FLASH MEMORY SENSING USING AVERAGING

by

Murugesh Subramaniam

A project

submitted in partial fulfillment

of the requirements for the degree of

Master of Science in Engineering, Electrical Engineering,

Boise State University

August 2003

The project presented by Murugesh Subramaniam, submitted for Master of Science with major in Electrical Engineering and titled "FLASH MEMORY SENSING USING AVERAGING", has been reviewed in final form, as indicated by the signatures and dates given below.

R. Jacob Baker, Advisor — Date

Joe Hartman, Committee Member — Date

Stephen A. Parke, Committee Member — Date

John R. Pelton, Graduate Dean — Date

ACKNOWLEDGMENTS

It has been an honor and a privilege to study at Boise State University. It has been a wonderful learning experience - both technically and real life- throughout my program. I would like to take this opportunity to thank everyone who made this possible: Those who inspired me, those who provided technical guidance, and those whose friendship and camaraderie made even the most difficult times more enjoyable.

First of all, I would like to express my deepest gratitude to Professor Jake Baker for his patient guidance and support during my years of graduate school here. Without his support and guidance this research could never have been completed.

I would also like to thank Dr. Joe Hartman and Dr. Stephen Parke not only for agreeing to be a part of my committee and spending their valuable time but also for their timely advice and guidance throughout my Masters program.

Last, but not the least, I would like thank my parents for their unconditional love, guidance and support. To my brother, who guided me through the last two years, I shall be ever grateful. I dedicate this to you with love.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE: OVERVIEW OF PROJECT

Abstract

A flash memory Read operation consists of sensing data from a random location in the memory array. During the Read operation, the flash cell control gate voltage is controlled through a read regulator circuit.[6] Minimizing the drain-source voltage variation will minimize the variations in cell current. Instead of sensing the voltage variation in the bit line we are sensing the current flowing in a flash memory cell. This allows for more precise measurement of the charge stored on the floating gate. The main goal of this project is to design a circuit, which does not rely on a precision current.

Project Goals

- To design a circuit to sense the drain current flowing in a flash memory cell.
- To develop a robust circuit which does not rely on a precision current.

Project Organization

This project is divided into 4 chapters.

Chapter One gives an overview of the project and the goals that need to be attained.

Chapter Two describes the basic flash memory cell structure, flash read, flash programming and flash erase operations. The NAND and NOR flash architectures are also discussed in chapter Two.

Chapter Three describes the whole project as a block diagram, also the reason behind this sensing technique. It also describes design of the individual blocks of this design.

Chapter Four covers the summary of the design.

CHAPTER TWO: FLASH MEMORY DEVICE

What is Flash Memory?

Flash Memory retains digital information under certain conditions. This retained material might be operational code or data files, or a combination of the two. The ideal memory subsystem optimizes density, preserves critical material in a non-volatile condition, is easy to program and reprogram, can be read fast, and is cost-effective for the application.[1,6]

Most flash memories are programmed with hot electron injection technology. The erasing technology is the Fowler-Nordheim tunneling off the floating gate to the drain region.[1] More detailed explanation is given in later sections. A few of the latest flash memories use Fowler-Nordheim tunneling for both Write and Erase operations. Flash is similar to EEPROM memory.[7] The principal difference is that EEPROM requires data to be written or erased one byte at a time. Whereas flash memory allows data to be written of erased in blocks. This is reason it's called flash memory. Internal state machines in the latest flash memories carry out the Write and Erase operations.

Flash Cell Structure

Flash memory is built using floating gate CMOS technology. Most of the basic flash memories have a structure as shown in figure 2.1.[1,6] The floating gate poly acts as a storage element. The floating gate is isolated from the control gate by inter-poly dielectric and from the substrate by tunnel oxide.[1] The floating gate can store a value

by virtue of its isolation from both the substrate and the control gate.[7] The control gate is the word line and controls the value stored and hence the name.

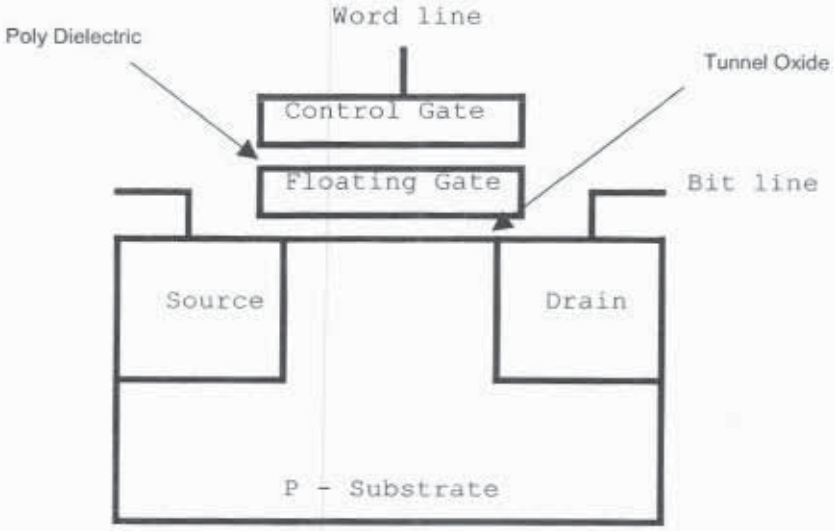


Figure 2.1. Flash Cell Structure

The drain is connected to the bit line and the source is either grounded or connected to the next cell depending on the type of flash memory. In a NOR flash, the source is grounded and in a NAND flash the cells are connected in series. This will be discussed further in later sections.

Flash Programming Operation

In one scenario a write operation, approximately 12v is applied between the control gate and source. Around 6v is applied between the drain and the source.[1] The potential difference between source and drain generates hot electrons, which move from source to drain. The higher voltage at the gate attracts electrons towards the floating gate from the substrate. After a length of time enough electrons are accumulated in the floating gate to change the cell from erased state "1" to programmed state "0". This accumulation of e^- charge causes the threshold voltage of the flash memory cell (the MOSFET formed between the control gate and substrate) to increase (so the cell is never

conducting). When reading the cell the large threshold voltage causes the cell to remain OFF.

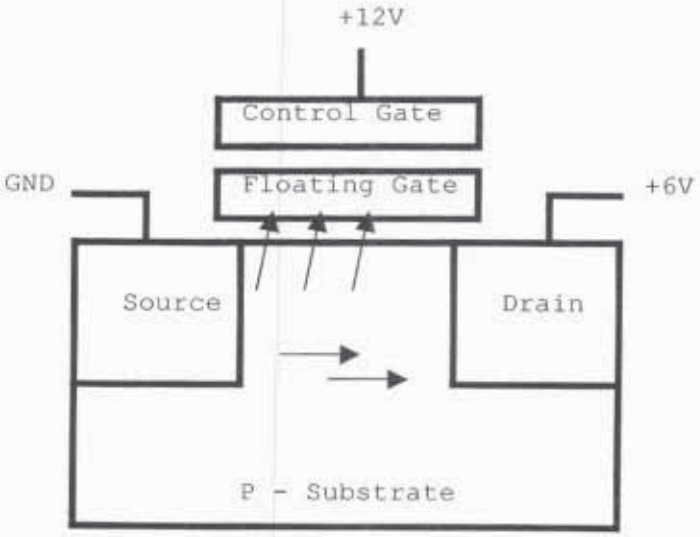


Figure 2.2. Flash Programming Operation

This method of programming is known as "Hot Electron Injection".[1,8] The voltage applied to the gate and the drain depends on the process technology used. These high voltages are obtained by using charge pump in the circuit. The programmed cell is represented by logic "0" during a Read operation.

Flash Erase Operation

The flash erase operation uses Fowler-Nordheim tunneling mechanism to remove charge from the floating gate and brings it to erased state.[1,8] This creates an electric field across the floating gate and the source through the thin oxide layer. Due to high voltage at the source, as compared to the voltage at the gate, electrons are attracted to the source. The erased cell is represented by a logic "1" during a Read operation. The threshold voltage of the erased Flash cell decreases. This keeps the MOSFET from turning OFF during a read operation.

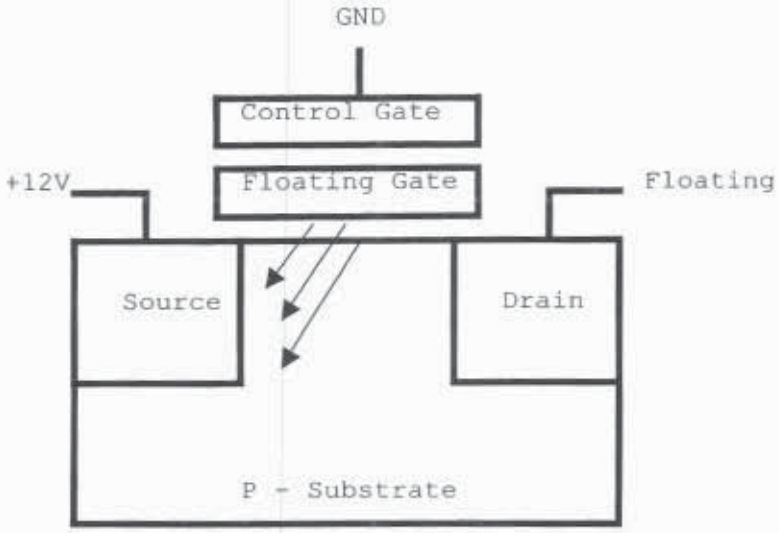


Figure 2.3. Flash Erase Operation

Flash Read Operation

During the flash read operation, around 5v is applied to both the gate and the drain, which is strong enough to turn *ON* an erased cell (the threshold voltage is < 5V); but too weak to turn *ON* a programmed cell (the threshold voltage is > 5V). When the 5V is applied to the control gate of an erased cell, the cell turns *ON*. This when detected by a sense amplifier produces a logic "1".

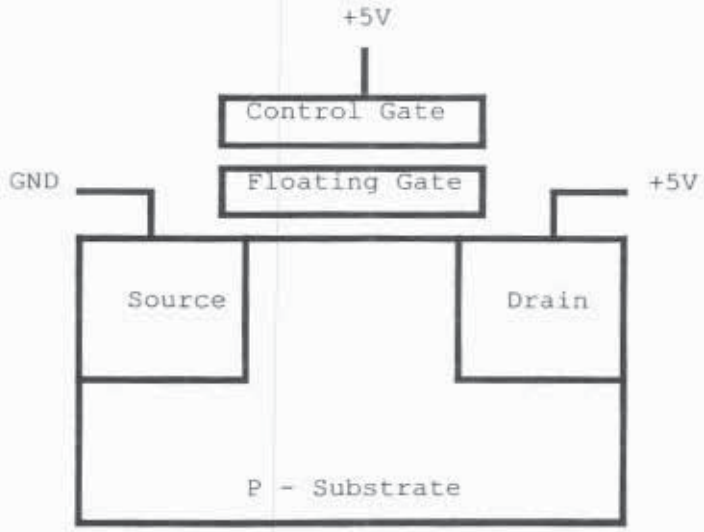


Figure 2.4. Flash Read Operation

When the 5V is applied to the control gate of a programmed cell, the cell remains *OFF*. This when detected by a sense amplifier produces logic "0".

