Analog VLSI Implementation of Synaptic Modification in Realistic Neurons

by

Jason Leonard

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degrees of Bachelor of Science in Electrical Science and Engineering and Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 22, 1998

June 1998

Author .....
Department of Electrical Engineering and Computer Science
May 22, 1998

Certified by .....
Chi-Sang Poon
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

Analog VLSI Implementation of Synaptic Modification in Realistic Neurons

by

Jason Leonard

Submitted to the Department of Electrical Engineering and Computer Science on May 22, 1998 in partial fulfillment of the requirements for the degrees of Bachelor of Science in Electrical Science and Engineering and Master of Engineering in Electrical Engineering and Computer Science

Abstract

An analog VLSI implementation of synaptic modification in realistic neurons is presented. The implementation uses CMOS integrated circuit technology to emulate the electrical behaviors of the neuron membrane, dendrite, and synapse, using principles based on the actual biology. The synapse circuitry includes a mechanism for the modification of the synaptic conductance. The circuits were simulated, layed out, and submitted for fabrication.

Thesis Supervisor: Chi-Sang Poon
Title: Principal Research Scientist, Harvard-MIT Division of Health Sciences and Technology

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Traditional electronic artificial networks implement mathematical or engineering abstractions of biological neurons. These are usually designed using digital circuits that operate up to a million times faster than actual neurons. There is an interest, however, in developing an artificial neural network that uses more life-like principles of neural computation. This type of neural network could potentially be used in electronic and electromechanical systems, such as artificial vision devices and robotic arms, to interact with real-world events in the same manner as biological nervous systems. This type of network could also be used as a research tool to better understand how biological neural networks communicate and learn. Today, the ability to simulate with software the behavior of biological networks is severely limited. Considering that a biological nervous system contains thousands or millions of interconnected neurons, such simulations could take many hours or days on the fastest of computer. An electronic circuit that emulates the analog behavior of actual biological neurons could perform the simulation in real time.

1.2 Background

Most efforts[2],[4] at creating electronic implementations of biological neurons have focused on emulating the input-output functional characteristics of the neuron, essentially treating the neuron as an abstracted black box. These implementations focus on the creating of an action potential, which is what neurons to communicate with one another, but with little regard for what actually produces the action potential in a biological neuron Mahowald and Douglas[1] have produced an analog integrated circuit with the functional characteristics of real nerve cells, but isomorphically emulates the membrane conduc-

tances within an actual neuron cell body.

1.3 Objective and Organization

This thesis discusses an analog VLSI implementation of a biological neuron. This implementation builds on the work of Mahowald and Douglas, but adds circuitry for the synapse through which neurons communicate, and the dendrite, the connection between the synapse and neuron cell body. Also included is circuitry for adaptation, or learning.

Chapter 2 discusses the biology of the neuron and presents electrical models of the neuron cell body and a Hebbian synapse. Chapter 3 is an overview of analog VLSI technology and discusses circuit blocks that are used in implementing the neuron circuits. Chapter 4 discusses in detail the implementation of the circuit for the neuron cell body, based on the work of Mahowald and Douglas. Chapter 5 discusses the circuit implementation of the Hebbian synapse, including the circuitry for learning and adaptation. Chapter 6 discusses the circuit implementation of the dendritic connection between the cell body and synapse. Chapter 8 presents HSPICE simulation results of a complete neural circuit. Chapter 9 discusses the layout in silicon of the circuits discussed in the previous chapters. Chapter 10 contains a summary of the results of this thesis, and offers suggestions for future work related to this thesis.

Chapter 2

Biological Background

2.1 Overview

The neuron is the basic anatomical unit of the nervous system[4]. It consists of a cell body equipped with a tree of filamentary structures called dendrites. The dendrites are covered with structures called synapses, where junctions are formed with other neurons. The synapses are the primary information processing elements in neural systems. The dendrites sum the synaptic inputs from other neurons, and the resulting currents are integrated on the membrane capacitance of the cell body until a threshold is reached. At that point, an output nerve pulse, called the action potential, is generated and then propagates down the neuron's axon, a long structure used to transmit data. The end of the axon consists of a tree of synaptic contacts that connect to the dendrites of other neurons.

2.2 Neuron Cell Body

The electrical activity in a neuron cell body occurs in the thin membrane that electrically

Ion	Concentration Inside (mM/l)	Concentration Outside (mM/l)	Reverse Potential (mV)
Potassium (K+)	400	10	-92
Sodium (Na+)	50	460	55
Chlorine (Cl-)	40	540	-65

Table 2.1: Typical concentrations of ions in the nerve cell membrane[4]

separates the neuron's interior from exterior fluid. An energy barrier is formed by the nerve membrane that is so high that few ions are able to surmount it[4]. Inside all nerve membranes are metabolic pumps that actively expel sodium ions from the cytoplasm while importing potassium ions from the extracellular fluid. The concentrations of these

ions inside and outside the cell are shown in Table 2.1.

The concentration gradient that exists across the membrane is used to power electrical activity. Ions diffuse in(out) of the membrane while electrically drifting out(in). When the voltage across the membrane reaches the reverse potential

$$V_r = -\frac{kT}{q}\ln\frac{N_{in}}{N_{ex}}, \tag{2.1}$$

the diffusion of ions will be exactly counterbalanced by the drift of ions. The reverse potentials for the three ions in the membrane are shown in Table 2.1. In operational terms, one can think of the sodium reverse potential as a positive power supply rail and the potassium reverse potential as the negative rail[4].

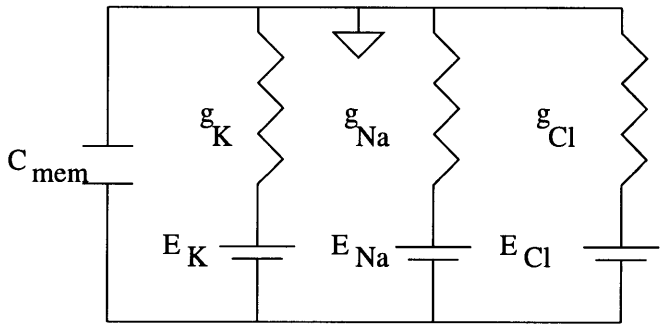

Figure 2.1: Electrical Model of Nerve Cell Membrane

Figure 2.1 summarizes the electrical characteristics of the nerve membrane. The voltage sources represent the reverse potentials of the ions, while the conductances represent the membrane permeability for that ion. The membrane capacitance is depicted as a lumped capacitor. For a given membrane voltage, the current through the membrane(i.e., the capacitor current) is

$$I_{mem} = (V_{memK} - V_K)G_K + (V_{Na} - V_{mem})G_{Na} + (V_{Cl} - V_{mem})G_{Cl} \tag{2.2}$$

Any net current will charge or discharge the membrane capacitance until the current is reduced to zero. Under normal conditions, the chlorine current is small and can be

neglected[4]. With this assumption, the voltage at which the net current is zero is:

$$V_0 = \frac{V_K G_K + V_{Na} G_{Na}}{G_K + G_{Na}} \tag{2.3}$$

V0 is called the resting potential and is typically equal to -85 millivolts, although it can vary considerably. A neuron at rest is termed "polarized" to a negative potential. The membrane becomes depolarized when its potential becomes more positive.

There are actually a number of different potassium ion currents. The delayed rectifier current, IKD, along with the sodium current, generates the nerve impulse. Two other potassium currents with slower dynamics, the so-called A-current(IKA) and the calcium dependent potassium current (IAHP or after-hyperpolarizing current), control the rate at which impulses are produced[1].

The activation and inactivation of the ion conductances in the membrane are themselves dependent on the membrane voltage and time. There is a sigmoidal relationship between the ion conductance and the membrane voltage. This creates a thresholding behavior that is responsible for the generation of the nerve pulse, or action potential. As the membrane becomes depolarized(in response to, for instance, an influx of synaptic current), there is a transient the sodium conductance, followed by a delayed but prolonged increase in the potassium conductance. The currents through these conductances create the action potential by acting on the capacitance of the membrane.

2.3 Hebbian Synapse

The model presented here is that proposed by Zador, Koch, and Brown[6]. Stimulation of the presynaptic neuron causes the release of neurotransmitters from its axon terminal. These transmitters include the amino acid glutamate.The transmitters bind to corresponding receptors on the postsynaptic membrane, causing ion channels to open up through these receptors. Glutamate binds to three types of receptors: n-methyl-D-aspartate

(NMDA), quisqualate, and kainate. LTP and LTD[3] are both mediated by the NMDA receptors, which carry primarily Ca^{2+} currents. The other receptors, termed the non-NMDA receptors, carry the rest of the synaptic current, consisting mainly of Na^{+}, with negligible Ca^{2+} content. These receptors are located on a spine head connected to the dendritic shaft.

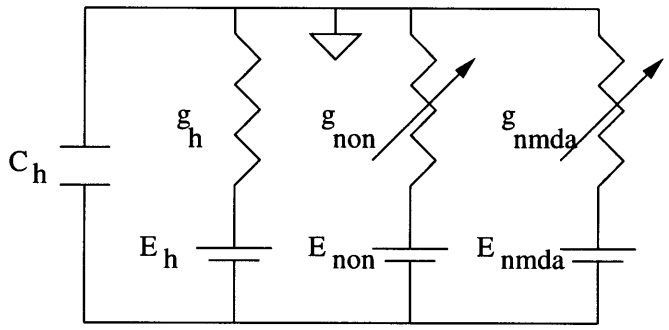

Figure 2.2: Electrical Model of the Spine Head

Figure 2.2 depicts the electrical model of the spine head. The total synaptic current consists of the sum of the NMDA and non-NMDA currents. There also exists a small leakage conductance and the capacitance that represents the membrane capacitance of the spine head. The current through the non-NMDA channels in response to a presynaptic stimulus is given by an alpha function:

$$I_{non} = (E_{non} - V_{head})\left(\kappa g_p t e^{-\frac{t}{P_t}}\right) \tag{2.4}$$

where $\kappa = e/t_p$, e is the base of the natural logarithm, $t_p = 1.5ms$, $g_p = 0.5nS$, and $E_{non} = 0$ .Note that the non-NMDA receptor conductance is purely ligand(neurotransmitter) dependent. The NMDA conductance, on the other hand, is both ligand dependent, due to the binding of neurotransmitters released from the presynaptic neuron, and dependent on the spine head membrane voltage. The current through the NMDA receptors is given by:

$$I_{NMDA}(t) = (E_{NMDA} - V_{head})g_n \frac{\left(e^{-\frac{t}{\tau_1}} - e^{-\frac{t}{\tau_2}}\right)}{1 + \eta[Mg^{2+}]e^{-\gamma V_{head}}} \tag{2.5}$$

where $\tau_1 = 80ms$, $\Upsilon_2 = 0.67ms$, $\eta = 0.33nM^{-1}$, $\gamma = 0.06mV^{-1}$, $E_{NMDA} = 0$, and $g_n = 0.2nS$. The voltage dependence of the NMDA receptor arises from the fact that the receptors inhibited by Mg^{2+} ions whose binding rate constant is dependent on the spine head membrane voltage. Near the resting membrane potential, the NMDA receptor channels are almost completely blocked by the Mg^{2+} ions, and thus little current flows. As the spine head membrane becomes partially depolarized, the Mg^{2+} ions become dislodged and more NMDA current flows.

The postsynaptic flow of Ca^{2+} ions through the NMDA receptor channels is crucial for the induction of LTP and LTD[3]. Upon entering the dendritic spine, the Ca^{2+} ions trigger a series of events that lead to the induction and maintenance of LTP or LTD. The precise mechanisms, however, are not well understood. One theory is that a second messenger, such as nitric oxide, is activated by the Ca^{2+} ions and certain calcium dependent proteins and then diffuses back to the presynaptic terminal, stimulating the release of more glutamate. Thus, this retrograde messenger operates as a positive feedback mechanism. This model is limited, though, in that it would only explain LTP. Another model suggests that rather than affecting presynaptic neurotransmitter release, the induction of LTP or LTD modulates the postsynaptic conductance of the non-NMDA receptor channels, which carry the bulk of synaptic current[3]. This model also relies on the influx of Ca^{2+} ions into the dendritic spines through the NMDA channel. During high frequency stimulation, the Ca^{2+} ions reach high concentrations in the compartmentalized spine head and preferentially activates a protein kinase. During low frequency stimulation, lower concentrations are reached and a protein phosphatase is released. Both proteins act on a common phos-

phoprotein, which triggers LTP or LTD by modulating the non-NMDA receptor channel conductance.

Chapter 3

Analog VLSI

This chapter discusses the analog circuits that are used as building blocks in the implementation of the neuron circuits. First, the principles of subthreshold behavior of CMOS transistors are discussed. Then, the circuit building blocks are discussed in detail.

3.1 Subthreshold MOSFET Operation

In traditional analog CMOS circuits, the transistors are biased in the saturation, or strong inversion, region of operation, where the drain current is given by

$$I_{DS} = \frac{k'W}{2L}(V_{GS} - V_t)^2(1 + \lambda V_{DS}) \tag{3.1}$$

where k' is a physical parameter, W/L is the ratio of the transistor channel width to channel length, V_{GS} is the gate-to-source voltage, V_t is the threshold voltage of the transistor, λ is the inverse of the Early voltage of the transistor, a parameter related to the output resistance of the device, and V_{DS} is the drain-to-source voltage.For $\lambda V_{D(S)} \gg 1$, the current is roughly independent of the drain-to-source voltage.

Generally, the transistor is not used in the subthreshold, or weak inversion, region of operation. In this region, the current levels are very small and the drain current varies exponentially with the gate voltage:

$$I_{DS} = k_x \frac{W}{L} e^{\frac{V_{GS}}{nV_T}} \left(1 - e^{\frac{-V_{DS}}{V_T}}\right) \tag{3.2}$$

where k_x and n are fabrication dependent parameters and V_T is the thermal voltage, which is approximately 26 millivolts at room temperature. For drain-to-source voltages greater than a few thermal voltages, the drain current is essentially independent of the drain-to-source voltage.

