic comparator in the path to prevent "chattering" or accidental flips because of noise.

When the circuit starts up, node A is held high by the power down device MPSW2. To prevent the node from collapsing in the time between MPSW2 turning off and MPM1B starting to carry current, we add a small capacitor C_0. Figure 7.4 shows the fine bandgap voltage with an external capacitor attached. Note that the fast charge block turns off (signal transitioning from high to low) when the output voltage is close to the final bandgap voltage.
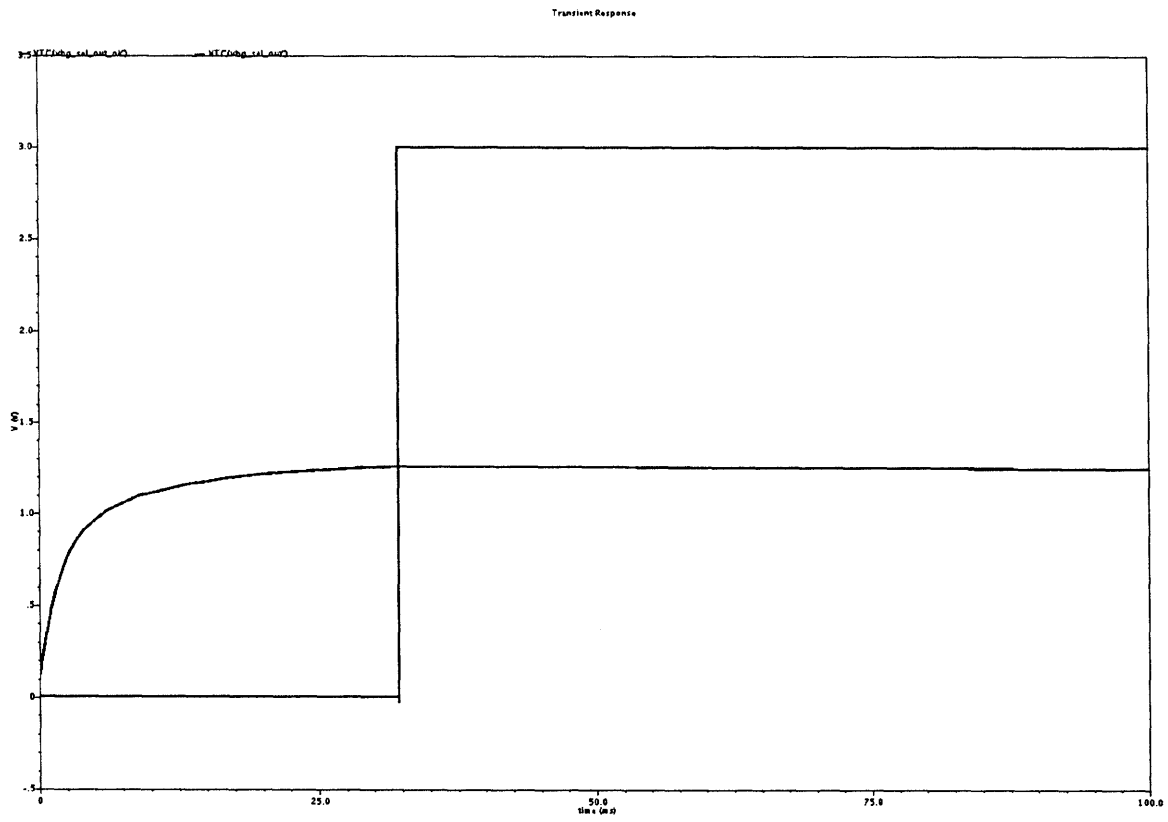
Figure 7.4: Fast Charge Plot

Chapter 8

Future Directions of Research and Overall Simulation Results

Section 8.1 outlines suggestions for future explorations/improvements in this field.