NOR and NAND Architecture

NOR Architecture

The NOR architecture is the simplest Flash architecture and is the most straight forward to program and erase. However, because of large layout area, it is not suitable for large data storage. NOR architecture can be used as a substitution for EEPROM.[1] The basic architecture of a NOR flash is shown in figure 2.5. In NOR architecture, each cell is connected to the bit line. When reading a value, the word line corresponding to the cell to be read is driven high.

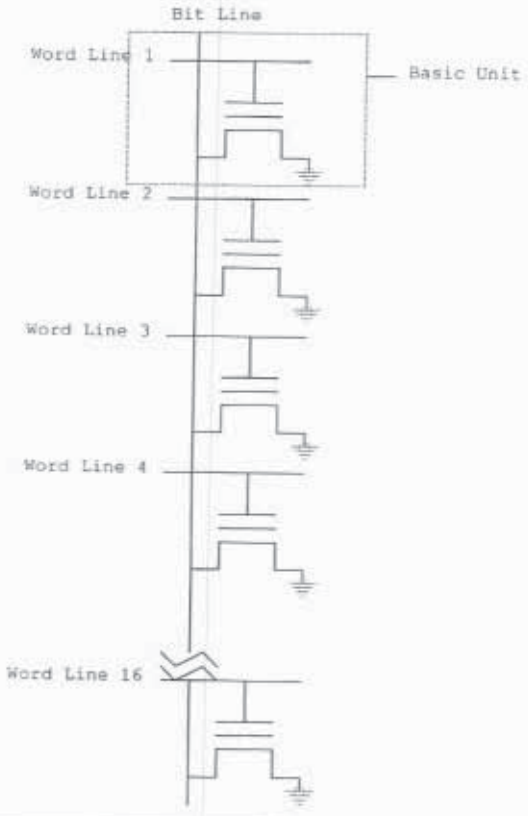


Figure 2.5. NOR Flash Architecture

If the cell were programmed, then the cell wouldn't turn *ON* and the bit line voltage would stay the same. If the cell were erased, then the cell would have turned *ON* and discharged the bit line.

NAND Architecture

NAND architecture is known for its high density (smaller layout area) and is capable of carrying out high-speed Programming, Read and Erase operations. The basic architecture of NAND is shown in figure 2.6.[1]

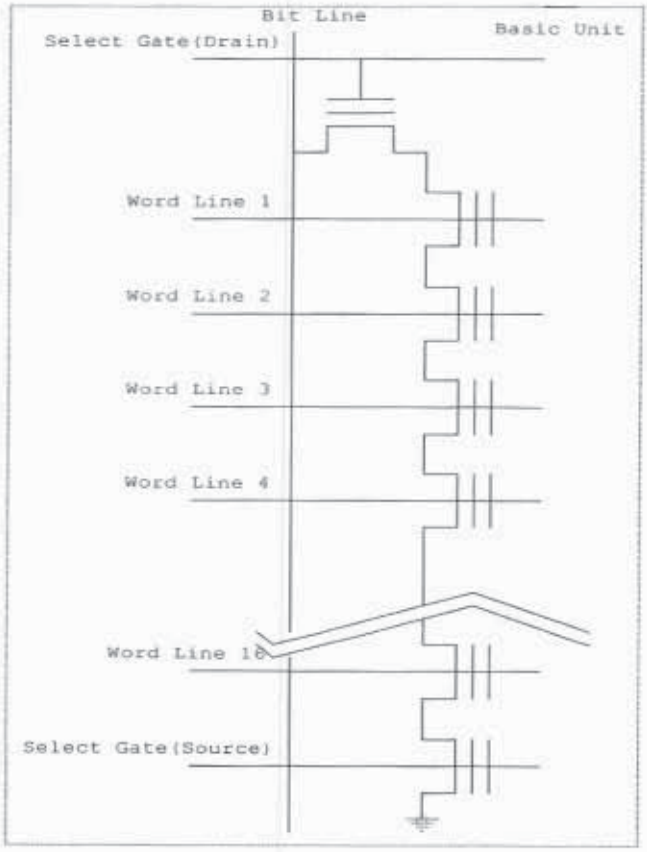


Figure 2.6. NAND Flash Architecture

In the NAND architecture the select gate source and select gate drain are turned *ON* connecting the cell to bit line and ground respectively. All the cells, except the one to be read, have their word lines (floating gates) driven to a high voltage so that they all

conduct.[1] The wordline of the cell to be read is driven to a lower voltage, say +5v.[1] If the cell were programmed, then it will not turn *ON* and there would not be a path from the select gate drain to the select gate source. Hence, the voltage at the bit line would stay high. If the cell were erased, then the small voltage applied to the word line would be sufficient to turn *ON* the transistor. This opens a path from the select gate drain to the select gate source. Hence, the bit line voltage would go to zero.

The sense amplifier we are developing will be used if the Flash cell is not perfect (i.e., if cell is *OFF* – Zero drain-source current and if cell is *ON* – Conducting significant current). Further, we hope to develop this scheme so that more than 2-levels can be stored in a Flash memory cell.

CHAPTER THREE: NEW SENSE AMPLIFIER

System Block Diagram

The block diagram of the sensing circuit is shown in figure 3.1. Here Rbit is used to model the bit resistance. The Rbit also models the Flash cell current. Here we are designing a circuit to sense any change in current through the bit resistance. The voltage on the bit line determines the value stored in a flash memory cell. In traditional sensing 0v left on the bit line, indicates that the particular cell being read is programmed as a "0". If Vbit has the precharged voltage, it shows that the particular cell being read is a "1".

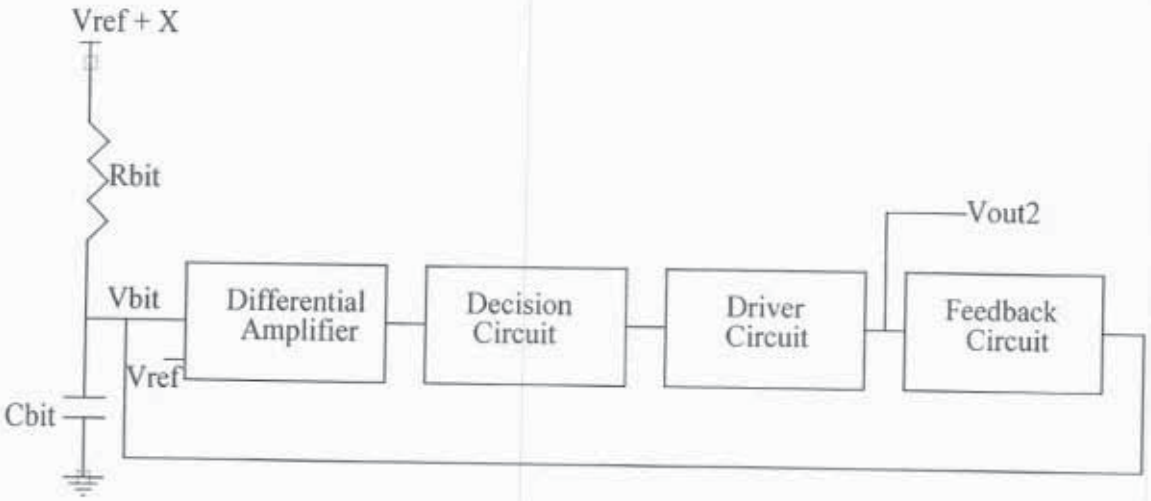


Figure 3.1. Circuit Block Diagram

In some cases the voltage at the Vbit need not go all the way to zero to show a "0" being programmed.

In the block diagram shown in figure 3.1, the current flowing in the Flash cell will be sensed. Here we are going to sense the change in current on the bit resistance Rbit. By

sensing the current we can determine more precisely, the value stored in the flash memory cell.

The first block shown in figure 3.1 shows a differential amplifier. One of the inputs has a reference voltage (2.5 V) and the other input has the Vbit voltage. The differential amplifier produces an output current dependent on the voltages at the gates of the differential amplifier transistors. The transistor with higher gate voltage will source a higher current. The differential amplifier is a preamp for the decision circuit.

The second block shown in figure 3.1 is the decision circuit. The decision circuit produces an output voltage depending on the current that was sourced by the differential amplifier. This voltage difference is of the order of a few millivolts. This signal is applied to the driver circuit.

The third block in figure 3.1 is the driver circuit where the signals are driven to logic levels. Special type of inverter is used for the driver circuit. The output of the driver circuit represents any change in current on the bit resistance. This signal is applied to the feedback circuit.

The fourth block in figure 3.1 is a feedback circuit. The current flowing through Rbit is accumulated on the bit line capacitance Cbit, and if the sensing circuit is working the feedback circuit will feedback a current equal and opposite of the current flowing in Rbit. This acts to hold the bit line voltage at a constant of Vref. The feedback circuit was designed using a switched capacitance resistor circuit.

Current Source

Many applications in CMOS require a constant current source. Differential amplifier use current sources for proper operation. For correct working of the differential pair, we would need a constant current. Here the basic concept of current mirror is discussed followed by the design of the current source used in the differential amplifier.

Current Mirror

Current mirrors and current sources work on the principle that identical devices with equal gate-to-source and drain-to-source voltages carry equal drain currents.[3]

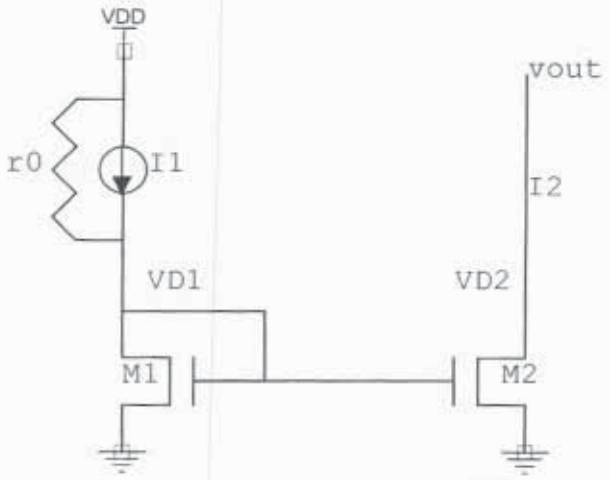


Figure 3.2. N-Channel Current Mirror

In figure 3.2, we can see that M1 is forced to carry current I1. This sets up the VD1 of M1. Since M1's gate voltage is equal to M2's, given both M1 and M2 are in saturation, both will carry the same current. The current in M1 is mirrored over to M2. Thus M2 can sink only the amount of current carried through M1 or less better.

Beta Multiplier

The Beta multiplier is the current reference used in the differential amplifier design. This circuit provides biasing the differential amplifier, which is explained in later sections.

Working

Beta multiplier's schematic is shown in figure 3.3.[2] The width of MB2 is K times larger than MB1 and L1 = L2 to provide $\beta_2 = K.\beta_1$.[2]

$$V_{GS1} = V_{GS2} + IR \qquad (3.1)$$

The gate to source voltage of M1 will be equal to the gate to source voltage of M2 plus the voltage across the resistance.