Most of the circuits discussed in this thesis have transistors biased in the subthreshold region, to take advantage of the exponential relationship between current and voltage, which is prevalent in actual neurons. CMOS technology is used instead of bipolar technol-

ogy because of the greater availability of CMOS processes. Also, the current levels, and hence time scales, of subthreshold CMOS transistors match up fairly well with the actual biological levels. Additionally, the nearly infinite resistance of the MOS gate is useful in many circuits.

3.2 Basic Circuit Building Blocks

3.2.1 Differential Pair

Perhaps the most important circuit is the differential pair shown in Figure 3.1. Assuming large enough drain-to-source voltages, the currents I_1 and I_2 can be expressed as

$$I_1 = k_x \frac{W_1}{L_1} e^{\frac{(V_1 - V)}{nV_T}} \tag{3.3}$$

$$I_2 = k_x \frac{W_2}{L_2} e^{\frac{(V_2 - V)}{nV_T}} \tag{3.4}$$

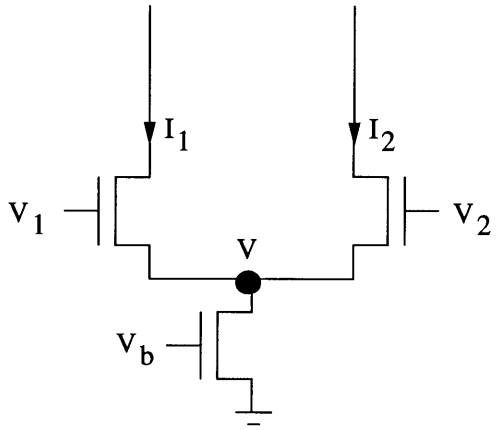

Figure 3.1: Differential Pair

These sum of these two drain currents must be equal to the current I_b through the bias transistor:

$$I_b = I_1 + I_2 = k_x e^{\frac{-V}{nV_T}} \left(\frac{W_1}{L_1} e^{\frac{V_1}{nV_T}} + \frac{W_2}{L_2} e^{\frac{V_2}{nV_T}} \right) \tag{3.5}$$

Solving this equation for the node voltage V yields:

$$e^{\frac{-V}{nV_T}} = \frac{I_b}{k_x} \frac{1}{\left(\frac{W_1}{L_1} e^{\frac{V_1}{nV_T}} + \frac{W_2}{L_2} e^{\frac{V_2}{nV_T}}\right)} \tag{3.6}$$

Substituting Equation 3.6 into Equations 3.3 and 3.4 yields expressions for the drain currents:

$$I_1 = I_b \frac{\frac{W_1}{L_1} e^{\frac{V_1}{nV_T}}}{\frac{W_1}{L_1} e^{\frac{V_1}{nV_T}} + \frac{W_2}{L_2} e^{\frac{V_2}{nV_T}}} \tag{3.7}$$

$$I_2 = I_b \frac{\frac{W_2}{L_2} e^{\frac{V_1}{nV_T}}}{\frac{W_1}{L_1} e^{\frac{V_1}{nV_T}} + \frac{W_2}{L_2} e^{\frac{V_2}{nV_T}}} \tag{3.8}$$

Essentially, if V_1 is more positive than V_2 by many nV_T, the voltage V rises to turn transistor M2 off, so that all of the current goes through M1 and $I_1 \approx I_b$ .The analogous situation occurs when V_2 is more positive than V_1.

3.2.2 Current Mirror

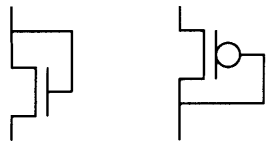

Figure 3.2: Diode-connected configuration for NMOS and PMOS devices

Figure 3.2 depicts the diode-connected circuit configuration, where the transistor's gate is connected to its drain. Thus, $V_{GS} = V_{DS}$, which for all useful gate-to-sources voltages means that the V_{DS} term in Equation 3.2 is negligible, guaranteeing that the transistor is saturated.

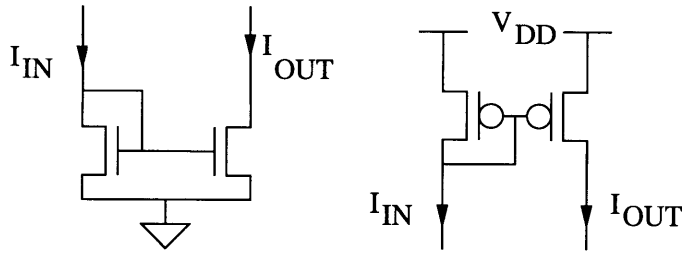

Figure 3.3: Current Mirrors

The circuits shown in Figure 3.3 are NMOS and PMOS current mirrors. In the NMOS current mirror, transistor M1 is diode-connected, and its gate is also connected to the gate of M2. Since the sources of M1 and M2 are common, they both have the same V_{GS}. Hence, assuming that the V_{DS} of M2 is large enough for M2 to be saturated also, the current I_{OUT} is:

$$I_{OUT} = \frac{W_2/L_2}{W_1/L_1} I_{IN} \tag{3.9}$$

Thus, the current mirror not only "mirrors" the input current, but it can scale it also.

3.2.3 Transconductance Amplifier

The simple transconductance amplifier[4] is shown in Figure 3.4. It consists of a differential pair "loaded" with a PMOS current mirror. If the transistors in the PMOS current mirror have the same dimensions, then the current I_1 is mirrored with unity gain to the other leg of the current mirror. Assuming that the transistors M1 and M2 of the differential pair are matched; i.e., the have the same W/L ratio, then the output current can be expressed as:

$$I_{out} = I_1 - I_2 = I_b \frac{e^{\frac{V_1}{nV_T}} - e^{\frac{V_2}{nV_T}}}{e^{\frac{V_1}{nV_T}} + e^{\frac{V_2}{nV_T}}}$$

$$I_{out} = I_b \tanh\left(\frac{1}{2}\frac{V_1 - V_2}{nV_T}\right) \tag{3.10}$$

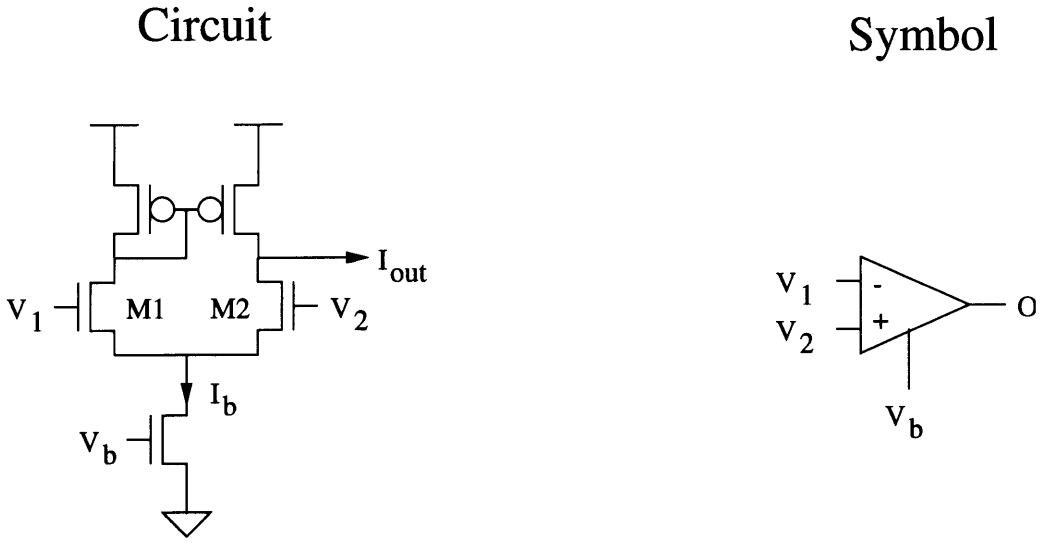

Figure 3.4: Simple transconductance amplifier: circuit and symbol

The output current is plotted as a function of V_1-V_2 in Figure 3.5 for n=?. Note that the output current is approximately linearly dependent on the input voltage difference for differences up to 100 millivolts, after which the output current begins to saturate to positive or negative I_b.

Figure 3.5: Output current of a simple transconductance amplifier as a function of the difference in input voltages. Vb=? n=?.

The transconductance of the simple transconductance amplifier in this linear region, with V_{in} defined to be V_1 - V_2, is

$$g_m = \frac{\partial I_{out}}{\partial V_{in}} = \frac{1}{2}\frac{I_b}{nV_T} \tag{3.11}$$

and is proportional to the bias current I_b.

The open-circuit voltage gain of the amplifier is

$$A = \frac{\partial V_{out}}{\partial V_{in}} = g_m r_{out} \tag{3.12}$$

where r_{out} is the finite output resistance of the amplifier. This is an effect of the deviation from ideal current source behavior of the amplifier. r_{out} is defined as:

$$r_{out} = \frac{\partial V_{out}}{\partial I_{out}} \tag{3.13}$$

and can be expressed as:

$$r_{out} = r_{oN} \| r_{oP} = \left(\frac{V_N}{I_b/2}\right) \| \left(\frac{V_P}{I_b/2}\right) \tag{3.14}$$

where r_{oN} and r_{oP} are the incremental output resistances of M2 and M4, respectively, and V_N and V_P are the Early voltages of NMOS and PMOS transistors, respectively. The Early voltages are proportional to channel length. Thus, the incremental voltage gain A of the amplifier is[4]

$$A = \frac{1}{nV_T}\left(\frac{1}{V_N} + \frac{1}{V_P}\right)^{-1} \tag{3.15}$$

One of the limitations of this simple transconductance amplifier is that its output voltage can has a significant lower limit. In other words, the output voltage cannot decrease below a certain value and the amplifier still work as desired. That critical value is:

$$V_{out-min} = \frac{1}{n}(min(V_1, V_2) - V_b) \tag{3.16}$$

Essentially, the amplifier will work as desired with its output voltage almost up to the power supply V_{DD} and down to V_b below the lowest input voltage. The fact that this lower bound depends on the input voltage limits the useful applications of the circuit.

3.2.4 Wide Output Range Amplifier

The circuit in Figure 3.6 avoids the output voltage limitation of the simple transconductance amplifier discussed in Section 3.2.3. This output of this circuit is capable of swinging almost down to ground and almost up to V_{DD}, independent of the input voltages.

Figure 3.6: Wide output range transconductance amplifier schematic

The current through M1 is mirrored through M3 and M4, while the current through M2 is mirrored through M5 and M6, and then again through M7 and M8. The currents flowing through M4 and M8 are just the two currents in the differential pair. the output current, then, is just the difference between I_1 and I_2.

The diode-connect transistors M3, M5, and M7 hold the drains of M1, M2, and M6, respectively, relatively constant.Hence, the output resistances of these transistors are not crucial to the operation of this circuit. Only the drain to source voltages of transistors Mb, M4 and M8 operate over a large range in this circuit. The channel lengths of these transistors can be made long to increase their output resistances, and therefore make the output currents of this circuit largely independent of the output voltage. The high output resistances of M4 and M8 also give the circuit a large voltage gain.

3.2.5 Wide Input Range Amplifier

In some applications, it is necessary for the transconductance amplifier to approximate a linear conductance over a range larger than the approximately 100 millivolts of the simple amplifier discussed in section 3.2.3. The amplifier shown in Figure 3.7 has a linear range twice that of the simple transconductance amplifier. This amplifier has a diode-connected transistor in each of the legs of the differential pair. Consider the currents in each leg of the differential pair. I_1 and I_2 can be expressed as

$$I_1 = I_0 e^{(V_1 - V_A)/(nV_T)} = I_0 e^{(V_A - V_N)/(nV_T)} \tag{3.17}$$

$$I_2 = I_0 e^{(V_2 - V_B)/(nV_T)} = I_0 e^{(V_B - V_N)/(nV_T)} \tag{3.18}$$

Solving for V_A and V_B,

$$e^{(-2V_A)/(nV_T)} = e^{(-V_N)/(nV_T)} e^{(-V_1)/(nV_T)} \tag{3.19}$$

$$e^{(-2V_B)/(nV_T)} = e^{(-V_N)/(nV_T)} e^{-(V_2)/(nV_T)} \tag{3.20}$$

$$e^{(-V_A)/(nV_T)} = e^{(-V_N)/(2nV_T)} e^{(-V_1)/(2nV_T)} \tag{3.21}$$

$$e^{(-V_B)/(nV_T)} = e^{(-V_N)/(2nV_T)} e^{(-V_2)/(2nV_T)} \tag{3.22}$$

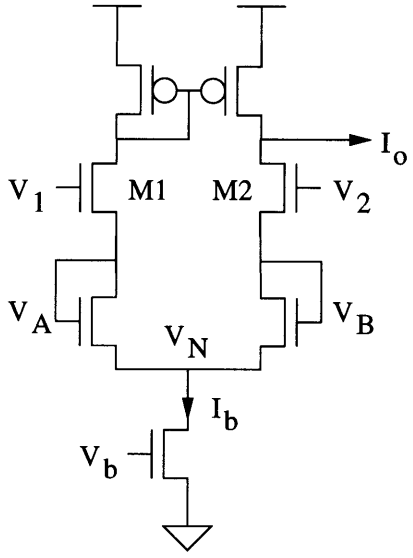

Figure 3.7: Wide input range transconductance amplifier

Noting that the currents I_1 and I_2 must sum to the bias current I_b,

$$I_b = I_1 + I_2 = I_0 e^{(-V_N)/(nV_T)} (e^{V_A/(nV_T)} + e^{V_B/(nV_T)}) \tag{3.23}$$

$$I_b = I_0 e^{(-V_N)/(nV_T)} (e^{V_N/(nV_T)} e^{V_1/(2nV_T)} + e^{V_N/(nV_T)} e^{V_2/(2nV_T)}) \tag{3.24}$$

$$I_b = I_0 e^{(-V_N)/(2nV_T)} (e^{V_1/(2nV_T)} + e^{V_2/(2nV_T)}) \tag{3.25}$$

$$e^{(-V_N)/(2nV_T)} = \frac{I_b}{I_0} \frac{1}{(e^{V_1/(2nV_T)} + e^{V_2/(2nV_T)})} \tag{3.26}$$

Plugging Equations 3.20-21 and 3.26 into Equations 3.17-18, we get expressions for I1 and I2 in terms of V1 and V2:

$$I_1 = I_b \frac{e^{V_1/(2nV_T)}}{(e^{V_1/(2nV_T)} + e^{V_2/(2nV_T)})} \tag{3.27}$$

$$I_2 = I_b \frac{e^{V_2/(2nV_T)}}{(e^{V_1/(2nV_T)} + e^{V_2/(2nV_T)})} \tag{3.28}$$

From these expressions, we can write the output current as

$$I_{out} = I_1 - I_2 = I_b \tanh\left(\frac{1}{4}\frac{V_1 - V_2}{nV_T}\right) \tag{3.29}$$

This is of the same for as the output current for the simple transconductance amplifier. However, the argument of the hyperbolic tangent function contains an additional factor of one-half. Hence, the input voltage difference can be twice as large for this amplifier as the simple transconductance amplifier and still maintain an approximately linear conductance.