To conclude this thesis, we present some overall simulation results for the whole bandgap block. Section 8.2 contains some plots and a table outlining key metrics, proposed specification and results from simulation.

8.1 Future ideas for further exploration

Although there has been a lot of progress in the present design, there are a few areas where we feel further optimization is possible.

The biggest area for improvement is getting rid of the external capacitor. External components cost money, board area and decrease reliability and manufacturability. The primary stumbling block in the present design to removing the external capacitor is the noise requirement. Without an external capacitor, the output noise of the present

design is around 19 uV in the 20-20kHz band. To get rid of the capacitor, the noise level needs to come down to about 3 μV. Promising research is being conducted in the area of low-noise bandgap designs in[14], which may make this feasible.

If the bipolar transistors in the coarse bandgap could be biased without requiring such an output leg, and an alternate way of supplying the current to the fine bandgap was found, there could be substantial power savings in the coarse bandgap stage. As standby power is an important metric for cellphones, this power savings could prove to be valuable.

The temperature compensation schemes used in this design are simple, first order techniques. One can achieve better temperature coefficients by using curvature compensation techniques like second order compensation and exponential compensation.

There are three calibration steps in the present design - for the coarse bandgap, fine bandgap and refgen_i. Each of these adds to the cost and production tester time required for validation. By coming up with a scheme to eliminate some or all of these trimming steps, we could save significant amounts of production tester time.

There are a few useful features that could be added to this design. Instead of supplying currents that vary with resistance, it could supply constant-g_m or absolutely accurate currents. This would minimize variations in other blocks due to bias current variation. The internally generated PTAT currents could also be used to estimate the chip die temperature by comparing it to the bandgap voltage.

Finally, there are multiple opportunities for power/area savings in the auxiliary blocks. For example, the present scheme for generating Vperm/3 in vbg_out uses large resistors, and consumes a significant amount of power in the low-power mode. An alternative implementation to detect when the coarse bandgap voltage is up and running could save some power and area.

8.2 Overall Results

The following table summarizes the specifications and results for this block. The actual values are got from simulation, and represent worst-case values.

Metric	Specification	Actual Value
Power Supply Rejection 20kHz(No Capacitor)	80 dB	91.95 dB
Power Supply Rejection 20kHz(With Capacitor)	80 dB	158 dB
Power Supply Rejection 217Hz(No Capacitor)	120 dB	140.6 dB
Power Supply Rejection 217Hz(With Capacitor)	120 dB	167.6 dB
Audio Band Noise (20Hz-20kHz)(No Capacitor)	3 μV	18.5μV
Audio Band Noise (20Hz-20kHz)(With Capacitor)	3 μV	1μV
Stability - Phase Margin	-	59 degrees
Stability - Gain Margin	-	15.3 dB
DC Variation - Fine Bandgap Mode	+/-0.4%	+/- 0.25%
DC Variation - Coarse Bandgap Mode	-	+/- 0.25%
Temperature Variation - Fine Bandgap Mode	+/- 0.4%	+/-0.2%
Temperature Variation - Coarse Bandgap Mode	-	+/-0.4%
Turn on Time - Without Capacitor	-	17 μs
Turn on Time - With Capacitor	-	12.5 ms
Current - Low Power Mode(Core-V_{Batt})	30μA	20μA
Current - Low Power Mode(Aux-V_{Perm})	-	20μA
Current - High Power Mode(Core-V_{Batt})	300μA	280μA
Current - High Power Mode(Aux-V_{ana})	-	287μA