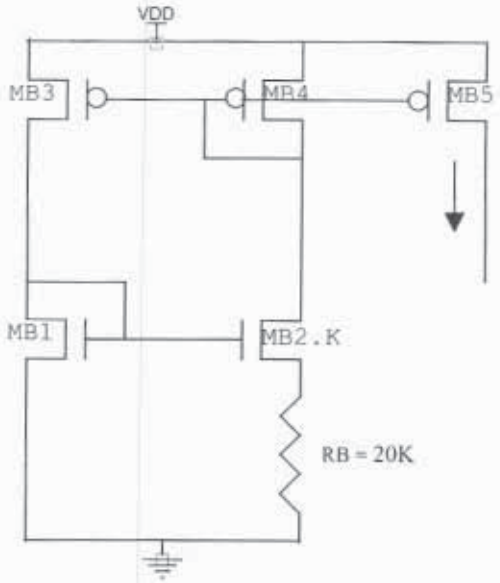


Figure 3.3. Basic Beta Multiplier

Drain current of MOSFET in saturation is,

$$I_D = \frac{\beta_1}{2}(V_{GS} - V_{THN})^2 \qquad (3.2)$$

$$V_{GS1} = \sqrt{\frac{2I_D}{\beta}} + V_{THN} \qquad (3.3)$$

$$V_{GS2} = \sqrt{\frac{2I_D}{k.\beta}} + V_{THN} \qquad (3.4)$$

From Equation 3.1,

$$V_{GS1} = V_{GS2} + IR \qquad (3.5)$$

$$V_{GS2} = V_{GS1} - IR \qquad (3.6)$$

Substituting (3.6) in (3.4),

$$V_{GS1} - IR = \sqrt{\frac{2I_D}{k.\beta}} + V_{THN} \qquad (3.7)$$

Using (3.4) in (3.7),

$$\sqrt{\frac{2I_D}{\beta}} + V_{THN} - IR = \sqrt{\frac{2I_D}{k.\beta}} + V_{THN} \qquad (3.8)$$

$$I = \frac{2}{R^2 \beta_1}\left[1 - \sqrt{\frac{1}{k}}\right]^2 \qquad (3.9)$$

Here we are generating a 10μA current source for the differential amplifier pair. In this case, the design variables that determine the current are the ratio of the gate widths 'k' and the value of the passive resistor 'R'. In this design, the ratio of the widths 'k' is taken as 4. i.e., W2 = 4 W1. For I = 10μA and k = 4, we can calculate the resistor value. Substituting in equation 3.9, we have

$$10\mu A = \frac{2}{R^2 K_{PN}} \frac{L1}{W1}\left(1 - \sqrt{\frac{1}{4}}\right)^2 \qquad (3.10)$$

$$R^2 = \frac{2}{10\mu A.50\frac{\mu A}{V^2}\frac{5}{2}0.25} \qquad (3.11)$$

$$R = 20K\Omega$$

We can use cascode MOSFETs to reduce the influence of output resistance [2] Cascoded beta multiplier will have a higher output resistance compared to the basic beta multiplier. The design of cascode beta multiplier is shown in figure 3.4.

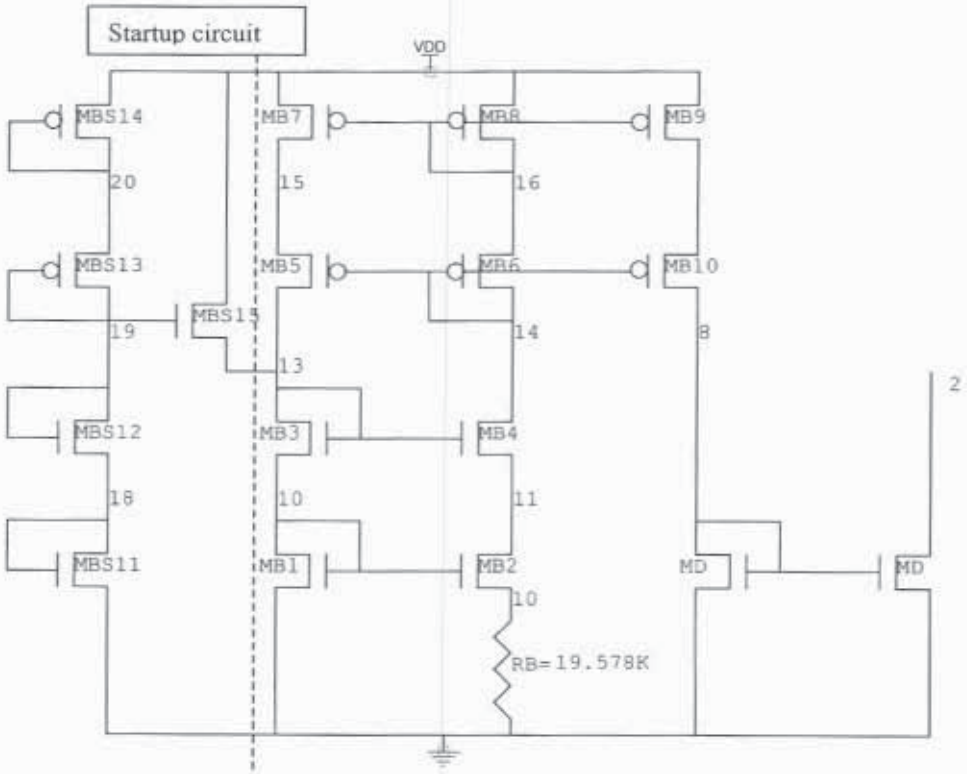


Figure 3.4. Cascoded Beta Multiplier

The startup circuit was used to keep node 13 from dropping to lower voltage and turning *OFF* transistors MB4 and MB3.

Simulations

The transient and DC analysis for the beta multiplier in figure 3.4 are shown in figure 3.5 and 3.6 respectively.

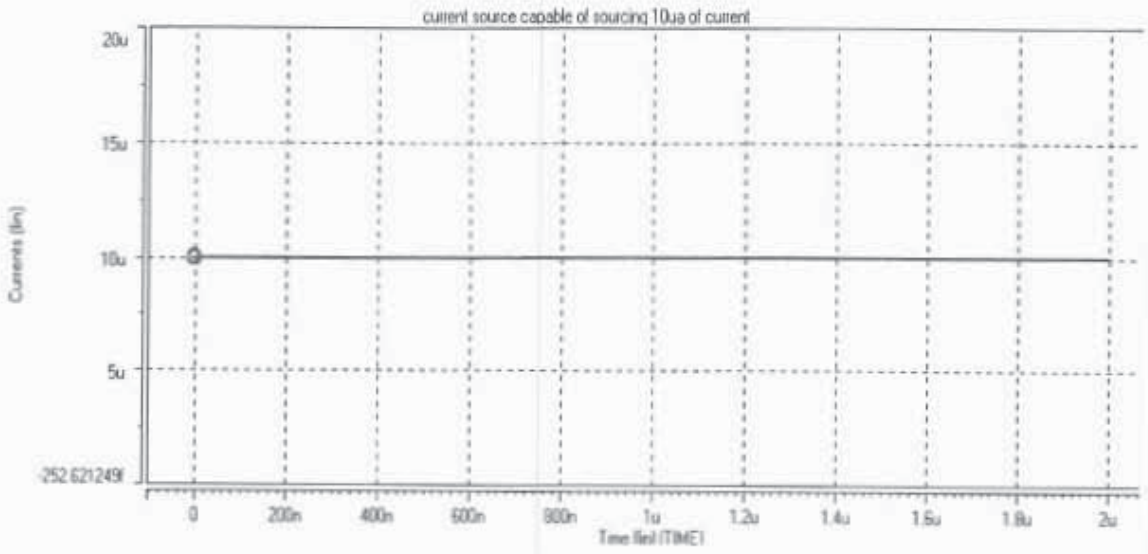


Figure 3.5. Beta Multiplier Transient Analysis

Figure 3.5 shows the nominal operating current of 10uA. Figure 3.6 shows that as the VDD of the beta multiplier is swept from 0 to 5 volts, the current varies from zero to 10μA. For CMOS circuits the power supply must be at least equal to the sum of threshold voltages of NMOS and PMOS transistors. We can see that the circuit operates around 1.5v, which is approximately equal to sum of threshold voltages of the NMOS and PMOS transistors. In the simulation, an 'R' value of 19.578K was used for the calculated value of 20K.

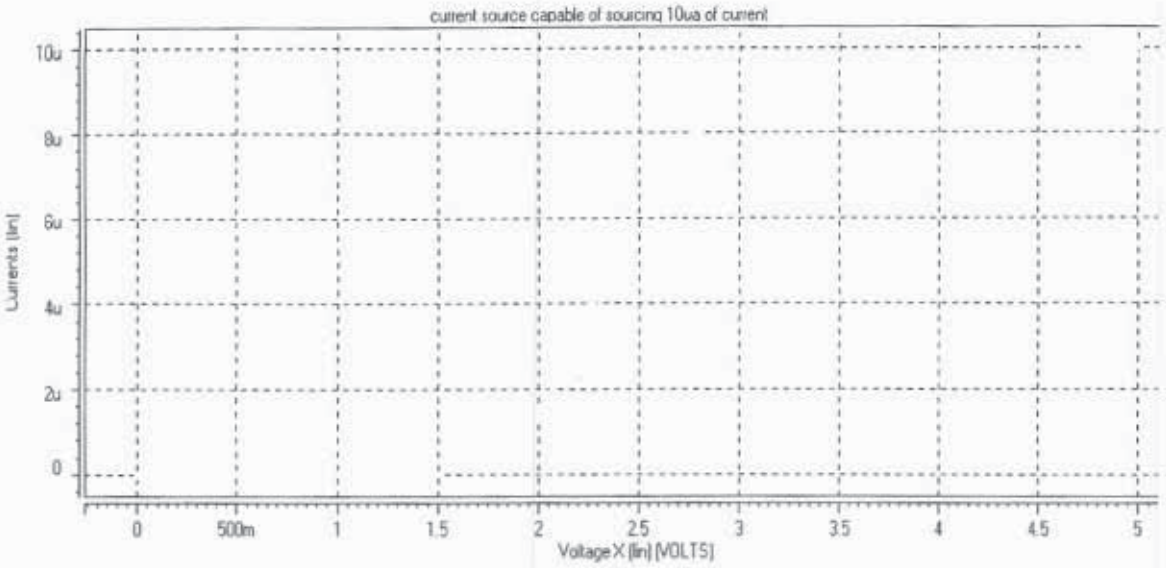


Figure 3.6. Beta Multiplier DC Analysis

Differential Amplifier

Differential pair is one of the most widely used circuit building block. The input stage of every op-amp is a differential pair. Differential pairs have two matched transistors with their source shorted together and connected to a current source. The current source used in this design was beta multiplier discussed in the previous section. These devices must always stay in the active region.

Working

The circuit shown in figure 3.7 is a differential amplifier with active loads.[4] The currents through MD1 and MD2 depend on the voltage applied to their respective gates. Higher voltage applied to the gate would cause more current to flow through the drains of MD1 and MD2. However the maximum current that can be sourced by the differential amplifier depends on the current source of the differential amplifier. These currents are mirrored over to MD5 and MD6. The gain in this stage can be improved by increasing the size of MD5 and MD6 proportional to MD3 and MD4. Making the widths of MD5

and MD6 'N' times larger would cause 'N' times the current of ID1 and ID2 to flow through them (because $\beta_5 = N\beta_1$; ID α β and ID5 α N ID1).