3.3 Follower Configurations

3.3.1 Voltage Follower

The unity-gain voltage follower, or buffer, configuration is shown in Figure 3.8. The output voltage follows changes in the input voltage, without affecting the circuit producing the input. The output voltage of the amplifier is the voltage gain of the amplifier multiplied by the difference in the voltages at its input terminals:

$$V_{out} = A(V_{in} - V_{out}) \tag{3.30}$$

Hence the transfer function of the amplifier is:

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 + \frac{1}{A}} \approx 1 - \frac{1}{A} \tag{3.31}$$

which for large values of A is close to 1. The simple transconductance amplifier can achieve gains on the order of 100-200, while the wide output range amplifier can achieve a gain of over 1000. Determining which amplifier to use, including whether to use a traditional operational amplifier, depends on the application and the desired accuracy

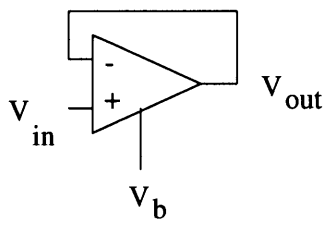

Figure 3.8: Follower Configuration

3.3.2 Follower-Integrator

The follower-integrator connection is shown in Figure 3.9. It contains an amplifier of the types discussed in Sections 3.2.3-5 connected as a follower, with a capacitor connected between the amplifier's output and ground. The output voltage can be described by the differential equation

$$C\frac{dV_{out}}{dt} = I_b \tanh\left(\frac{1}{2}\frac{V_{in} - V_{out}}{nV_T}\right) \tag{3.32}$$

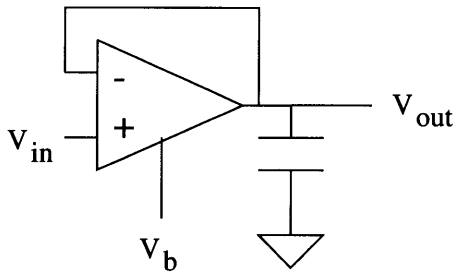

Figure 3.9: Follower-integrator connection

When the argument of the tanh function is small enough so that it is approximately linearly dependent on its input, the differential equation becomes

$$C\frac{dV_{out}}{dt} = g_m(V_{in} - V_{out}) \tag{3.33}$$

where the conductance g_m is defined as in Equation 3.11. This is a first-order differential equation analogous to that of an RC circuit in which a voltage source drives a series combination of a resistor and a capacitor.In the circuit shown here, the conductance of the transconductance amplifier corresponds to the resistor in the RC circuit. The transfer function of this circuit, using frequency domain notation(i.e., Laplace transform), can be expressed as

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 + s\tau} \tag{3.34}$$

where $\tau = C/g_m$. This circuit acts a low-pass filter with breakpoint $1/\tau$. The response of

this circuit to a small (small enough so that we may assume the tanh function is approximately linear) step change can be expressed as

$$V_{out}(t) = V_{out0}(0) + \Delta V_{in}(1 - e^{-t/\tau}) \tag{3.35}$$

Since the follower-integrator connection has an infinite input resistance, multiple follower-integrators can be cascaded without loading one another, which is not true of a passive RC network. Thus, the circuit can be cascaded to form a delay line whose transfer function and delay is well prescribed. Furthermore, the time constant of the follower-integrator connection can be varied over orders of magnitude by varying the bias voltage of the transconductance amplifier.

Chapter 4

Circuit Implementation of Nerve Cell Membrane

This chapter discusses the implementation of the electrical model of the nerve cell membrane shown in Figure 2.1. All circuits in this section and throughout the thesis were implemented using the parameters of the Orbit 2.0 micron fabrication process. Device sizes and bias voltages were chosen based on this process.

4.1 Cell Membrane and Leakage Conductance

The cell membrane capacitance shown in Figure 2.1 is implemented with a real capacitor, connected between the 'membrane' node and ground, whose structure is discussed in Chapter 8:Layout and Fabrication. The value of this capacitance is approximately 1 pico-farad. There is also a leakage conductance (not shown in Figure 2.1) in the nerve cell membrane. The value of this conductance can be treated as constant.

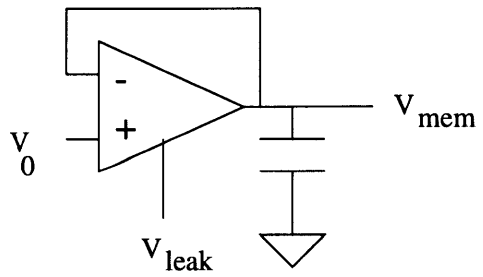

Figure 4.1: Membrane capacitance and leakage conductance(implemented using a transconductance amplifier).

This leakage conductance can be implemented using a transconductance amplifier connected in a follower-integrator configuration, with the output voltage being the membrane voltage and the capacitor being the membrane capacitance[5], as shown in Figure 4.1. The input voltage of the amplifier should be the resting potential voltage(Equation.

2.3), since there is no leakage current when the membrane voltage is at the rest resting potential. The value of the leakage conductance is determined by the bias voltage of the transconductance amplifier, which determines its g_m. The use of a transconductance amplifier yields a good approximation of the linear conductance. To improve the range over which there is an approximately constant conductance, the wide input range transconductance amplifier of Section 3.2.5 can be used instead of the simple transconductance amplifier.

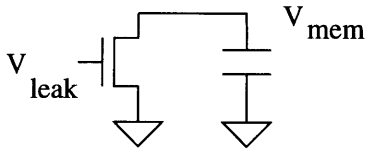

Figure 4.2: Membrane capacitance and leakage conductance(implemented using a single transistor)

Alternately, the leakage conductance can be implemented using a single transistor, whose gate voltage is set to control the leakage current[2], as shown(with the membrane capacitor) in Figure 4.2. This implementation does not emulate a true conductance however, as the current through the transistor does not vary, to first order, with its drain to source voltage. Nonetheless, since the leakage current is relatively small, the loss of accuracy comes with the benefit of a very simple and small implementation with a single transistor.

4.2 Potassium Ion Channel

Figure 4.3 depicts the circuits used to generate the delayed rectifier potassium current (IKD). This is the topology developed by Mahowald and Douglas[1]. This topology was implemented using the parameters of the Orbit 2.0 micron fabrication process. Device sizes and bias voltages were chosen based on this process.

The membrane voltage is delayed(low-pass filtered) by a follower-integrator, whose output voltage is compared to a reference voltage KDKNEE of a differential pair. The current flowing out of the differential pair is mirrored by a PMOS current mirror into a NMOS 'conductance' transistor, whose source is connected to the power supply voltage EK(corresponding to the potassium reverse potential in Figure 2.1) and whose drain is connected to the membrane node. The conductance transistor is designed with dimensions that make its behavior more ohmic than typical CMOS transistor, although the ohmic behavior is far from ideal. This transistor sinks current from the membrane capacitance.

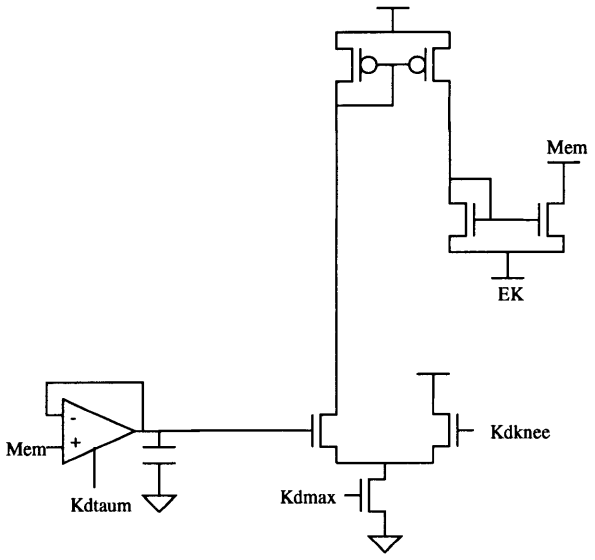

Figure 4.3: Circuit implementation of the potassium(IKD) ion current in the cell membrane.

The voltage dependence of the potassium conductance is provided by the differential pair, which Mahowald and Douglas refer to as the activation transistors. The potassium current IKD does not become inactivated, unlike the sodium current, and thus does not require an inactivation signal, as will be seen in the following section for the sodium conductance.The time dependence of the conductance is provided by the follower-integrator, whose time constant is determined by a fixed capacitor(of value 1 picofarad) and the vari-

able bias voltage KDTAUM of the transconductance amplifier. The voltage KDMAX controls the maximum amount of current that flows in the differential pair.

4.3 Sodium Ion Channel

The circuits used to generate the sodium ion current in the cell membrane are shown in Figure 4.4. Unlike the potassium current IKD, the sodium current depends on the interaction of activation and inactivation circuits.

The sodium current activates very quickly, so there is no need for a follower-integrator to delay the membrane voltage signal to the activation differential pair. The output current from the activation subcircuit represents sodium activation. The inactivation of sodium is delayed, so the low-pass filtered version of the membrane voltage is applied to the input of the inactivation differential pair. The output current of this differential pair, representing the sodium inactivation current is then mirrored by the PMOS current mirror.

A straightforward application of Kirchoff's Current Law at the node NAVG yields

$$I_{NAVG} = I_{activation} - I_{inactivation} \tag{4.1}$$

This current is then mirrored to a conductance transistor, connected between ENA and the membrane node. ENA represents the sodium reverse potential shown in Figure 2.1.

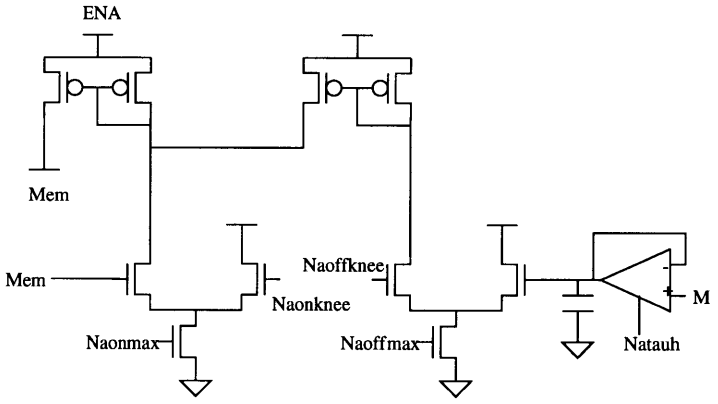

Figure 4.4: Circuit implementation of the sodium ion conductance in the nerve cell membrane.

As in the potassium circuit, the voltages NAONMAX and NAOFFMAX control the maxi-

mum amount of current that flows in the activation and inactivation subcircuits, respectively.

4.4 Voltage Scale

The nerve cell membrane voltage can vary between the limits set by the reverse potentials of potassium(EK) and sodium(ENA). In real neurons, these limits are approximately -100 millivolts and +50 millivolts, respectively. Biological conductances are roughly ten times as sensitive to voltage as CMOS transistors[1]. Furthermore, the circuits discussed in the previous sections are driven by a 0-5 Volt power supply for the Orbit 2.0 micron process transistors. For convenience, the circuit or electronic reverse ion potentials are set to EK = 1.5 V and ENA = 3.0 V. In other words,

$$V_{electronic} = 10V_{biological} + 2.5V \tag{4.2}$$

This choice of voltage provides 10 times as large a range of membrane voltage and is halfway between the power supply and ground.

4.5 HSPICE Simulation of Nerve Membrane Circuits

Figures 4.5 and 4.6 show the results of an HSPICE simulation of the nerve cell membrane circuits using the Orbit 2.0 micron fabrication process parameters. In this simulation, the only ion conductances that are included are the sodium and potassium(IKD) currents.