The following figures (Fig. 8.1 - 8.6) showcase the overall performance of the bandgap circuit:
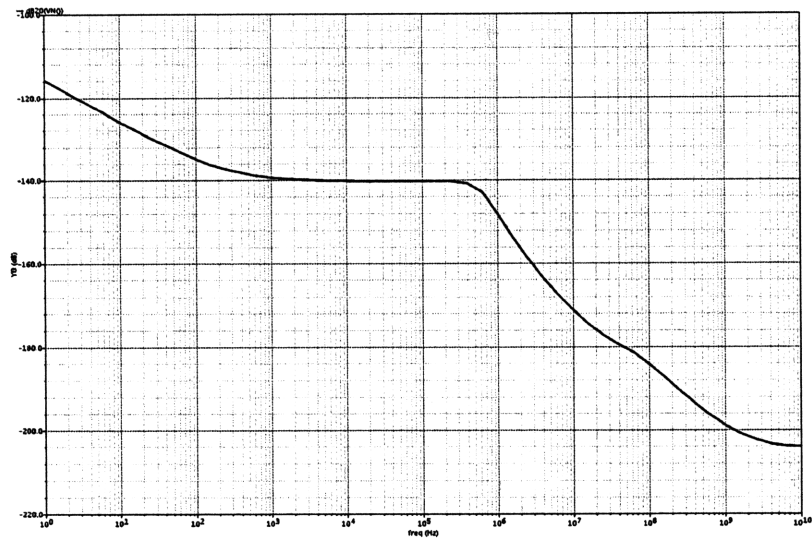
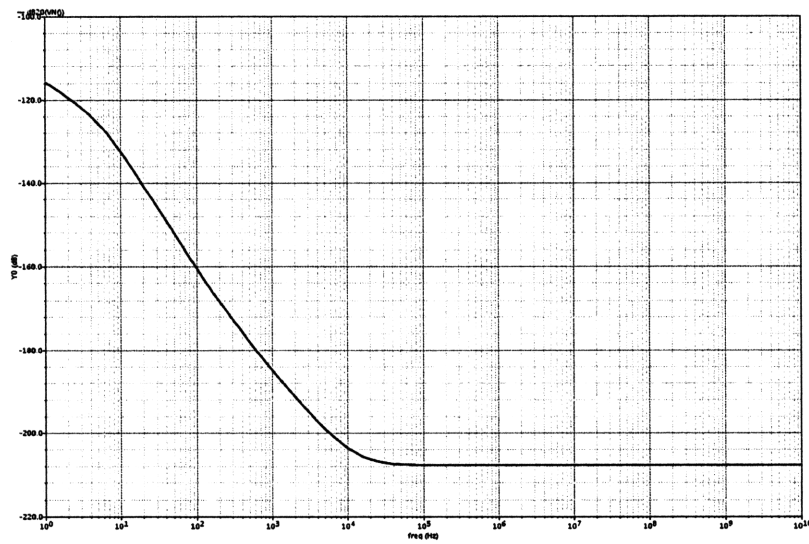Figure 8.1: Output Noise Spectrum Without External Filtering Capacitor

Figure 8.2: Output Noise Spectrum With External Filtering Capacitor
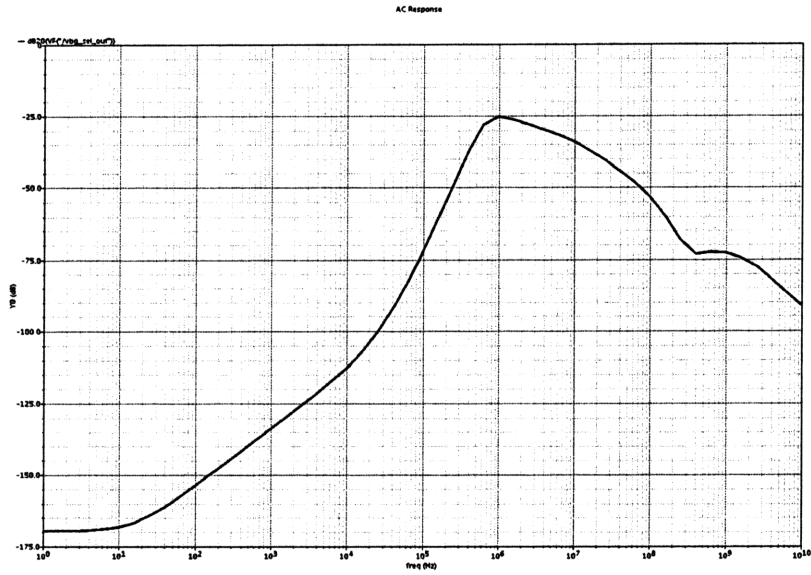
82