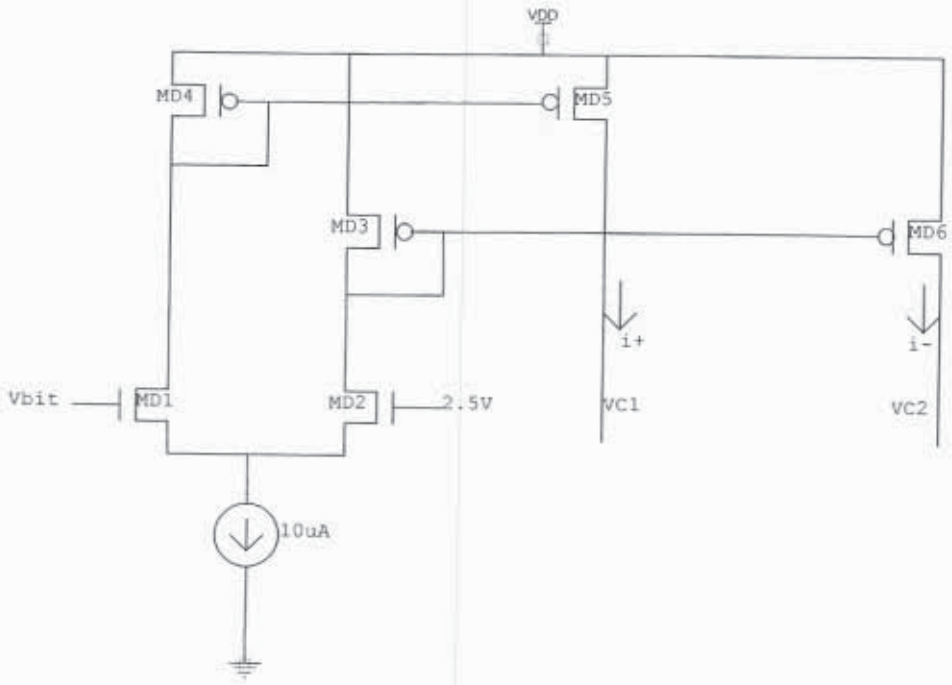


Figure 3.7. Differential Amplifier with Active load

Simulations

The differential amplifier shown in figure 3.7 was simulated with one input tied to a reference voltage; Vref of 2.5V and the other swept from 2.4V to 2.6V. The currents at nodes VC1 and VC2 are shown in figure 3.8. The reference voltage is supposed to be equal to the precharged bit line voltage. Here we are assuming that the precharged bitline voltage be 2.5V.

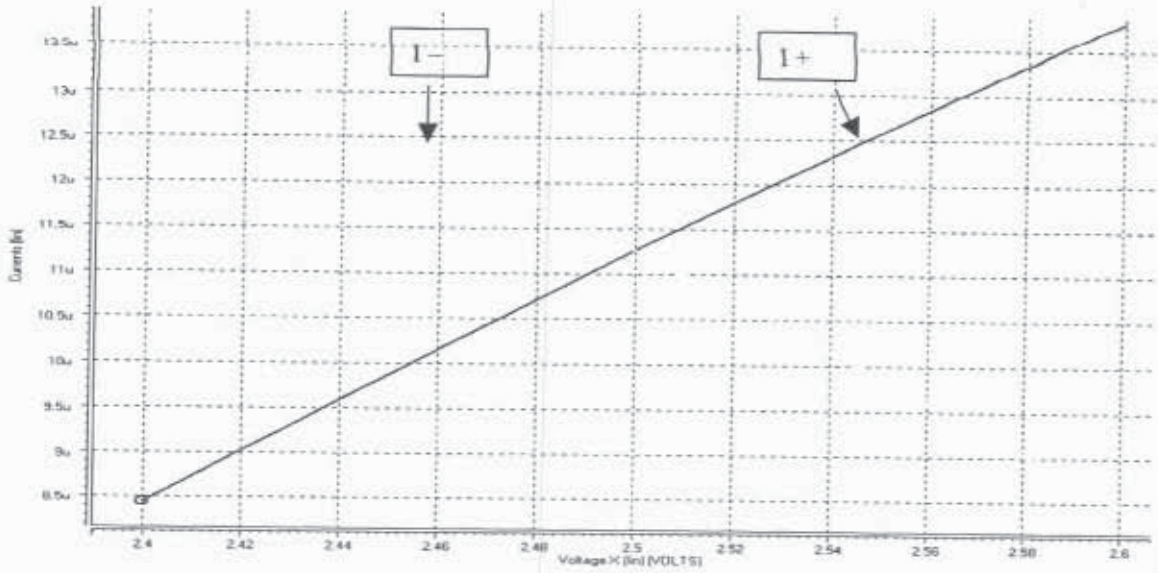


Figure 3.8. Current at nodes VC1 and VC2

When we sweep Vbit voltage from 2.4v to 2.6 volts, we can see that the current at VC1 and VC2 (I+ and I – respectively) are equal when Vbit is 2.5v. As the Vbit voltage increases, we can see that the current through VC1 (I+) also increases. For a 100mV difference between the inputs of the differential amplifier, we get a 5.5uA difference between the currents I+ and I–.

Decision Circuit

The decision circuit is an important part of the design. This is used to amplify the small voltage level change that occurs due to the current from the differential pair. The difference between the voltage VC1 and VC2 is in the order of a few millivolts.[4] The decision circuit enables us to detect this small voltage change and determine the value stored in the flash memory cell.

Working

The decision circuit is shown in red in figure 3.9. Depending on the voltage on Vbit, i+ would be higher than i– (or) vice versa.

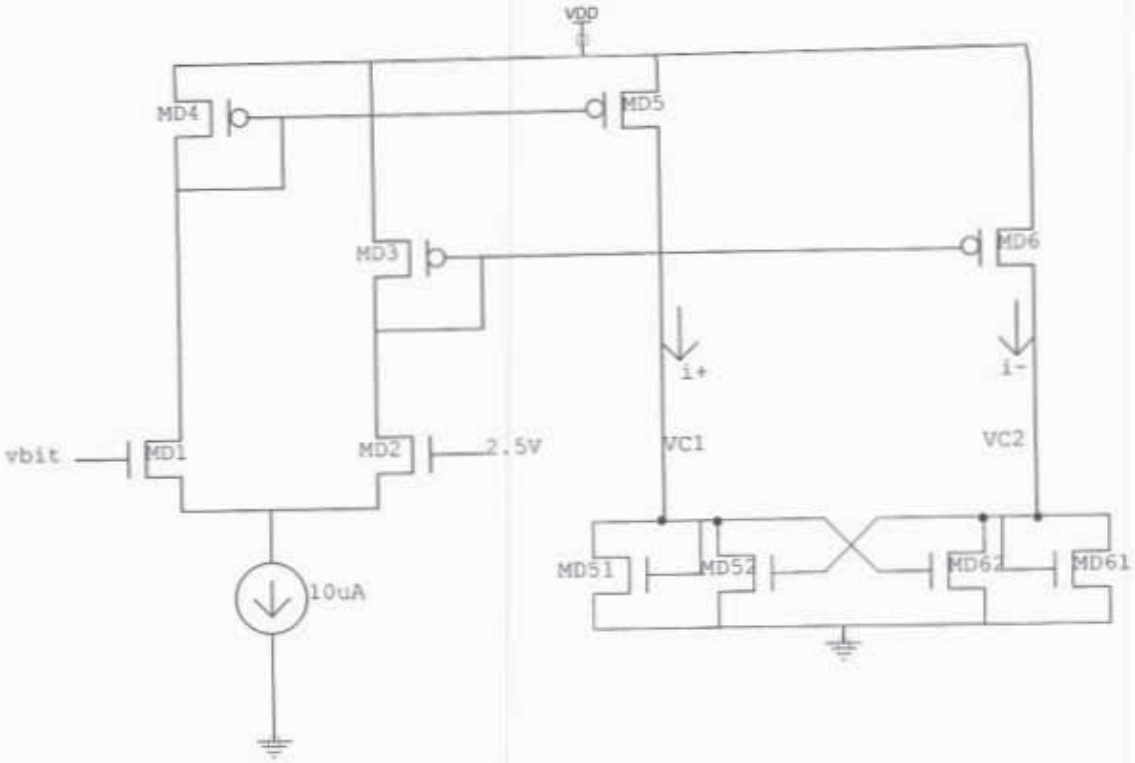


Figure 3.9. Decision circuit

When i+ is higher than i-, the voltage VC1 would be higher than VC2 since they are driving the same load. This causes MD51 and MD62 to stay *ON*. MD52 and MD61 are *OFF*. When this condition prevails, VC2 is approximately 0v and VC1 will be

$$i+ = \frac{\beta}{2}(V_{GS51} - V_{THN})^2 \qquad (3.12)$$

$$V_{GS51} - V_{THN} = \sqrt{\frac{2i+}{\beta}} \qquad (3.13)$$

$$V_{C1} = V_{GS51} = \sqrt{\frac{2i+}{\beta}} + V_{THN} \qquad (3.14)$$

When i– goes higher than i+, the voltage at VC2 would be higher than the voltage at VC1. This causes MD61 to turn *ON* stronger and this causes MD52 to turn *ON*. As MD52 turns *ON* stronger, it pulls off all the current causing MD51 to turn *OFF*. The

switching takes place when the voltage at VC2 equals the threshold voltage of MD52 and i+ equals i-.

Simulations

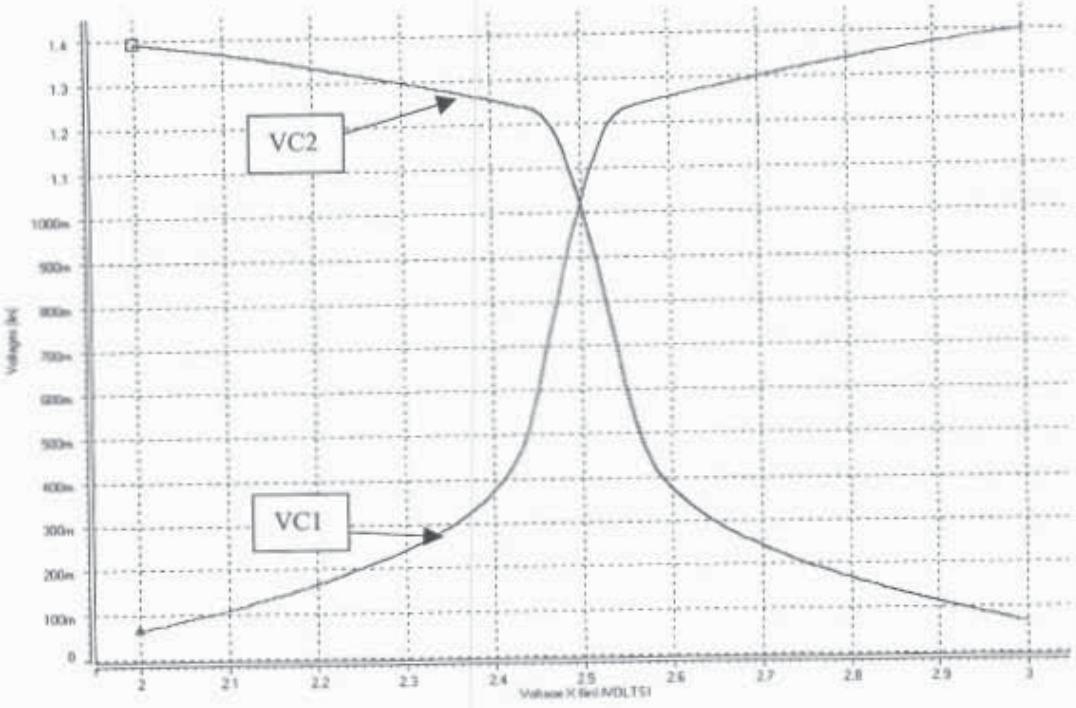


Figure 3.10. Decision Circuit Output

The voltages at node VC1 and VC2 are shown in figure 3.10. We can see that the voltage at VC1 is higher than that of VC2 when Vbit is greater than 2.5v. For a 100mV difference in the differential pair input, the voltage difference between VC1 and VC2 is 900 mV. This voltage, being very small, has to be amplified before any decision can be taken as to which of the input is larger.