A step of excitory (synaptic) current is injected into the nerve cell membrane, as shown in the schematic of Figure 4.5. This stimulus and the response of the neuron cell membrane are shown in Figure 4.6. The voltage on the membrane capacitor builds up as the input current dumps charge onto it. Once the membrane voltage reaches a certain value, the sodium activation circuit is activated. This circuit rapidly injects current into the membrane capacitor, causing its voltage to rise rapidly. Soon afterward the potassium current and the sodium inactivation current start conducting. The sodium inactivation cancels

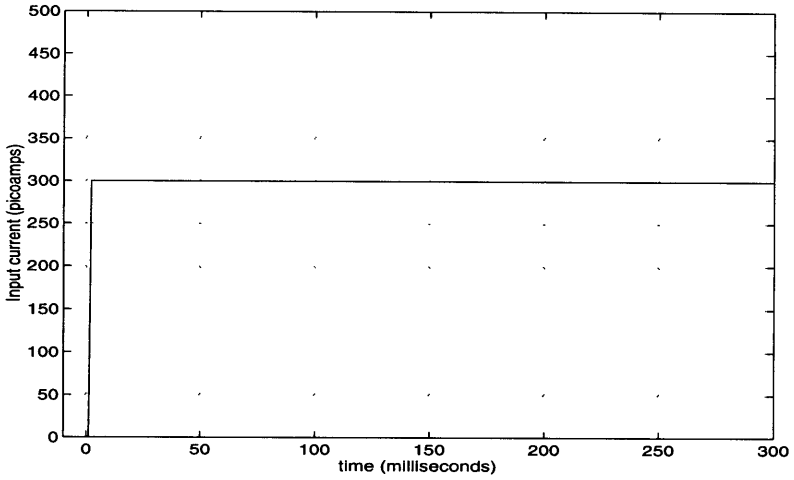

Figure 4.5: Step of input current into nerve membrane

out the sodium activation, while the potassium current drains current from the membrane capacitor, rapidly decreasing its voltage, and effectively deactivating the sodium and potassium circuits. Thus, the nerve impulse is created. Since in this simulation the excitory

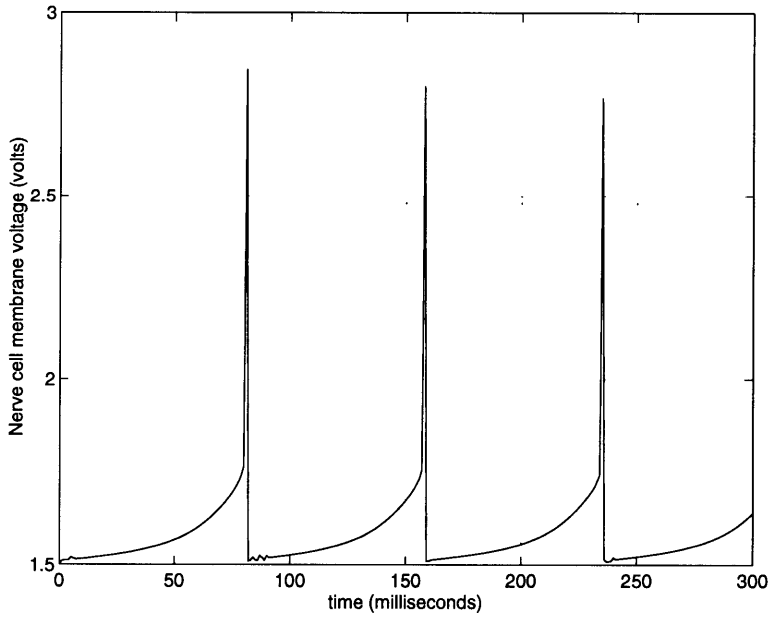

Figure 4.6: Nerve cell membrane voltage in response to step in input current

stimulus is a constant step in current, the membrane voltage again begins to build up and eventually another nerve impulse is created. In this simulation, the frequency with which

the nerve impulses are created is dependent upon the value of the step in input current, the size of the membrane capacitance, and the conductances of the sodium and potassium ion channels.

Chapter 5

Circuit Implementation of a Hebbian Synapse

5.1 NMDA Channel

The circuit implementation of the NMDA receptor channel is shown in Figure 5.1. Recall from Equation 2.5 that the conductance of the NMDA channel not only varies with time, but is also dependent on the spine head membrane voltage. This functional dependence on voltage is implemented using the differential pair in Figure 5.1. Applying and rearranging the Equation 3.7to this differential pair yields:

$$I_{diffpair} = I(t)\frac{1}{1 + Ze^{\frac{V_{Ref} - V_{Head}}{nV_T}}} \tag{5.1}$$

where Z is the ratio of the sizes of the transistors in the two legs of the differential pair.

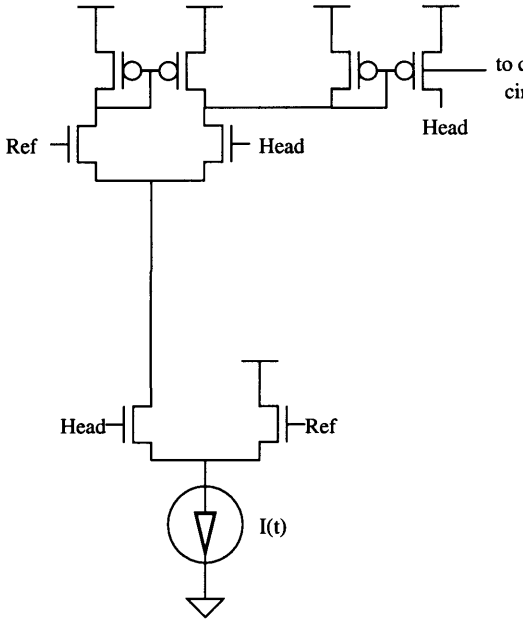

Figure 5.1: Circuit implementation of NMDA receptor channel

$I_{diffpair}$ then serves as the bias current for the transconductance amplifier. Assuming the transconductance amplifier is operating in its linear range, its output current is

$$I_{out-amp} = (V_{Head} - V_{Ref})g_m = (V_{Head} - V_{Ref})\frac{1}{2}\frac{I^{diffpair}}{nV_T} \tag{5.2}$$

This current is then mirrored(with appropriate scaling) through the PMOS current mirror back to the spine head membrane (and also to the calcium concentration circuit described in Section 5.2). If I(t) is chosen appropriately, then Equation 5.2 will have the same form as described in Equation 2.2. The actual implementation of I(t) will be discussed in Section 5.5.

In order to maintain the proper exponential dependence on voltage in implementing Equation 2.2 with the circuit in Figure 5.1, whose voltage dependence is described by Equation 5.1, It is necessary to shift and scale the voltage ranges over which the circuits operate, just as was described in Chapter 4 when implementing the neuron cell body. However, in this case the voltage scaling will be different.

Typically, the spine head voltage varies between -80 mV and 0 mV[6]. For convenience, the biological 0 V is again chosen to be 2.5 V. Thus, there is a 2.5 V shift between an actual biological channel and the electronic implementation. The voltage range must also be scaled to maintain the same behavior in the exponential. From Equations 2.2 and 5.1, it is apparent that

$$\gamma = \frac{1}{nV_T}x \tag{5.3}$$

where x is the ratio of the electronic implementation voltage to the actual biological voltage. For typical values of n, x should be approximately 2. In other words, while the biological voltage changes over a range of approximately 80 mV, the actual implementation will vary range over 160 mV. Formally, the voltages are related according to:

$$V_{electronic} = 2V_{biological} + 2.5V \tag{5.4}$$

Thus, the spine head voltage, in the electronic implementation, will vary from approximately 2.34 V to 2.5 V. This implies that V_{Ref} in Figure 5.1 should equal 2.5 V. Also, since the differential input of the transconductance amplifier is potentially up to 160 mV, greater than the linear range of the simple transconductance amplifier, it is best to replace the simple transconductance amplifier with one with a wide input range, as discussed in Section 3.2.5, which has a linear range of nearly 200 mV.

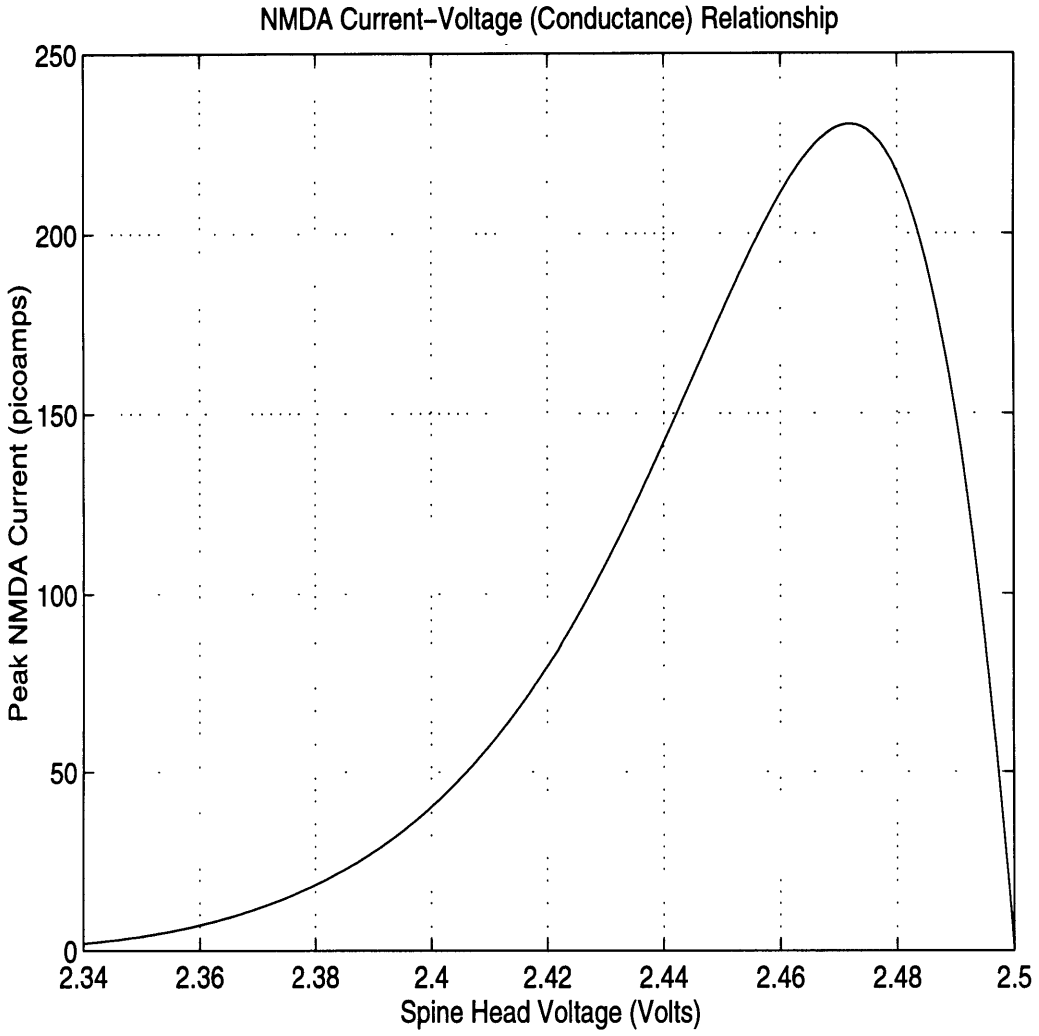

Figure 5.2: Peak NMDA current as a function of spine head voltage

The voltage dependence of NMDA current produced by the circuits in Figure 5.1 is shown in Figure 5.2. Here, the peak current is plotted as a function of voltage. For lower vales of the spine head voltage, corresponding to negative highly polarized values of biological potential, the magnesium ions block the NMDA channel and the current is small. As the spine head becomes depolarized (to a more positive potential), the NMDA channel conducts more. When the spine head voltage gets very close to its resting potential(biological = 0 mV; electronic = 2.5 V) very little current flows because the voltage gradient is small.

5.2 Calcium Measurement

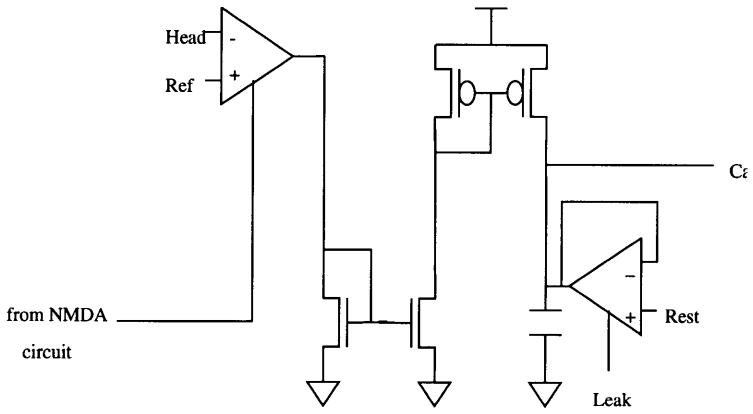

Figure 5.3: Calcium concentration measurement circuit

The circuit in Figure 5.2 measures the concentration of Ca^{2+} ions in the compartmentalized spine head. Recall from Section 2.3 that these ions are the impetus for the induction of LTP or LTD. The Ca^{2+} ions enter the spine head primarily from the NMDA channel current, which mostly consists of Ca^{2+} ions. Thus, the flow of Ca^{2+} ions into the spine head is roughly proportional to the NMDA current. Hence, a fairly accurate measure of the Ca^{2+} concentration in the spine head is the measure of net charge over time that flows in the NMDA current. For an arbitrary current I(t), the total charge is

$$Q(t_2) - Q(t_1) = \int_{t_1}^{t_2} I(t)dt. \tag{5.5}$$

This charge can be accumulated on a capacitor, whose voltage is thus proportional to the charge. In other words, the concentration of Ca^{2+} ions can be measured by accumulating a scaled copy of the NMDA current onto a capacitor. The voltage on the capacitor will then represent the concentration of Ca^{2+} ions. This is implemented in the circuit in Figure 5.2. The NMDA current is mirrored and scaled by the PMOS current mirror. This current then is accumulated onto a capacitor connected to ground. The voltage on the capacitor can be expressed as

$$V_{Ca^{2+}}(t) = \frac{1}{C}\int_0^t I_{NMDA}(t)dt + V_{Ca^{2+}}(0). \tag{5.6}$$

Also shown in the figure is a transconductance amplifier connected in the follower configuration with its output connected to one terminal of the capacitor. This provides for a leakage current, whose time constant is determined by the g_m (and thus is controllable by modifying the bias voltage) of the amplifier and the capacitance of the capacitor. The voltage V_{Rest} determines the resting level of the capacitor voltage when there is not current flowing onto it(and none leaking off).