Figure 8.3: PSRR Without External Filtering Capacitor
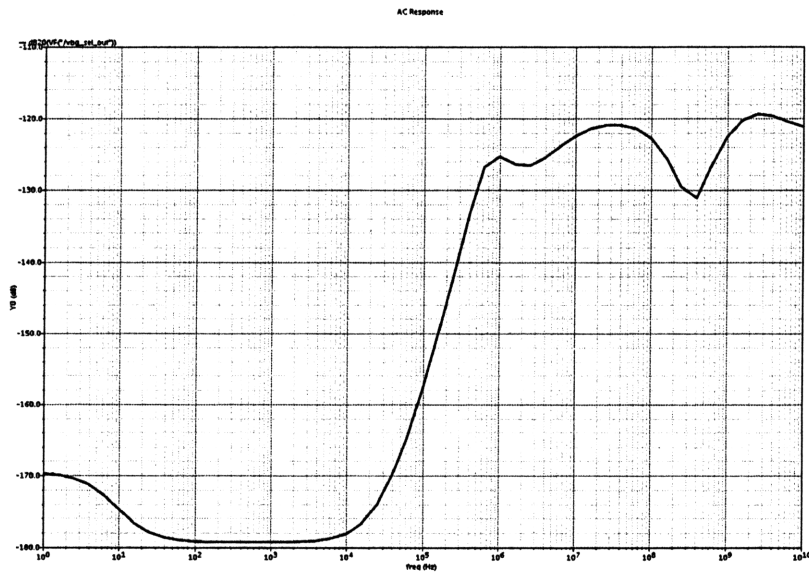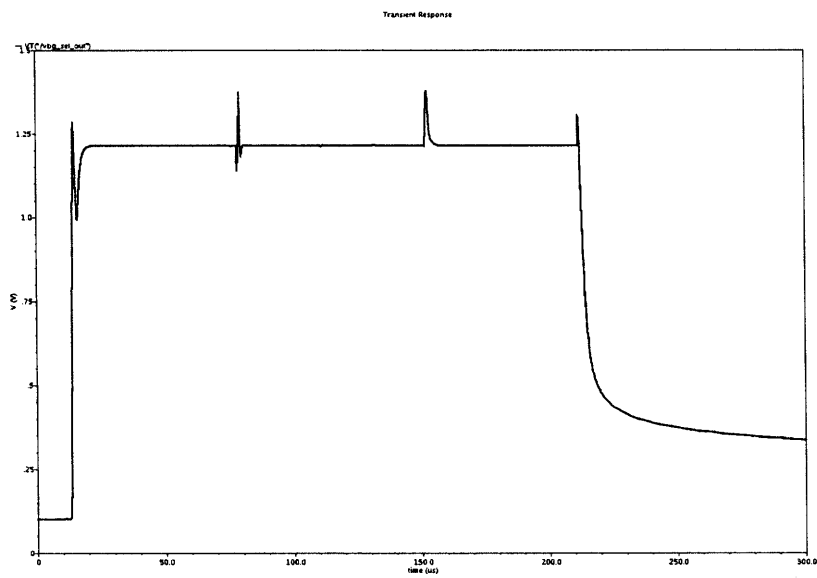