Driver Circuit

Working

The driver circuit is a special kind of inverter used for switching point adjustment. The basic circuit diagram of the driver circuit is shown in figure 3.11.

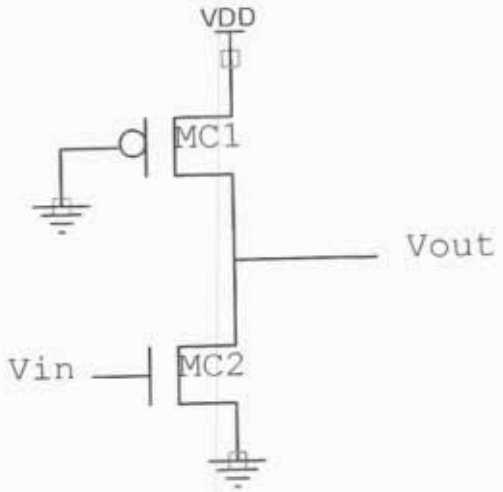


Figure 3.11. Driver Circuit

The transistor MC1 is *ON* all the time. This is because the gate of the PMOS transistor (MC1) is tied to ground. This will cause the output to be pulled high all the time. The input to this circuit is given to the gate of transistor MC2. The input capacitance of this circuit is lower when compared to that of a regular inverter. The maximum output voltage that can be obtained is VDD. However the output node will not go all the way to zero. This is because of the fact that there is a path from Vout to VDD through MC1. When Vin is strong enough to turn on MC2 it will pull Vout low.

Simulations

The input voltage Vin is swept from 0 to 5v in steps of 0.1v. We can see from figure 3.12 that as the input voltage goes above 1v it is strong enough to turn on transistor MC2. Turning on MC2 causes a path from Vout to ground pulling Vout towards ground. Here we have to note that Vout cannot go all the way to zero volts because there is always a path from Vout to VDD. We can see that very low switching point voltage of 1V was achieved with this circuit.

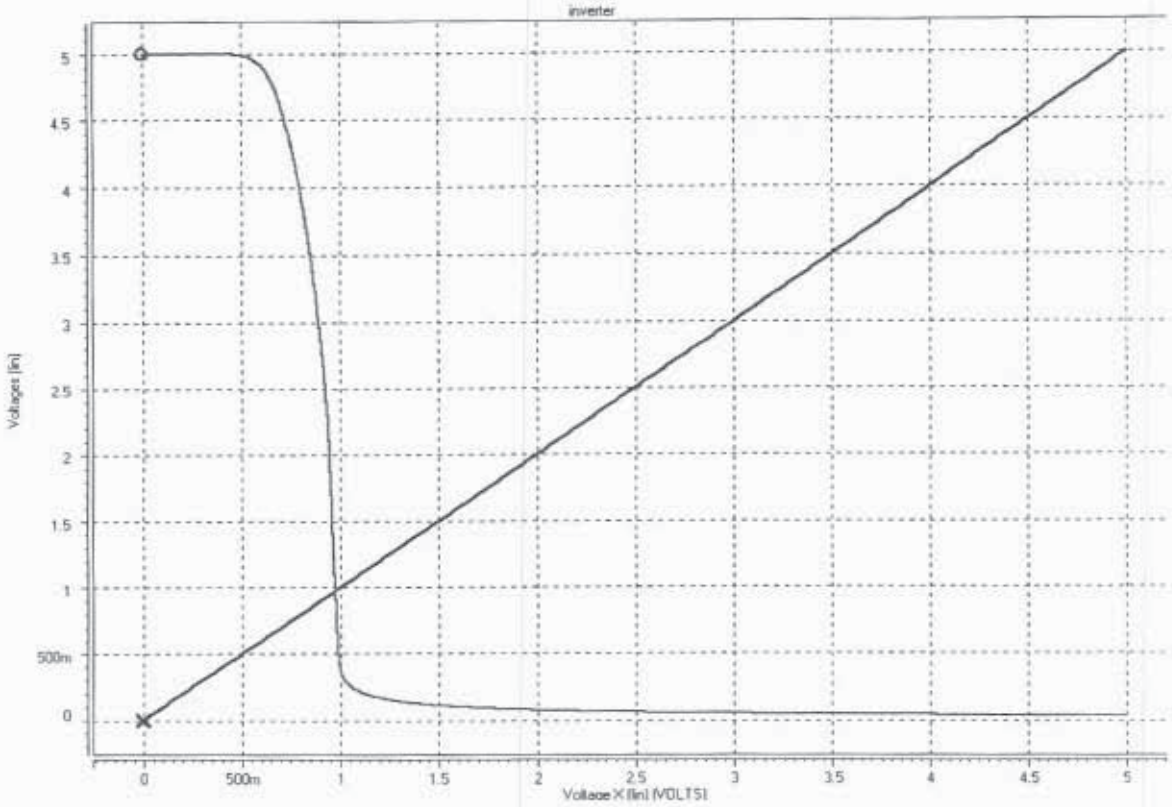


Figure 3.12. Driver Circuit Simulation

Switched Capacitor Circuit

Working

Switched capacitance circuits are used when we have to realize a high precision, large resistor value.[5] The basic circuit for a switched capacitor is shown in figure 3.13.

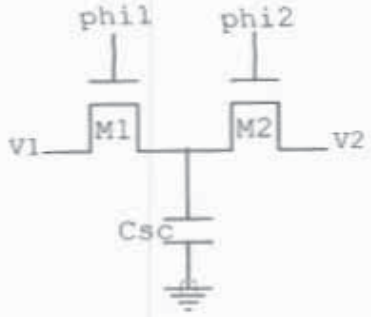


Figure 3.13. Switched Capacitor

PHI1 and PHI2 are two non-overlapping clocks of a particular frequency. M1 and M2 act as switches. Due to non-overlapping clock cycles, only one of the two MOSFETS will be *ON* at any particular time. The charge on Csc, if V1 and V2 are not equal, is Csc (V1 – V2). The average current transferred at a given time will be:

$$I = \frac{Q}{T} \qquad (3.15)$$

$$I = \frac{C_{SC}(V1 - V2)}{T} \qquad (3.16)$$

Also,

$$I = \frac{(V1 - V2)}{R_{SC}} \qquad (3.17)$$

$$\Rightarrow \frac{C_{SC}(V1 - V2)}{T} = \frac{(V1 - V2)}{R_{SC}} \qquad (3.18)$$

$$R_{SC} = \frac{T}{C_{SC}} = \frac{1}{f_{CLK} C_{SC}} \qquad (3.19)$$

Therefore, we could get a large resistor value by using appropriate capacitance value and clocking the switches at a particular frequency. A precise charge can be removed using the switched capacitor circuit.

Analysis

Suppose we want to use a very large resistor in the range of 5Meg, it would take a lot of layout area using a diffused or implanted resistor. Switched capacitor can be used to solve this problem. Consider the voltage divider circuit shown in figure 3.14. M1, M2 and Cf together will act like a resistor of a particular value.

For R = 5Meg and Cf = 20fF we can calculate PHI1 and PHI2

$$R_{SC} = \frac{1}{f_{CLK} C_{SC}}$$

$$f_{CLK} = \frac{1}{R_{SC} C_{SC}} = \frac{1}{5Meg.20f} = 10MHz \qquad \rightarrow T = 100\ nS.$$

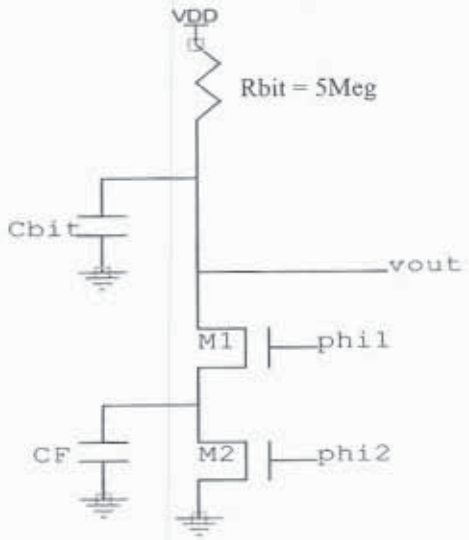


Figure 3.14. Switched Capacitor Example

Simulations

The Rbit being 5Meg and the Rsc realizing a 5Meg resistor, we have a voltage divider here. With VDD at +5v the output voltage vout is,

$$Vout = \frac{R_{sc}}{R_{bit} + R_{sc}} V_{DD}$$

$$Vout = \frac{5Meg}{5Meg + 5Meg}.5 = \frac{1}{2}.5 = 2.5v$$

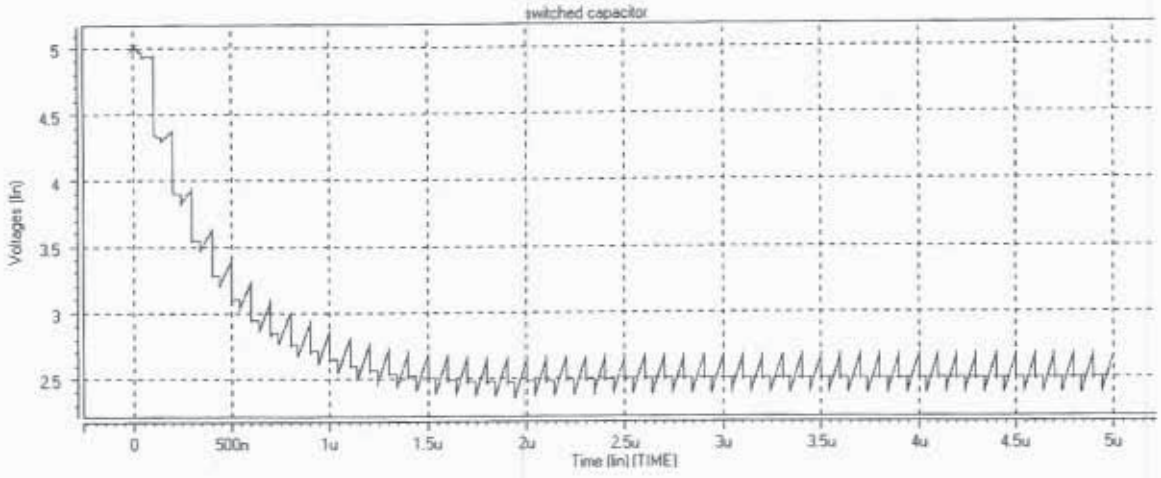


Figure 3.15. Switched Capacitance Resistance

The simulated result seen in figure 3.15 shows an output of 2.5v.