5.3 Threshold Detection

When the concentration of Ca^{2+} ions in the spine head reaches certain levels, LTP or LTP may be triggered. Thus, it is desirable to have some type of threshold detection to determine when V_{Ca}^{2+} has crossed certain voltage levels. Thus, a circuit of the type shown in Figure 5.4 is desired. Each comparator compares V_{Ca}^{2+} to a reference voltage and produces a digital 1 if V_{Ca}^{2+} is greater than the reference and a digital 0 if V_{Ca}^{2+} is less than the reference. The voltage references are generated by a resistive voltage divider of the power supply voltage. This assumes that the comparator draws little current at its input

terminal(a very reasonable assumption). This technique does dissipate static power, however. This can be reduced by selecting large-valued resistors, but these would require significant area on an integrated circuit.The voltage references could also be generated using more sophisticated techniques utilizing band-gap or zener diode references. The comparators can be implemented using commercial comparators, or more simply, a high gain differential voltage amplifier whose output is connected to a digital buffer. When the inputs of the amplifier are slightly different, the amplifier saturates and the output voltage is very near one of the power supply rails. The digital buffer then assures that a good logic value is obtained, as well as buffering the amplifier output.

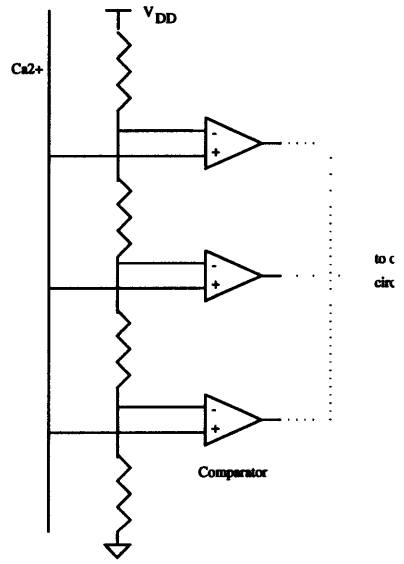

Figure 5.4: Threshold detection circuit

A simple alternative to a full-blown comparator is to use only digital buffers(which is essentially an amplifier) to compare V_{Ca}^{2+} to a reference. Each digital buffer would assume the role of the comparator in Figure 5.3, while the voltage reference would be the switching voltage of the digital buffer. This can be adjusted by appropriately changing the relative sizes of the p-channel and n-channel devices in the buffer.

However, this has the disadvantage that the references are subject to mismatch between transistors and are not adjustable once implemented.

5.4 Non-NMDA Channel

The basic circuit implementation of the non-NMDA receptor channel is shown in Figure 5.5 The output current of the transconductance amplifier is:

$$I_{out} = I(t)\tanh\left(\frac{1}{2}\frac{V_{Ref} - V_{Head}}{nV_T}\right). \tag{5.7}$$

Approximating the tanh with its argument yields

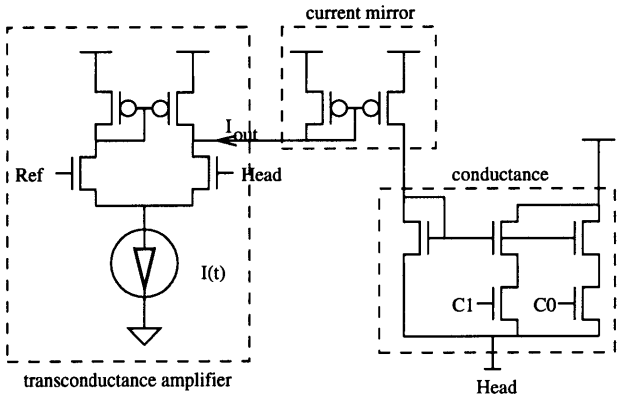

Figure 5.5: Circuit implementation of non-NMDA receptor channel

$$I_{out} = I(t)\frac{1}{2}\frac{V_{Ref} - V_{Head}}{nV_T} \tag{5.8}$$

which has the same form as the desired form of Equation 2.4 For consistency, the voltage range used in this circuit is the same as that used in the NMDA circuit discussed in Section 5.1. This implies that the wide input range amplifier should be used instead of the simple transconductance amplifier. This current is then scaled and mirrored by the p-channel current mirror(which is required since the output voltage of the amplifier is near the supply voltage) into a set of conductance transistors.

These conductance transistors form what is essentially an n-channel current mirror with multiple legs, in the form of a simple digital-to-analog converter. The legs of the current mirror contain control transistors that turn on or off each leg of the conductance; in other words, these binary control inputs regulate the amount of peak amount of non-NMDA current that flows into the spine head. Each set of control input values corresponds to the different conductance levels associated with LTP and LTD. In Figure 5.5, each succeeding leg of the D/A converter is scaled by a factor of two, so that the conductance changes are binary, corresponding to the control values. The peak non-NMDA current is plot as a function of spine head voltage(for a nominal set of control inputs) in Figure 5.6. Ideally, the slope of the curve should be constant, as the non-NMDA conductance is not a function of the spine head voltage. Nonetheless, the slope, and hence, conductance, is relatively constant over the range of spine head voltages.

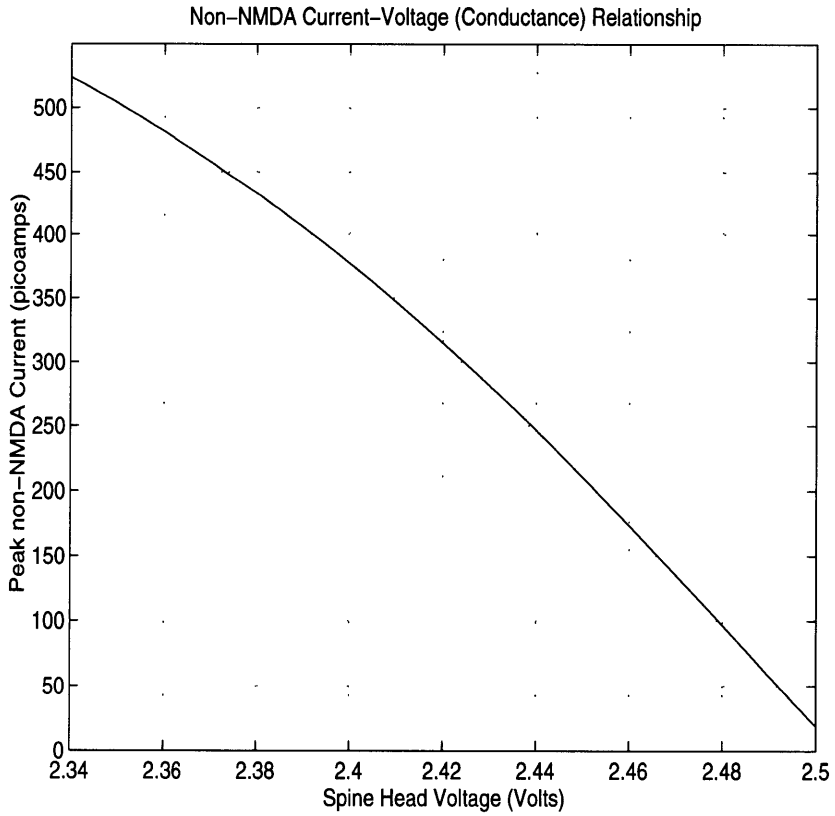

Figure 5.6: Peak non-NMDA current as a function of spine head voltage.

In real neurons, the conductance changes associated with LTP and LTD occur smoothly in value and time. However, in this circuit, the conductance changes in discrete steps.In the circuit implementation of the non-NMDA channel shown in Figure 5.5, only two control inputs are shown, although more could be added to increase the number of different conductance levels. As the number of control inputs increase, the size of the discrete steps in conductance values decrease, and thus the circuit more closely approximates the smoothness of real neurons. Thus, there is a trade-off between silicon area (number of transistors) and desired precision and smoothness of the conductance values.

Recall that induction of LTP or LTD, and hence the regulation of the control inputs in the non-NMDA circuit, is determined by the Ca2+ concentration in the spine head. The control inputs, which are digital in nature, are thus determined in part by the threshold detection circuitry discussed in Section 5.3 The outputs of that circuit are manipulated digitally, using information about the current conductance value, to produce the control inputs for the non-NMDA circuit. This digital circuitry is discussed in Section 5.6.

5.5 Implementation of time varying current

Thus far, the implementation of the time varying current I(t) has not been discussed. The NMDA and non-NMDA channels open in response to a presynaptic stimulus -- the action potential of another neuron. For each stimulus, current flows according to Equations 2.1 and 2.2. The time dependence for each of these equations is an alpha function. These can be treated as the impulse responses of the spine head channels to a presynaptic stimulus. The Laplace transform of both impulse responses can be written as

$$H(s) = \frac{1}{(1+\tau_1 s)(1+\tau_2 s)} \tag{5.9}$$

For the non-NMDA channel, $\tau_1 = \tau_2$.

I(t) can be implemented in a number of ways. One way of approximating the alpha function behavior is by applying the presynaptic input to a pair of cascaded follower integrators, with the output of the second follower integrator connected to the gate of an NMOS transistor, as shown in Figure 5.7. Recall from Section 3.3.2 that the transfer function of a single follower integrator is

$$\frac{V_{out}}{V_{in}} = \frac{1}{1+s\tau} \tag{5.10}$$

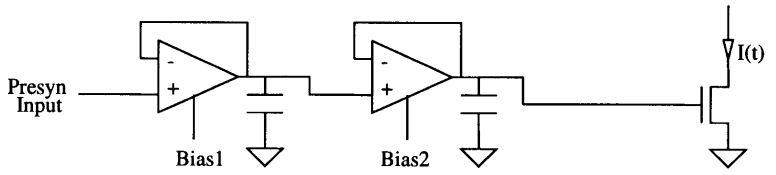

Figure 5.7: Cascaded follower-integrators

Thus, in Figure 5.?, the transfer function from input to the voltage on the gate of the transistor is

$$\frac{V_{GS}}{V_{in}} = \frac{1}{(1+\tau_1 s)(1+\tau_2 s)}, \tag{5.11}$$

which is of the same form as Equation 5.9. If the input voltage is a unit impulse, then the gate voltage on the NMOS transistor will change incrementally like an alpha function.

If this change in gate voltage is small, then the NMOS transistor can be analyzed using the small signal incremental model of the MOSFET in subthreshold shown in Figure 3.? Thus, the incremental change in current through the transistor will be approximately

$$\Delta I(t) = g_m \Delta V_{GS} \tag{5.12}$$

Thus, if the change in V_{GS} is an alpha function, then, to first order so will the change in I(t). Of course, since the current through the NMOS transistor is changing exponentially, because it is in subthreshold, this linearized model of the transistor is valid for only small changes in V_{GS}. While V_{GS} may change enough to violate this, Equation 5.12 is still a

good first order approximation of the change in current through the transistor. However, the inherit nonlinearities of CMOS transistors limit the accuracy to which an alpha function time dependence can be approximated.

While in practice creating an impulse is impossible, a pulse of unit area and of short duration is mathematically equivalent to a unit impulse as long as the duration of the pulse is shorter than all characteristic time constants in the circuit. Thus, if the presysnaptic were a short pulse, then the incremental change in V_{GS} will be of the form of an alpha function. The presynaptic input is, however, the nerve impulse, or action potential, of another neuron, as discussed in Chapter 4 and shown in Figure 4.5. If this action potential were buffered with a digital buffer, the output would be a short, essentially constant valued pulse as desired. The result of buffering the action potentials shown in Figure 5.8.

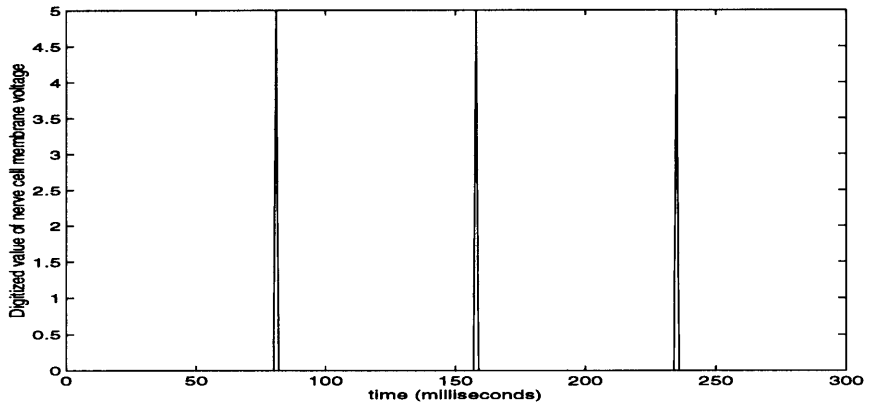

Figure 5.8: Buffered(digitized) action potentials

One problem with this implementation is that even when there is no presynaptic input, there is current that flows through the NMOS transistor. This arises from the fact that there is a non-zero voltage on the capacitors in Figure 5.7. This is because the output voltage of a simple transconductance amplifier in the follower configuration does not follow the input voltage when the input is zero. There exists some offset, as shown in the voltage sweep of Figure?. To minimize this problem, the simple transconductance amplifiers can

be replaced by wide output range amplifiers, which are capable of swinging closer to ground. Even in this case, there will be some current, albeit smaller, even when the input voltage is zero. This small amount of current can either be tolerated, or can be compensated by providing for a leakage transistor, as shown in Figure 5.9. The bias voltage for the leakage can be determined by sensing circuitry or it may be externally set.

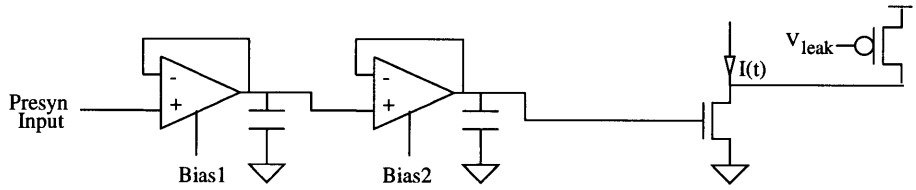

Figure 5.9: Cascaded follower-integrators with leakage transistor.

The time varying current I(t) can also be implemented by using a current mirror similar to that used for the non-NMDA conductance, with control inputs determining when each leg of the current mirror conducts, as shown in Figure 5.10. The size of the transistor in each successive leg is scaled by two to provide a binary implementation. The control inputs can then be controlled so that the total output current approximates an alpha function. The accuracy of this implementation is improved by adding more legs in the current mirror. One drawback of this circuit is the requirement of digital control logic to determine the control inputs. On the other hand, this implementation can potentially provide a very accurate approximation of an alpha function current, and no current flows when there is no input, unlike the use of follower integrators.