Figure 8.4: PSRR With External Filtering Capacitor

Figure 8.5: Overall Transient simulation. The coarse bandgap is enabled, followed by the fine bandgap. Then the fine bandgap is disabled, and finally, the coarse bandgap is disabled.
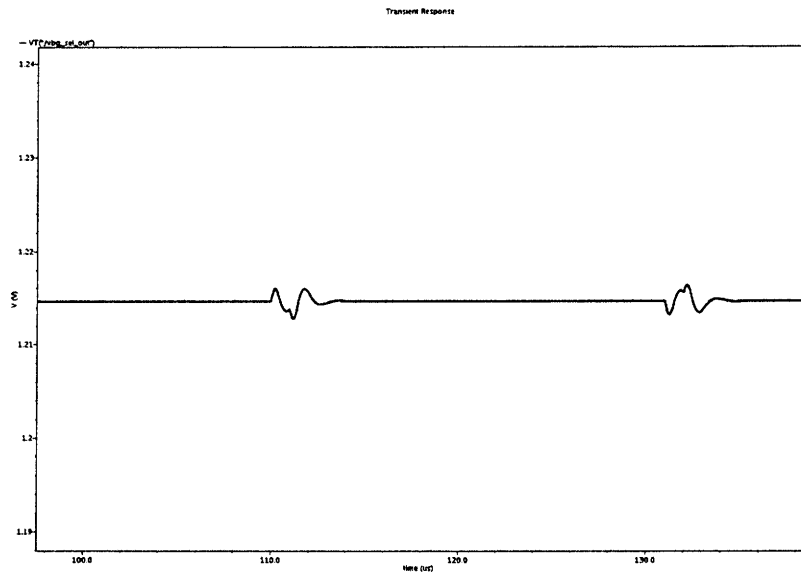
Figure 8.6: Output Voltage when Vbatt steps 300 mV from 4.2 to 4.5V and then back down to 4.2V. Notice the Output Voltage is well within the +/-5% specification

Bibliography

[1] A.P. Brokaw: "A Simple Three-Terminal Bandgap Reference," IEEE J. Solid State Circuit,Dec 1974

[2] Y.P. Tsividis: "Accurate Analysis of Temperature Effects in I_c-V_{BE} Characteristics with Application to Bandgap Reference Sources," IEEE J. Solid State Circuits, Dec 1980

[3] G.A. Rincon-Mora: "Voltage References: From Diodes to Higher Order bandgap Circuits," published by IEEE Press,2002

[4] B.P. Song, P.R. Gray: "A Precision Curvature Compensated CMOS Bandgap Reference," IEEE J. Solid State Circuits, Dec 1983

[5] G.M. meijer et. al.:"A New Curvature Corrected Bandgap Reference," IEEE J. Solid State Circuits, Dec 1982

[6] P.E. Allen, D.R Holberg: "CMOS Analog Circuit Design." published by Oxford University Press,2002

[7] I. Lee et. al.: "Exponential Curvature Compensated BiCMOS Bandgap References," IEEE J. Solid State Circuits, June 1993

[8] G.A. Rincon-Mora, et. al.: "A 1.1 V Current Mode and Piecewise Linear Curvature Corrected Bandgap reference," IEEE J. Solid State Circuits, Oct 1998

[9] P. Miller, D. More: "Precision Voltage References," http://www.thaler.com/thcpdf/TI_article.pdf

[10] W. Jung: "Voltage References and Low Dropout Regulators," http://www.analog.com/UploadedFiles/Technical_Articles/3752210Publication_V-Ref.pdf

[11] N. Weste, D. Harris, A.Banerjee: "CMOS VLSI Design: A Circuits and Systems Perspective" published by Pearson Education, 2005

[12] R. Gregorian, G. Temes: "Analog MOS Integrated Circuits for Signal Processing," published by John Wiley and Sons, 2004

[13] R. Baker, H. Li, D. Boyce: "CMOS Circuit Design, Layout and Simulation," published by IEEE Press, 1998

[14] K. Sanborn, D. Ma, V. Ivanov: "A Sub 1-V Low-Noise Bandgap Voltage Reference" IEEE J. Solid State Circuits, November 2007