Building the Circuit

This project focuses on detecting any small changes in the bit line current resulting from the flash memory device. In a NAND flash memory cell, the select gate drain and select gate source are turned *ON* so that there is a path from the drain to the source. All the word lines are kept high except the one to be programmed. Around +5v is applied to the cell that has to be read.[1] If that cell were programmed, then it wouldn't turn *ON* due to higher threshold voltage caused by the excess charge trapped at the floating gate. This will cause the bit line voltage to stay high, as there is no path to ground. If the cell were erased, then the +5v applied to its gate would be sufficient enough to turn *ON* the floating gate MOSFET. Once the FET is *ON*, there would be a path from drain to source. This causes a variation of current through the bit line. Any small change in bit line current has to be detected by the sense amplifier.

This project concentrates on building a circuit that can average the current through the bit line and produce an output signal from which the flash cell current can be determined.

We know that initially any change in the bit line current when sensed by the differential amplifier produces an output current. This output current depends on the difference between the voltage at Vref and Vbit nodes. The decision circuit amplifies the small voltage level change that occurs due to the current from the differential pair. The simulation is shown in figure 3.18.

As we sweep the vbit voltage from 2 to 3v, we can see that switching takes place between VC1 and VC2. The millivolt difference between VC1 and VC2 is now in the order of volt.

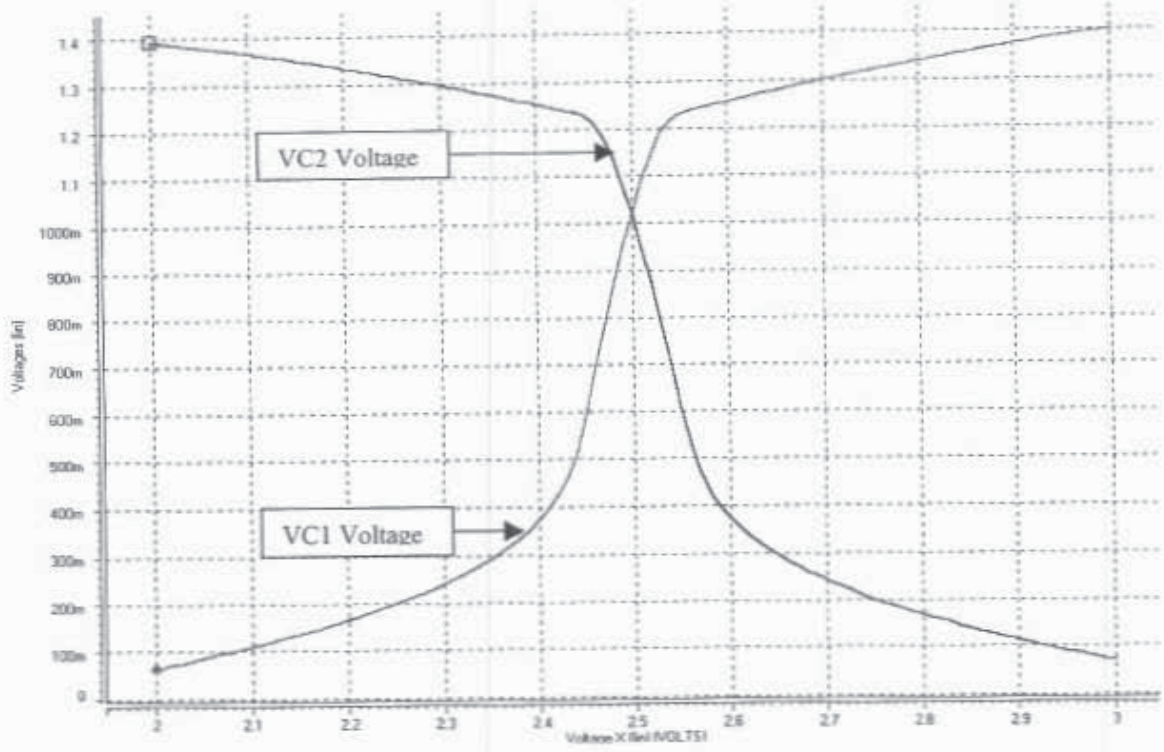


Figure 3.16. Decision Circuit Output

Now we will have to sense the voltage difference and produce rail-to-rail output voltage. The next stage we are using here is a driver circuit. This circuit can be used to sense the volt difference and produce a logic output.

Driver Circuit

The driver circuit works as an inverter. We have to sense any change in Vbit voltage and produce logic high or low at the output. As the Vbit voltage increases, we can see from figure 3.16 that the voltage at VC1 increases and the voltage at VC2 decreases. Since we want a high at the output when Vbit exceeds the reference voltage we want to connect VC2 node to the driver circuit. The size of the MOSFETS can be adjusted to set the switching point voltage at 1v.

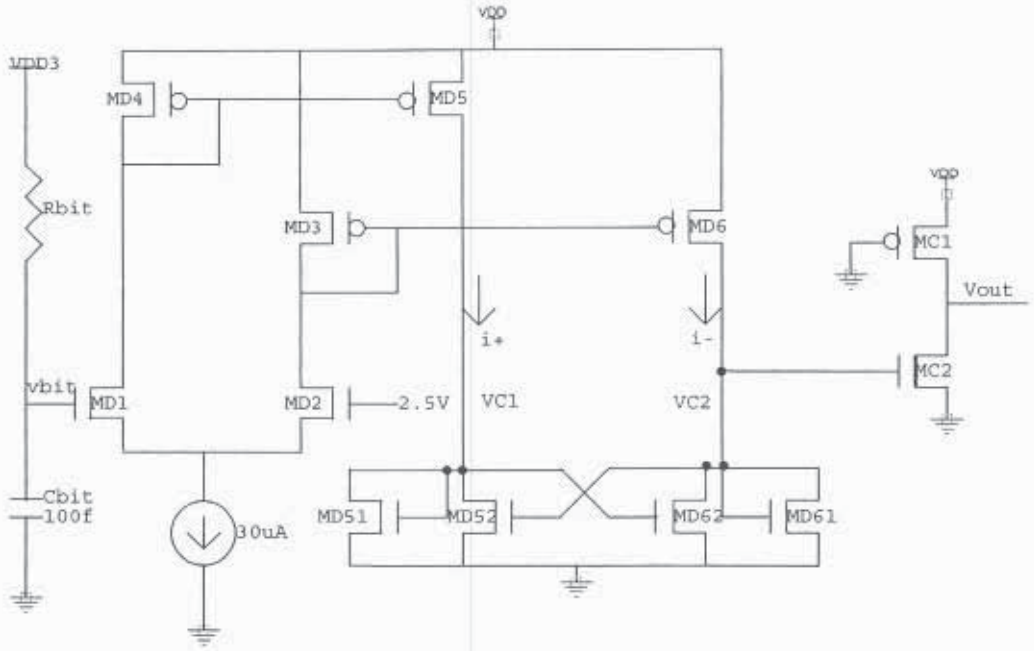


Figure 3.17. Driver Circuit

The Vbit voltage is swept from 2.4v to 2.6v. The simulation in figure 3.18 shows that as the VC2 voltage goes low with an increase in Vbit voltage above the reference voltage, the output of the driver circuit is driven high.

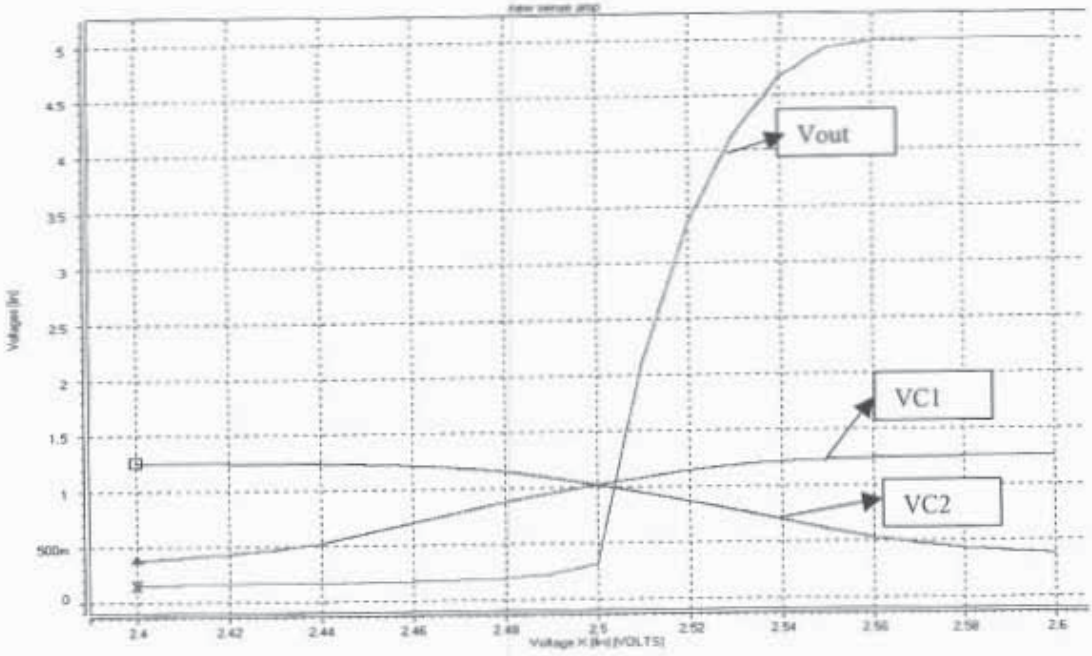


Figure 3.18. Driver Circuit Output

The problem with the above circuit is that Vout will not go all the way to zero. This is because the PMOS transistor is *ON* all the time. This restricts the Vout from going all the way to zero volts. The next circuit is a feedback circuit. Since the feedback circuit is connected to the bit line through an NMOS switch, it is better to have the Vout signal varying from VDD to zero. If not the feedback circuit will not be completely isolated from the bit line when required to. This leads to adding two more inverters after the driver circuit to make sure that the signals vary from rail to rail. The circuit with the inverter is shown in figure 3.19.