If the alpha function time dependence is not critical, but rather the area under the alpha function(i.e., the total amount of charge), then I(t) can simply be a constant-valued pulse of current, which could be implemented using the circuit in Figure 5.10, but with only one control input. However, this implementation would be a deviation from the concept of using life-like principles in design the circuits.

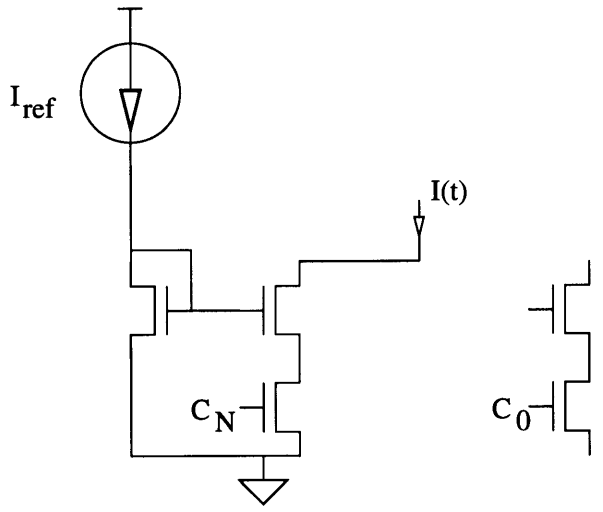

Figure 5.10: Alternative implementation of time-varying current

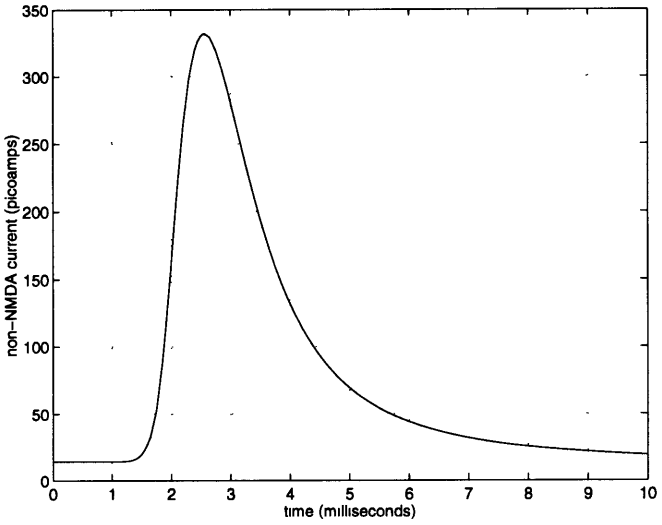

Figure 5.11: NMDA current as a function of time following a single presynaptic stimulus

Figures 5.5 and 5.6 show the NMDA and non-NMDA currents as a function of time following a single presynaptic stimulus. The time varying current was implemented using the cascaded follower-integrators of Figure 5.7, with no leakage transistor.

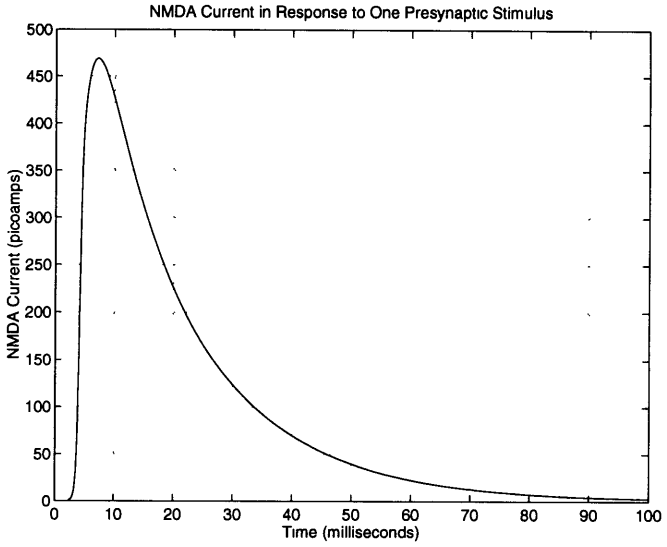

Figure 5.12: Non-NMDA current as a function of time following a single presynaptic stimulus

5.6 Digital Control Circuitry for Induction of LTP and LTD

The induction of LTP is characterized by a prolonged increase in the conductance of the non-NMDA receptor channel, while the induction of LTD is characterized by the decrease in conductance of the non-NMDA receptor channel. This is achieved in the circuit shown in Figure? by the control inputs which turn on or off the legs of the conductance transistors. The digital circuitry that determines what those control inputs are is discussed in this section.

LTP and LTD are initiated when the concentration of Ca^{2+} ions in the spine head reach certain levels. The circuitry discussed in sections 5.2 and 5.3 measure the concentration of Ca^{2+} ions in the spine head and then determine when it has reached certain levels. The outputs of the circuitry in Figure 5.4 are digital values. These digital signals are then used to determine the value of the control inputs, C1 and C0, in Figure 5.5. Thus the value of the non-NMDA conductance can range from its most depressed state, when C1=C0=0, to its most potentiated state, when C1=C0=1. The manner in which this is done depends upon the type of algorithm for LTP/LTD that is being implemented.

5.6.1 Simple Algorithm

A simple theory of the induction of LTP or LTD is that the conductance value of the non-NMDA current depends solely on the concentration of Ca^{2+} ions, regardless of the initial conductance state.The changes in conductance value are triggered by crossing one of the thresholds, Th0-Th3. Note that if none of the thresholds are crossed, then the conductance is in the nominal state. Depression of the conductance value is triggered only when the calcium concentration rises above the first threshold. This is because LTD can only be initiated when there is a stimulus(low-frequency). If there is no stimulus, there should be no reason for LTD(or LTP) to be initiated.

The truth table for this algorithm is shown in Table 5.1. Th0-Th2 are the outputs of the threshold detection circuitry described in Section 5. It turns out that for this algorithm, with only two control inputs C1 and C0, only three threshold values are necessary. Note that many of the rows of the truth table are "don't care." This is because the combinations of Th2-Th0 in those rows are not possible, since Th2-Th0 are correlated with one another. For instance, if Th2= 1, then that requires that Th1=1 and Th0=1.

Th2	Th1	Th0	C1	C0	Level of Potentiation
0	0	0	0	1	Normal
0	0	1	0	0	Depressed
0	1	0	*	*	Don't Care
0	1	1	1	0	Potentiated
1	0	0	*	*	Don't Care
1	0	1	*	*	Don't Care
1	1	0	*	*	Don't Care
1	1	1	1	1	Most Potentiated

Table 5.1: Truth table for simple LTP/LTD algorithm

From this table, logic expressions for C1 and C0 can be derived in terms of Th2-Th0:

$$C1 = Th1 \cdot Th0 = Th1 \tag{5.13}$$

$$C0 = \overline{Th2} \cdot \overline{Th1} \cdot \overline{Th0} + Th0 \cdot Th1 \cdot Th2 = \overline{Th0} + Th2\,. \tag{5.14}$$

It is important to note that C1 and C0 are not updated continuously in time, but rather at discrete intervals. The duration of the interval is the length of time for which LTP or LTD is in effect. The circuitry that accomplishes this is shown in Figure 5.13. The outputs of the logic circuits implementing Equations 5.13-14 are fed into registers, whose outputs C1 and C0, respectively. The period of the clock signal for the registers determines how long LTP or LTD is in effect.

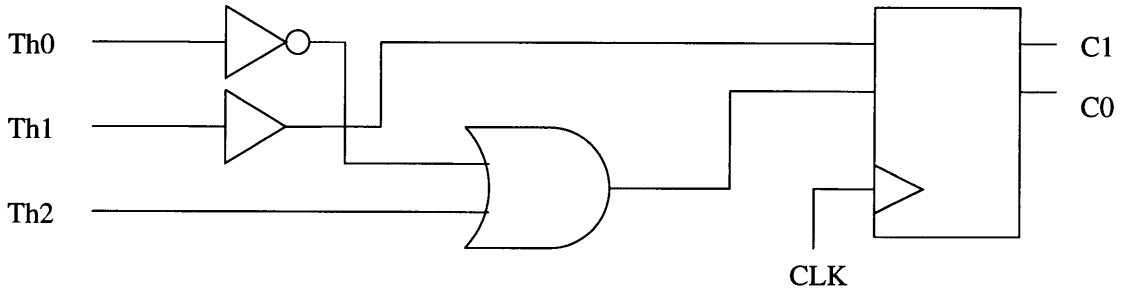

Figure 5.13: Circuit implementation of simple LTP/LTD algorithm

5.6.2 LTP/LTD Reversal Algorithm

One theory of LTP and LTD is that they can reverse each other. In other words, the current conductance state depends on the previous conductance state. If the non-NMDA conductance is initially in a potentiated state, and the calcium concentration dictates that LTD should be initiated, then the next conductance state would be a more nominal level, as opposed to fully depressed. With this algorithm, it is the change in conductance, rather than the actual conductance level, that is affected by the concentration of calcium ions. This dependence is shown qualitatively in Figure 5.14.

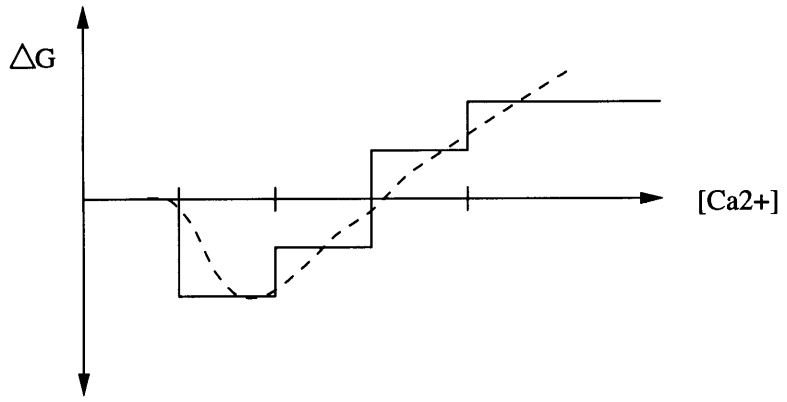

Figure 5.14: Dependence of the change in conductance on the calcium concentration

This can be expressed analytically as

$$G[n+1] = G[n] + \Delta G, \tag{5.15}$$

where G[n] represents the current conductance state.

Th3	Th2	Th1	Th0	DG2	DG1	DG0	Change in Potentiation
0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	-2
0	0	1	1	1	0	1	-1
0	1	1	1	0	0	1	+1
1	1	1	1	0	1	0	+2

Table 5.2: Truth table for threshold logic of the LTP/LTD reversal algorithm

For the four threshold outputs Th3-Th0 of Figure?, ΔG can be represented by a two's complement three-bit number DG2-DG0. The truth table for DG2-DG0 is shown in Table 5.2. DG2-DG0 represents a change in C1-C0 ranging from -2 to +2. From this truth table, logic expressions for DG2-DG0 can be derived in terms of Th3-Th0:

$$DG2 = \overline{Th2} \cdot Th0 \tag{5.16}$$

$$DG1 = Th3 + \overline{Th1} \cdot Th0 \tag{5.17}$$

$$DG0 = \overline{Th3} + Th1 \tag{5.18}$$

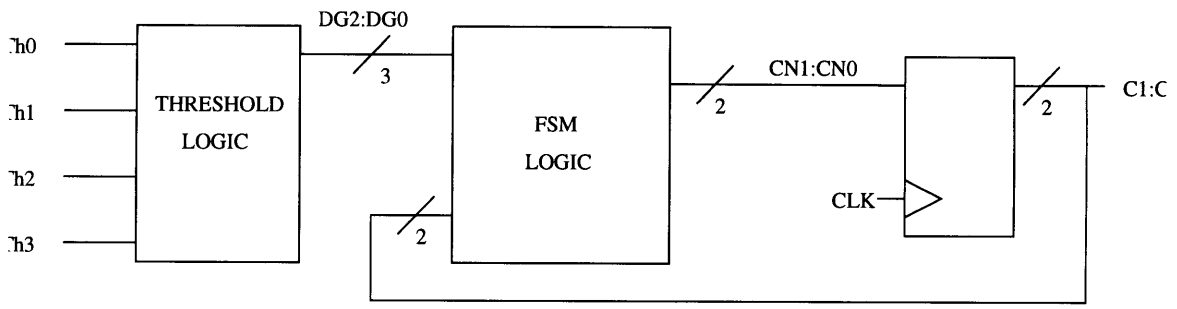

Figure 5.15: Block diagram of circuit implementation of LTP/LTD reversal algorithm

Figure? depicts a block diagram of the circuits that implement this LTP/LTD reversal algorithm. DG2-DG0 and the current conductance state control signals C1 and C0 are inputted into a FSM, whose outputs are the next control signals, CN1 and CN0. These are then fed into registers, which are clocked by a signal whose period determines the length of LTP or LTD. The truth table for the FSM that calculates CN1 and CN0 is shown in Table 5.3. From this truth table, logic expressions for CN1 and CN0 can be derived:

$$CN1 = C1 \cdot \overline{DG2} \cdot \overline{DG1} + C0 \cdot \overline{DG2} \cdot DG0 + DG2 \cdot DG1 \cdot \overline{DG0} + C1 \cdot C0 \cdot DG2 \cdot DG0 \tag{5.19}$$

$$= \overline{DG2} \cdot (C1 \cdot \overline{DG1} + C0 \cdot DG0 + DG1 \cdot \overline{DG0}) + C1 \cdot C0 \cdot DG2 \cdot DG0$$

$$CN0 = \overline{DG2} \cdot C1 \cdot C0 + \overline{C0} \cdot \overline{DG2} \cdot DG0 + C0 \cdot \overline{DG2} \cdot \overline{DG1} \cdot \overline{DG0} + C1 \cdot C0 \cdot DG1 + \tag{5.20}$$

$$C1 \cdot \overline{DG2} \cdot DG1 + C0 \cdot \overline{DG2} \cdot DG1 + C1 \cdot \overline{C0} \cdot DG2 \cdot DG0$$

$$= (C1 + C0) \cdot \overline{DG2} \cdot DG1 + C1 \cdot C0 \cdot (\overline{DG2} + DG1) + \overline{C0} \cdot DG0 \cdot (C1 + \overline{DG2}) + C0 \cdot \overline{DG2} \cdot \overline{DG1} \cdot \overline{DG0}$$

C1	C0	DG2	DG1	DG0	CN1	CN0
0	0	0	0	0	0	0
0	0	0	0	1	0	1
0	0	1	0	1	0	0
0	0	1	1	0	0	0
0	0	0	1	0	1	0
0	1	0	0	0	0	1
0	1	0	0	1	1	0
0	1	1	0	1	0	0
0	1	1	1	0	0	0
0	1	0	1	0	1	1
1	0	0	0	0	1	0
1	0	0	0	1	1	1
1	0	1	0	1	0	1
1	0	1	1	0	0	0
1	0	0	1	0	1	1
1	1	0	0	0	1	1
1	1	0	0	1	1	1
1	1	1	0	1	1	0
1	1	1	1	0	0	1
1	1	0	1	0	1	1

Table 5.3: Truth table for the FSM logic of the LTP/LTD reversal algorithm

Chapter 6

Circuit Implementation of Dendrite

This chapter discusses the circuit implementation of the dendrite. The dendrite is the structure that connects the synapses to the nerve cell membrane. In other words, synaptic current is injected through the dendrite onto the nerve cell membrane. The dendrite can be can be modeled electrically as a passive, highly resistive network.