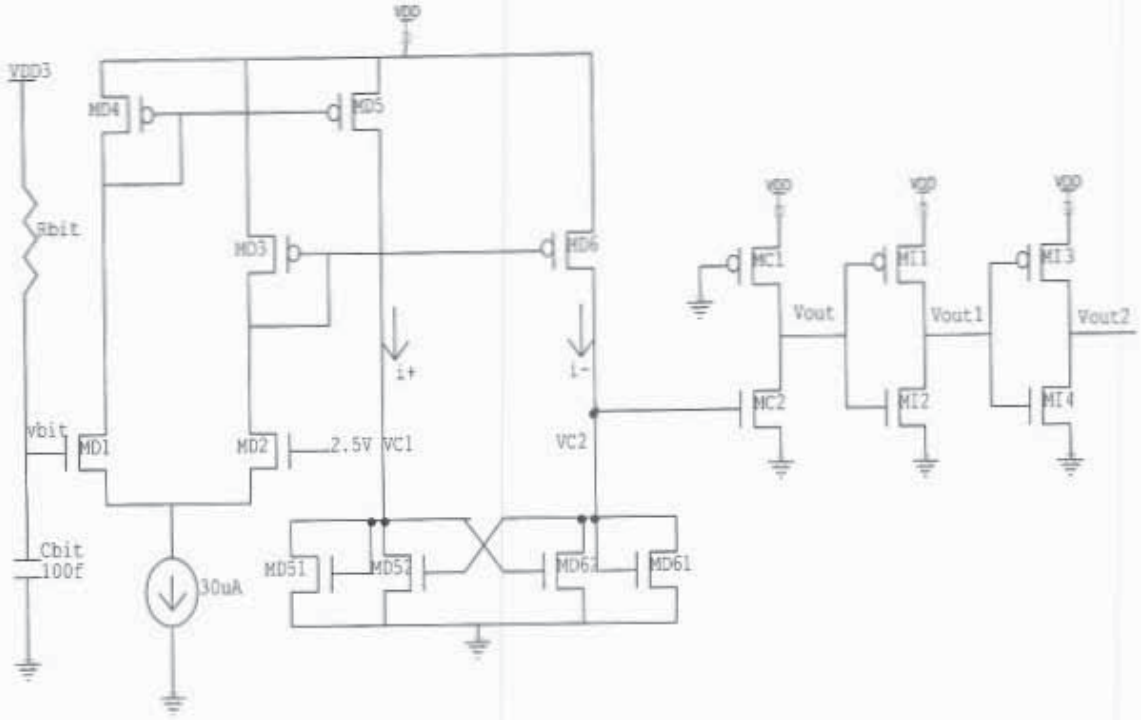


Figure 3.19. Driver Circuit with Buffer

We can see from figure 3.20 that with the output buffers the signal from the driver circuit is now going all the way form VDD to zero. The signal Vout2 is the output of the second inverter. The next part of the circuit will be the feedback circuit.

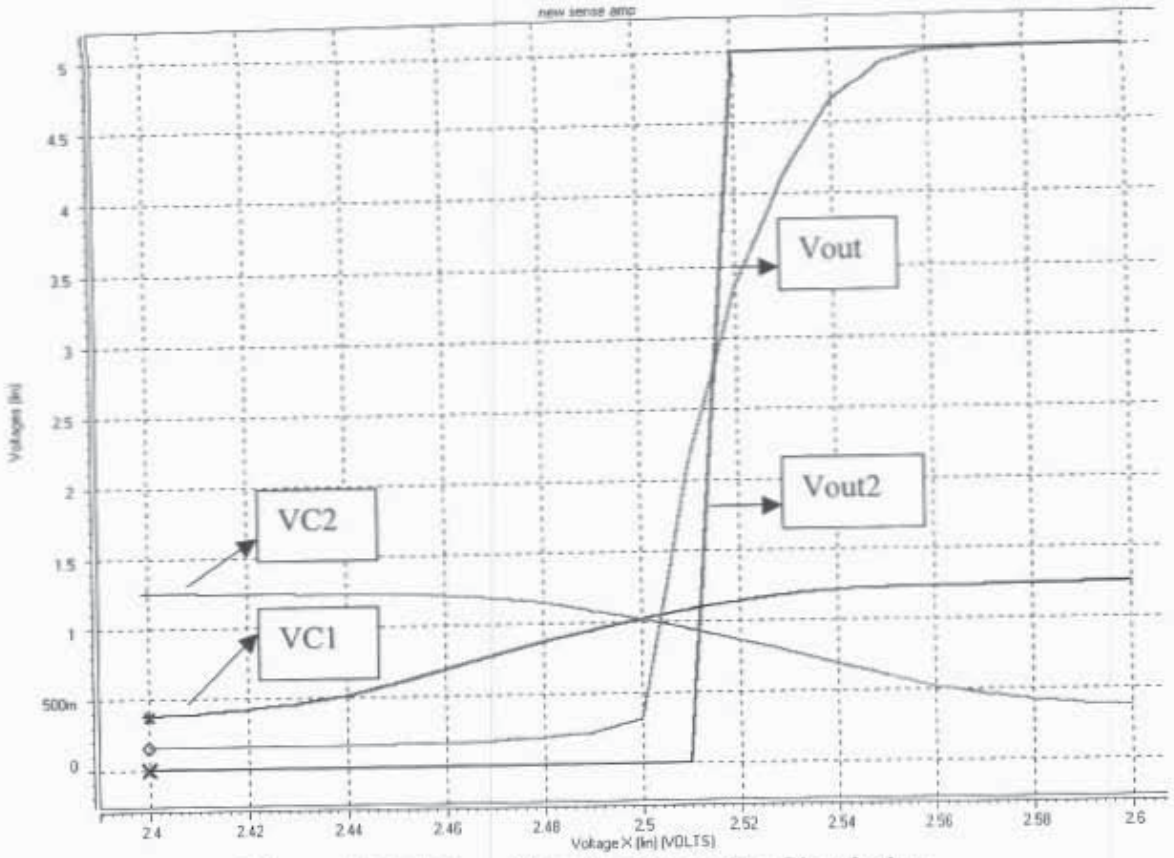


Figure 3.20. Driver Circuit with Buffer Simulation

Feedback Circuit

The last part of the design is the feedback circuit, which can sense changes in current through the bit line. The feedback circuit's input is taken from the output of the buffer. The buffer output voltage Vout2 would go high when Vbit goes above the reference value. The feedback circuit is made of switched capacitor resistor. The capacitor value and the frequencies of PHI1 and PHI2 can be adjusted to remove a precise amount of charge from the vbit node.

The feedback circuit is controlled by the buffer circuit output as shown in figure 3.21. The feedback circuit when turned *ON* produces a current opposite to the current flowing through the bit line. This causes the Vbit node to stay approximately at Vref.

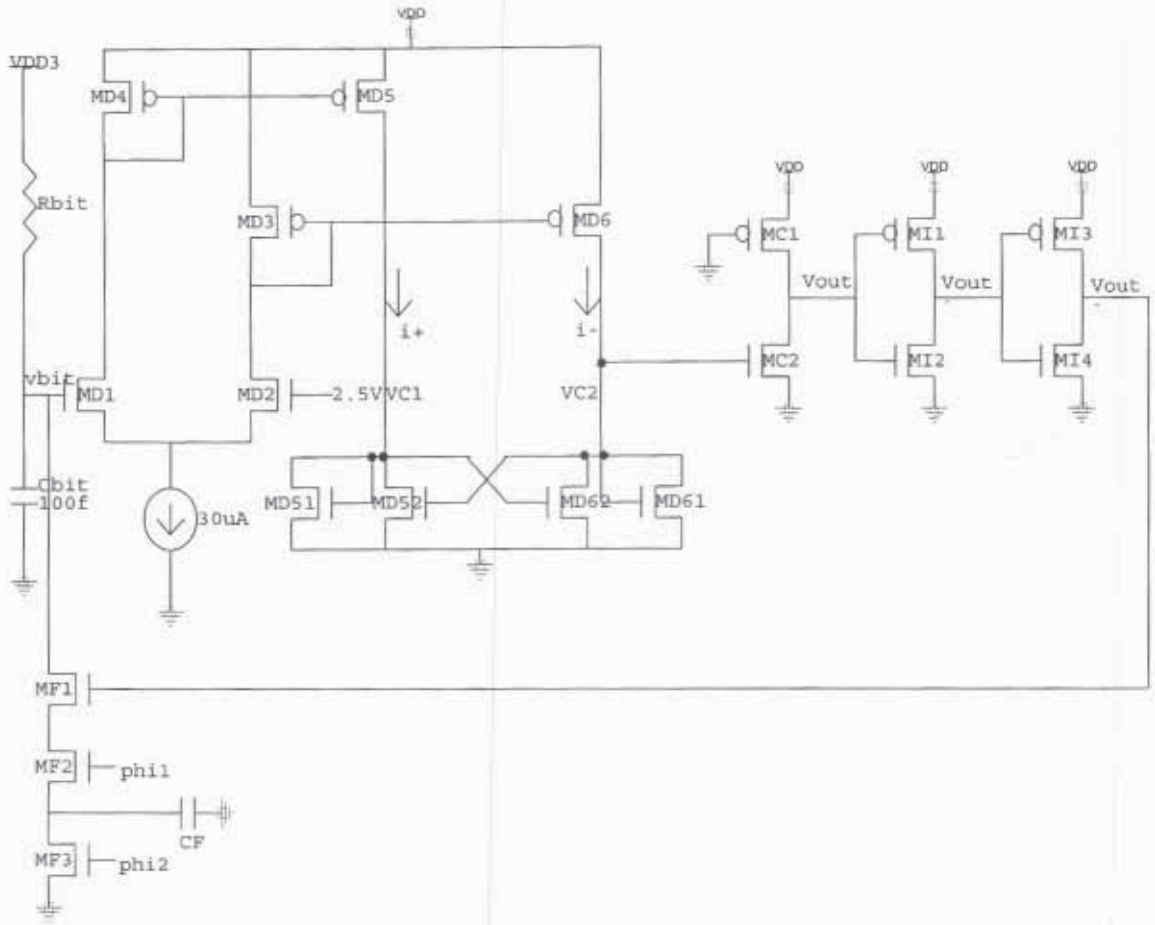


Figure 3.21. New Sense Amplifier

From figure 3.22, we see that Vbit is initially at 2.6v. Since this is higher than the reference voltage (2.5v), the buffer output Vout2 goes high. The output from the buffer is shown by Vout2 in figure 3.22. Vout2 going high causes the feedback circuit to be activated. The feedback circuit pulls out a precise charge from Vbit node. This simulation was for an Rbit value of 500k ohms. As charge is removed from Vbit node the Vout2 signal goes low because the Vbit voltage is now lower than the reference voltage (2.5v). After some time the Vbit node gets charged back to voltage VDD3, which is higher than Vref (2.5v) causing Vout2 to go high.

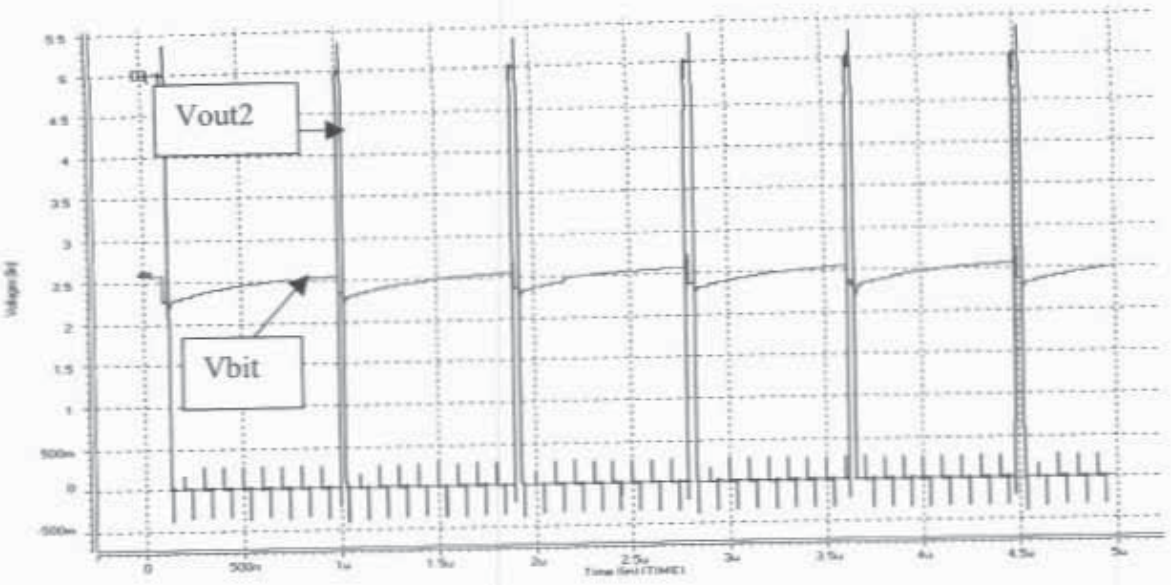


Figure 3.22. New Sense Amplifier Output

When the circuit is run for a considerable time, we can see the number of times vout2 goes high (figure 3.23) varies with the bit resistance Rbit which models the flash cell. For a larger bit resistance, it takes more time for vbit node to charge back above Vref (2.5v). The Rbit value can be estimated knowing the number of times Vout2 goes high in a given time.

Estimating the Bit Resistance:

Equating the current through the bit resistance with the current through the switched capacitance resistance can give us a relation to estimate the bit resistance.

$$\frac{V_{DD3} - V_{bit}}{R_{bit}} = \frac{V_{bit}}{R_{SC}} \qquad (3.21)$$

Solving for R_{bit},

$$R_{bit} = \frac{V_{DD3} - V_{bit}}{V_{bit}} R_{SC}$$

Where,

$$R_{sc} = \frac{1}{f_{CLK} C_F} \frac{\#ofClockPulses}{\#ofOutputHigh} \qquad (3.22)$$

Therefore

$$R_{bit} = \frac{V_{DD3} - V_{bit}}{V_{bit}} \cdot \frac{1}{f_{CLK}.C_F} \cdot \frac{50}{\#ofOutputHigh} \quad (3.23)$$

$$R_{bit} = \frac{2.6 - 2.5}{2.5} \cdot \frac{1}{10.10^6.40.10^{-15}} \cdot \frac{50}{\#ofOutputHigh}$$

$$R_{bit} = \frac{5.10^6}{\#ofOutputHigh}$$

The number of times the output is high is shown in figure 3.23 by red lines. From that information the resistance values are calculated below.

For 40 highs,

$R_{bit} = \frac{5.10^6}{40} = 125k$, The simulated resistance was 100KΩ

For 18 highs,

$R_{bit} = \frac{5.10^6}{18} = 277k$, The simulated resistance was 300KΩ

For 9 highs,

$R_{bit} = \frac{5.10^6}{9} = 555k$, The simulated resistance was 500KΩ

For 6 highs,

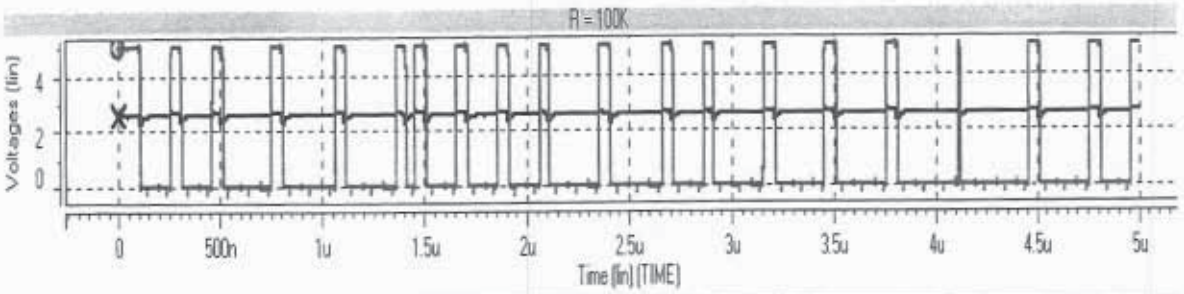
$R_{bit} = \frac{5.10^6}{6} = 833k$, The simulated resistance was 750KΩ

For 4 highs,

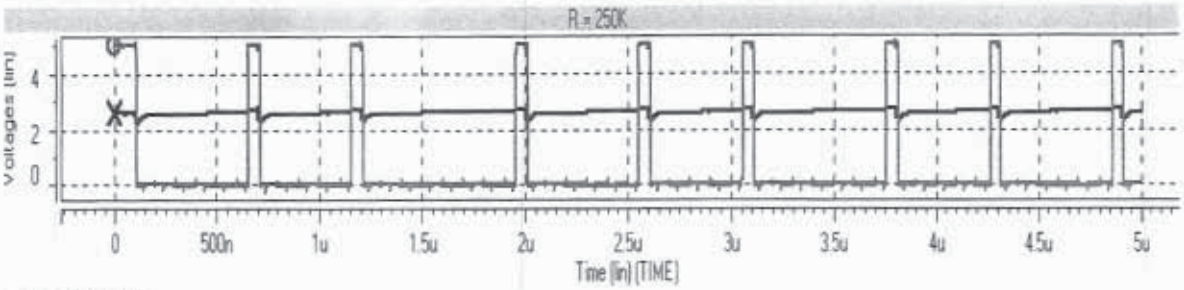
$R_{bit} = \frac{5.10^6}{4} = 1.25Meg$, The simulated resistance was 1MegΩ

We could see form figure 3.23 that the approximation is very close to the absolute value used in the simulation.

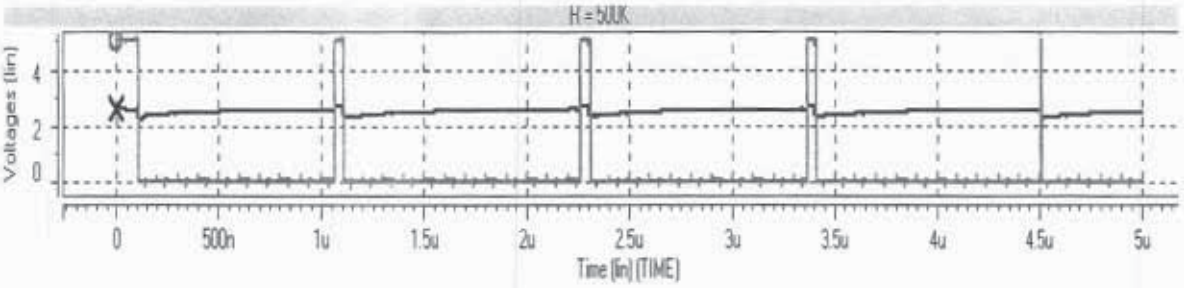
R = 100K



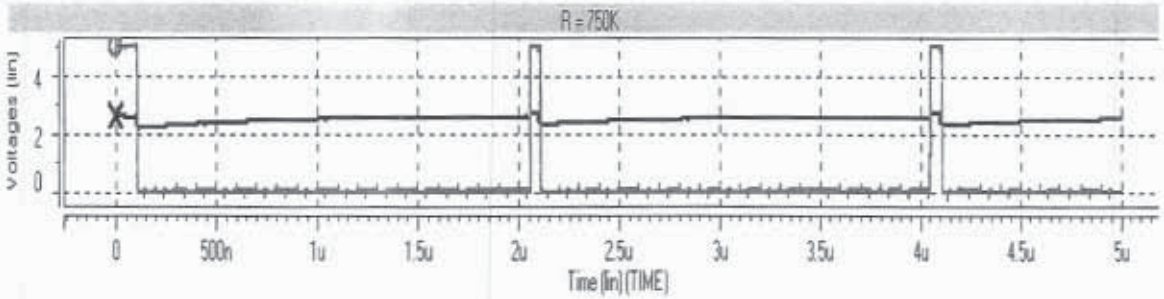
R = 250K



R = 500K



R = 750K



R = 1 MEG

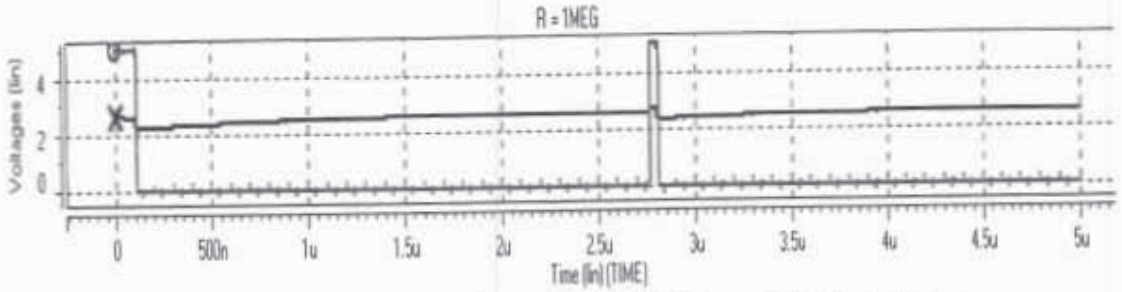


Figure 3.23. Complete Circuit with Different Bit Resistance

CHAPTER FOUR: SUMMARY

A circuit was designed, which uses averaging techniques to determine the value of the bit resistance. This is a robust design, and does not rely on precision circuitry. The output signal goes high when there is a change in the bit line voltage. From the output signal we can determine the exact value of the resistance. The output can be connected to a counter, which counts the number of times the output goes high. Resistance can be estimated using the counter value. In our results we saw that resistance's of various values can be estimated using this circuit design. This circuit can sense various current levels. With variation in process the resistance value might change. Since the flash memory is programmed in a similar way this circuit is also robust with process variation.

The decision circuit plays a major role in sensing the micro amp difference between the differential pair currents. This project can further be developed to work on low power supply and to do multi-level storage in flash memory cell.

REFERENCES

[1] Ashok K.Sharma, *Semiconductor Memories Technology, Testing and Reliability*, IEEE Press, Chapter 3. ISBN-0-7803-1000-4.

[2] R.Jacob Baker, Harry W.Li and David E.Boyce, *CMOS, Circuit Design, Layout and Simulation,* John Wiley and Sons publishers, Chapter 21. ISBN 0-7803-3416-7.

[3] Roubik Gregorian, *Introduction to CMOS OP-Amps and Comparators*, John Wiley & Sons, Chapter 3. ISBN-0-471-31778-0.

[4] R.Jacob Baker, Harry W.Li and David E.Boyce, *CMOS, Circuit Design, Layout and Simulation,* John Wiley and Sons publishers, Chapter 26. ISBN 0-7803-3416-7.

[5] Ibid, Chapter 27.

[6] Intel Technology Journal, *Intel Strata Flash Memory Technology Development and Implementation,* 2003, 23 March 2003. <http://www.intel.com/technology/itj/q41997/articles/art_2.htm>

[7] Micron Technotes, *Boot Block Flash Memory Technology,* 23 March 2003. <http://download.micron.com/pdf/technotes/FT01.pdf>

[8] Micron Technotes, *Designing for Smart Voltage Technology Flash,* 23 March 2003. <http://download.micron.com/pdf/technotes/FT04.pdf>