When there is no synaptic current, there is no voltage drop across the dendrite, and thus the spine head (synaptic) membrane and the nerve cell membrane are at the same potential. The dendrite can be represented as a single effective resistor of high value (on the order of megaohms) instead of a resistive network, especially if only the endpoint (nerve cell and spine head membrane) voltages are of interest. However, implementing a resistor of such high value on an integrated circuit is not a simple task. The first two sections of this chapter discuss different ways of implementing such a resistance.

Since the dendrite circuit connects the nerve cell circuits (Chapter 4) to the synaptic circuits (Chapter 5), the voltage on each of these membranes is related to the voltage on the other membrane. However, the voltage scales in these circuits are different. Section 6.3 addresses this problem and presents its solution.

Section 6.4 presents the complete circuit implementation of the dendrite

6.1 Polysilicon Resistor

A resistor can be build on an integrated circuit by laying down a resistive layer such as polysilicon. Some fabrication processes include a layer of undoped polysilicon that has a very high resistance and could be used to implement the dendritic resistance. However, this layer is not available in the Orbit 2 micron process available from MOSIS that is used

to layout the circuits in this project. Ordinary doped polysilicon or diffusions cannot produce high enough resistances to be of any use for this application, as they would simply take up too much silicon area. Thus, another method, described in the next section, must be used to implement the dendritic resistance.

6.2 Horizontal Resistor

The horizontal resistor circuit is a circuit developed by Carver Mead[4] that can approximate a large valued resistor whose resistance can be set electronically with a control input.

6.2.1 Basic Topology

This circuit consists of two pass transistors connected in series. The gate voltage of each transistor is set at a fixed value V_q above the respective input voltage, V_1 or V_2, such that

$$V_q = V_{G1} - V_1 = V_{G2} - V_2 \tag{6.1}$$

Figure 6.1: Horizontal resistor network

This voltage controls the current through the pass transistors and thus sets the effective resistance of the circuit. How V_q is set will be discussed in the next subsection. The current through the pass transistors is given by Equation 3.1. Applying this equation to the horizontal resistor circuit, one obtains

$$I = I_0 e^{\frac{V_q}{nV_T}}\left(1 - e^{\frac{V_2 - V_n}{nV_T}}\right) = -I_0 e^{\frac{V_q}{nV_T}}\left(1 - e^{\frac{V_1 - V_n}{nV_T}}\right) \tag{6.2}$$

From this equation, the voltage V_n at the node connecting the two pass transistors can be determined:

$$2e^{\frac{V_n}{nV_T}} = e^{\frac{V_1}{nV_T}} + e^{\frac{V_2}{nV_T}} \tag{6.3}$$

Substituting this equation into Equation 6.2 yields

$$I = I_0 e^{\frac{V_q}{nV_T}} \frac{e^{\frac{V_1}{nV_T}} - e^{\frac{V_2}{nV_T}}}{e^{\frac{V_1}{nV_T}} + e^{\frac{V_2}{nV_T}}} = I_0 e^{\frac{V_q}{nV_T}} \tanh\left(\frac{1}{2}\frac{V_1 - V_2}{nV_T}\right) \tag{6.4}$$

For small differences in V_1 and V_2, the circuit looks like a resistor with effective resistance

$$R = \frac{2nV_T}{I_0 e^{\frac{V_q}{nV_T}}} \tag{6.5}$$

6.2.2 Bias Circuit

This section discusses the circuitry that produces the bias voltages at the gates of the pass transistors of the horizontal resistor circuit. The bias circuitry must adjust the gate voltage such that V_q defined in Equation 6.1 is held constant even when the node voltages V_1 and V_2 change. The circuit that achieves this function for VG1 is shown in Figure 6.2. An identical circuit with V2 as the input generates the voltage VG2.

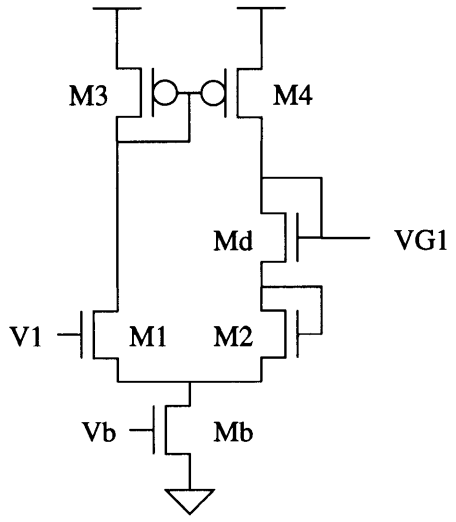

Figure 6.2: Biasing circuitry for horizontal resistor connection

The horizontal resistor bias circuit is essentially a transconductance amplifier connected in the follower configuration, with the addition of the diode connected transistor Md. The gate voltage of M2 follows the voltage V1. The voltage VG1 is thus equal to the

voltage V1 plus an offset equal to the voltage across the diode connected Md. This offset is the voltage required to maintain the current through Md, which is half of the bias current Ib. By Equation 6.1, this offset is in fact equal to V_q. The current through Md can be expressed as

$$\frac{I_b}{2} = I_0 e^{\frac{V_{G1} - V_1}{nV_T}} = I_0 e^{\frac{V_q}{nV_T}} \tag{6.6}$$

The dependence on the drain voltage in Equation 3.1 is negligible since the transistor is diode connected. Hence, Equation 6.6 relates V_q to the bias current I_b. In other words, setting the gate voltage of the bias transistor Mb determines V_q, independent of V_1 or V_2. This in turn determines the resistance of the horizontal resistor connection, given by Equation 6.5. Combining Equations 6.5 and 6.6 yields an expression for the effective resistance of the connection that is solely dependent on I_b and the thermal voltage:

$$R = \frac{4nV_T}{I_b} \tag{6.7}$$

6.2.3 Application

Since I_b can depends exponentially on the gate voltage of Mb, the effective resistance of the horizontal resistor connection can vary over many orders of magnitude. The current I is plotted in Figure 6.3 as a function of bias voltage for a V_1-V_2 equal to 50 millivolts. As the bias voltage increases, more current flows, corresponding to a decrease in effective resistance.

The current I of the horizontal resistor configuration is shown in Figure 6.4 as a function of V_1-V_2. For V_1-V_2 less than approximately 100 millivolts, the slope of the curve is relatively constant, corresponding to a constant effective resistance given by Equation 6.7. For V_1-V_2 greater than 100 millivolts in magnitude, the current saturates, and the circuit loses its linear resistive behavior.

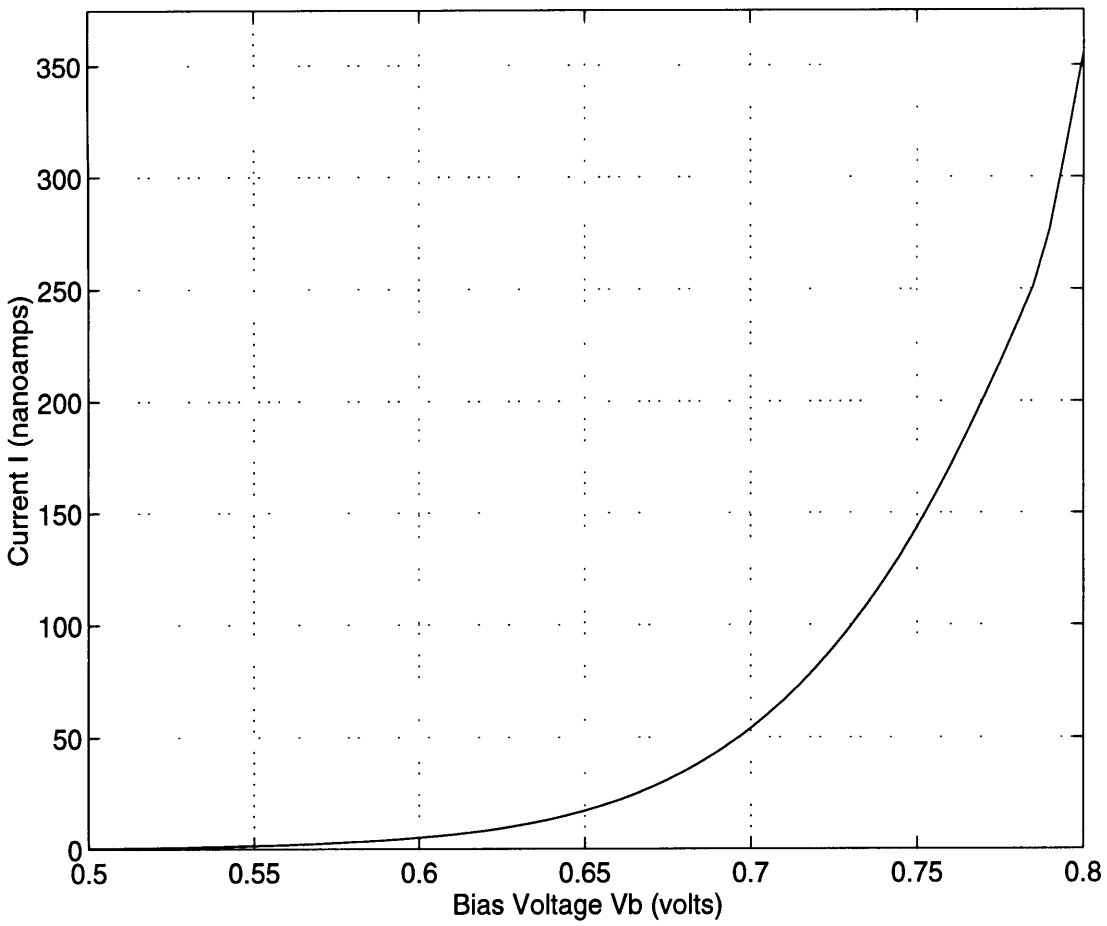

Figure 6.3: Simulated(HSPICE) current in the horizontal resistor circuit as a function of bias voltage. (V1 - V2 = 0.05 V)

This 100 millivolt linear range of the horizontal resistor configuration may not be enough in certain applications. However, multiple horizontal resistor circuits can be placed in series to increase the range over which the circuits behave like a constant resistance.

The horizontal resistor circuit is a practical and flexible way to implement the dendritic resistance of the neuron. A single connection may suffice, although multiple circuits may be cascaded to increase the linear resistive range. However, this is probably not necessary since the spine head voltage and the nerve cell voltage do not differ much in volt-

age, except at the instance of an action potential, which is of relative short duration. Thus the saturation of the resistor current at this time is tolerable

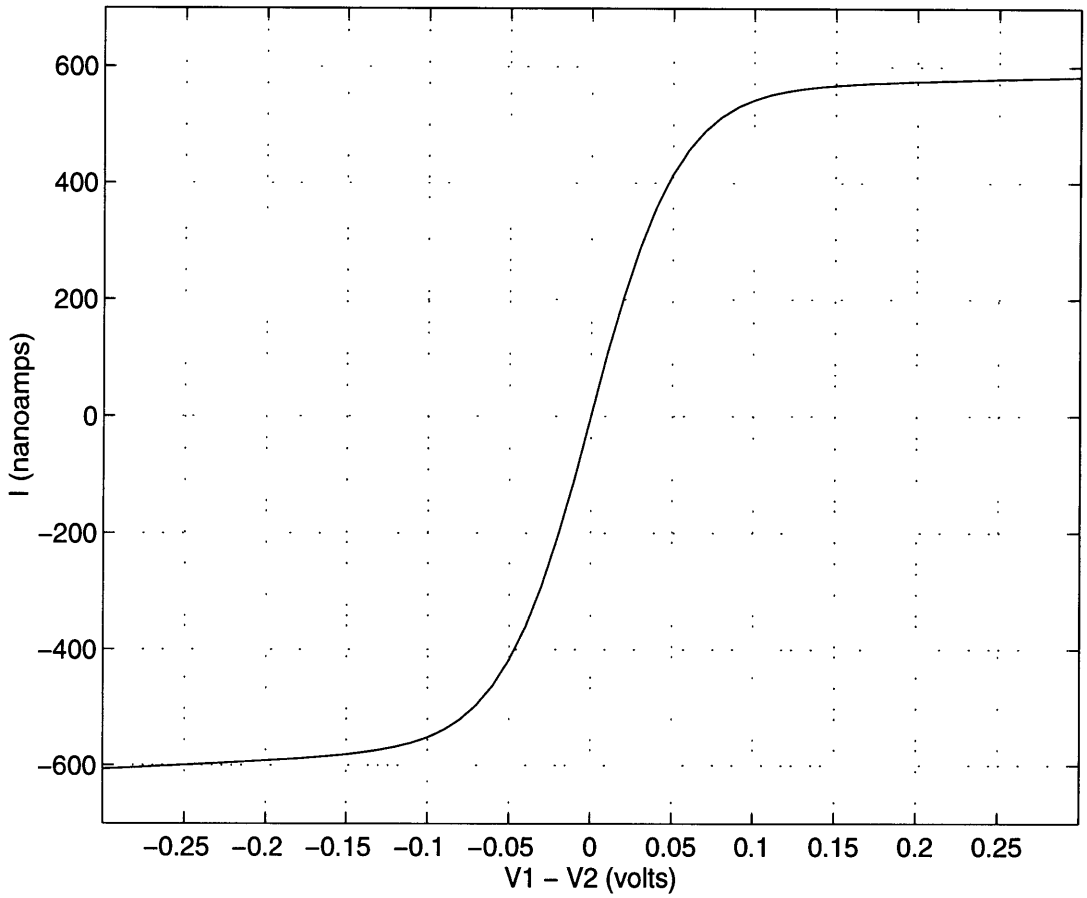

Figure 6.4: Simulated(HSPICE) current in the horizontal resistor circuit as a function of difference in input node voltages. (Vb = 0.7 V)

6.3 Scaling Circuitry

The dendritic circuit connects the nerve membrane circuits (Chapter 4) to the spine head(synaptic) circuits (Chapter 5). However, these two different groups of circuits operate over different voltages ranges, given by Equations 4.2 and 5.4. both voltage ranges possess the same ground, or reference, voltage, equal to 2.5 volts. However the nerve membrane circuits operate over a range that is five times as big as the range for the spine head circuits. Recall that this difference in voltage scale resulted from the different expo-

nential dependence on voltage that the two different types of circuits exhibited. Since the nerve membrane circuits are connected to the spine head circuits through the dendritic resistance, this difference in voltage scales is a problem.

However, there is a solution. First, note that the spine head node in the circuits of Figures 5.1 and 5.4 is connected to the gates of transistors and to the drains of transistors. The connections to the transistor gates are essentially the inputs of the circuit, while the connections at the transistors drain serve as the outputs. These outputs essentially behave like current sources; in other words, the output current is negligibly affected by the head voltage at these drain connections. Thus, at the drain connections, the voltage scale does not really matter, so this can be on the same scale as nerve cell membrane circuit. The spine head input connections at the transistor gates must be on the voltage scale given by Equation 5.4 for the spine head circuits. Figures 6.5 and 6.6 show the NMDA and non-NMDA circuits of Chapter 4 relabeled with the input spine head nodes (HEADIN) and the output notes (HEADOUT).

The HEADOUT node can be connected to one end on the horizontal resistor network, with the other end connected to the nerve cell membrane. The HEADIN voltage must be scaled version of the HEADOUT voltage to achieve the necessary voltage scale. In particular HEADIN must be related to HEADOUT according to

$$V_{HEADIN} = V_{Ref} - \frac{1}{5}(V_{Ref} - V_{HEADOUT}) \tag{6.8}$$

where V_{Ref} is equal to 2.5 volts. This behavior can be implemented by using the circuit shown in Figure 6.5. The amplifiers shown are operational amplifiers, although any amplifier with large gain, small output resistance, and very large input resistance(i.e., a MOS amplifier) should suffice. This circuit, although ultimately envisioned to be implemented on an integrated circuit, was for this project implemented using discrete, off-chip, compo-

nents. In particular, the operational amplifier part used was a LMC684, consisting of four low power MOS input stage operational amplifiers.

Operational amplifier (op-amp) #1 is connected in the unity buffer configuration. Its input voltage is VHEADOUT, which is thus equal to its output voltage V1. The buffer configuration is used so that the scaling circuitry does not load the other circuitry and draw current from it. Op-amp #4 is also connected in the unity buffer configuration. Its input voltage is one half of the power supply voltage of 5 volts, created by the resistors labeled R1. This op-amp is used to generate the reference voltage Vref = 2.5 volts. Op-amps #2 and #3 are each connected in the differential amplifier configuration. It can be shown that the output voltage of op-amp #2 is

$$V_2 = \frac{R_2}{R_1}(V_{Ref} - V_1) = \frac{R_2}{R_1}(V_{Ref} - V_{HEADOUT}) \tag{6.9}$$

The output voltage of op-amp #3 is similarly

$$V_{HEADIN} = V_{Ref} - V_2 = V_{Ref} - \frac{R_2}{R_1}(V_{Ref} - V_{HEADOUT}) \tag{6.10}$$

which is exactly the same as the desired Equation 6.8 with $R_2/R_1 = 1/5$.

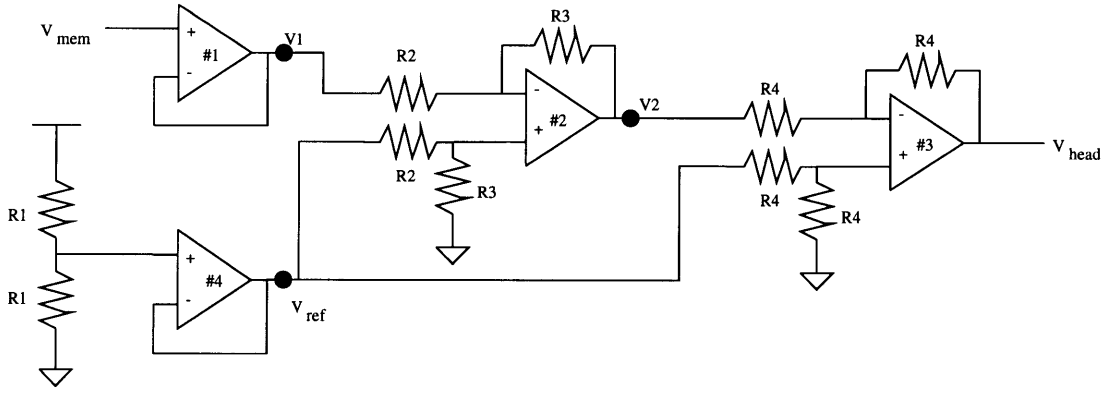

Figure 6.5: Scaling circuitry. R1=R4=100 kilohm; R2=22 kilohm; R3=110 kilohm

Chapter 7

Simulation of Complete Neural Circuit

7.1 Simulation setup

A complete neuron consisting of a cell membrane, dendrite, and a synapse was simulated. The spine head node of the synapse circuit was connected to one node of the dendrite circuit. The other node was then connected to the nerve cell membrane node, as discussed in Chapter 6.

7.2 Simulation Results

In the first simulation, presynaptic stimuli at 100 Hz. were inputted into the synapse circuit for a duration of one second, as shown in Figure 7.1 The nerve cell membrane was also stimulated by injecting a constant current into the membrane. This caused the voltage

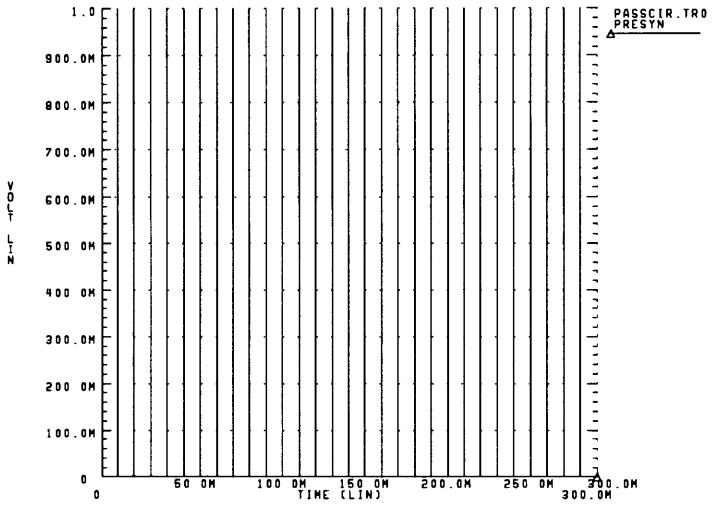

Figure 7.1: Presynaptic stimuli at 100 Hz.

to change and action potentials to be generated. This voltage is shown in Figure 7.2.

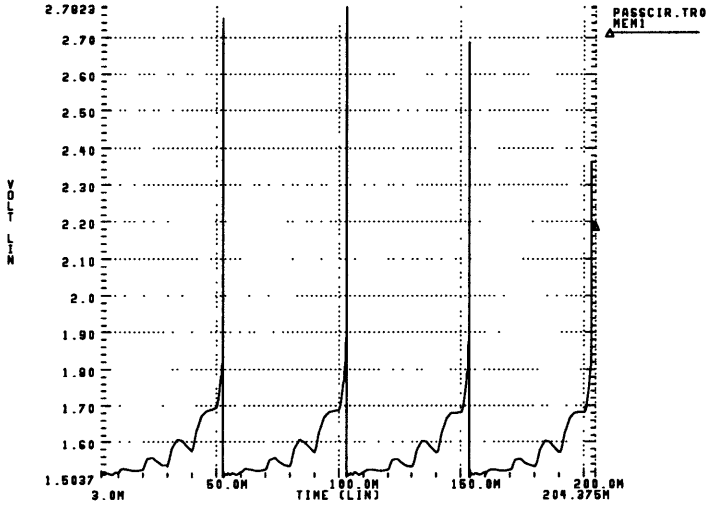

Figure 7.2: Nerve cell membrane voltage

This current was injected so that the membrane voltage would be changing significantly, representing many other synapses injecting current.

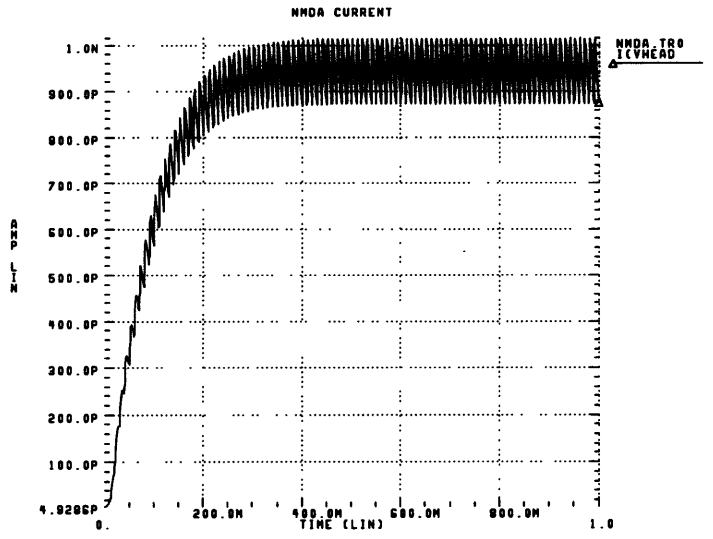

Figure 7.3: NMDA current

The resulting NMDA current is shown in Figure 7.3. This current is integrated to

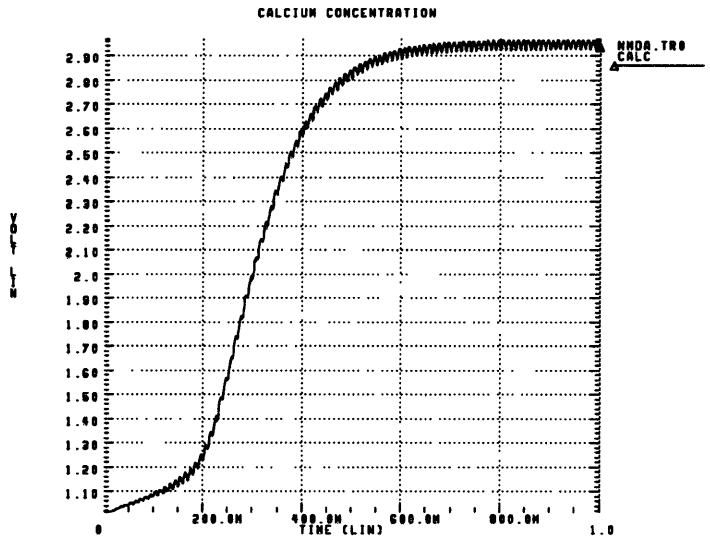

Figure 7.4: Calcium concentration

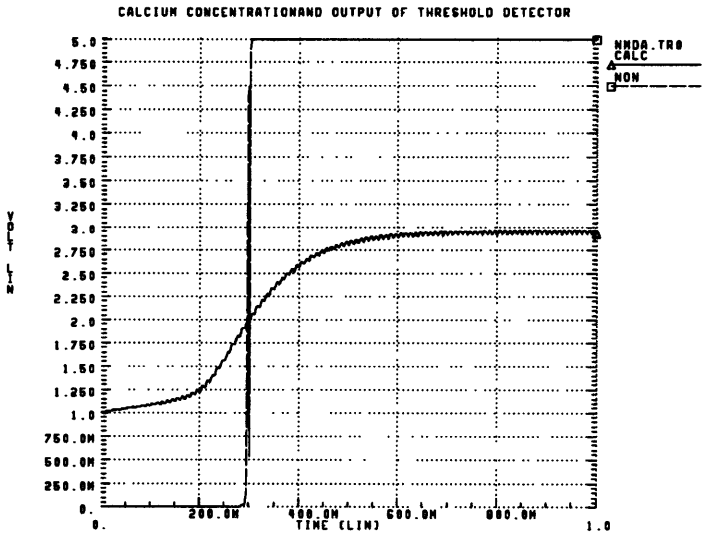

Figure 7.5: Calcium concentration and threshold detector output(Th1)

form the calcium concentration, with leakage, as shown in Figure 7.4. This voltage is then compared to the multiple thresholds. In this simulation the second threshold is crossed, resulting in an increase in conductance, corresponding to the induction of LTP. The threshold detector output (Th1) is shown with the calcium concentration in Figure 7.5. The induction of LTP results in the modification of the non-NMDA conductance. Figure 7.6 illustrates the effect. The non-NMDA current is shown as a function of time in response to a single presynaptic stimulus, before and after the induction of LTP. The peak current

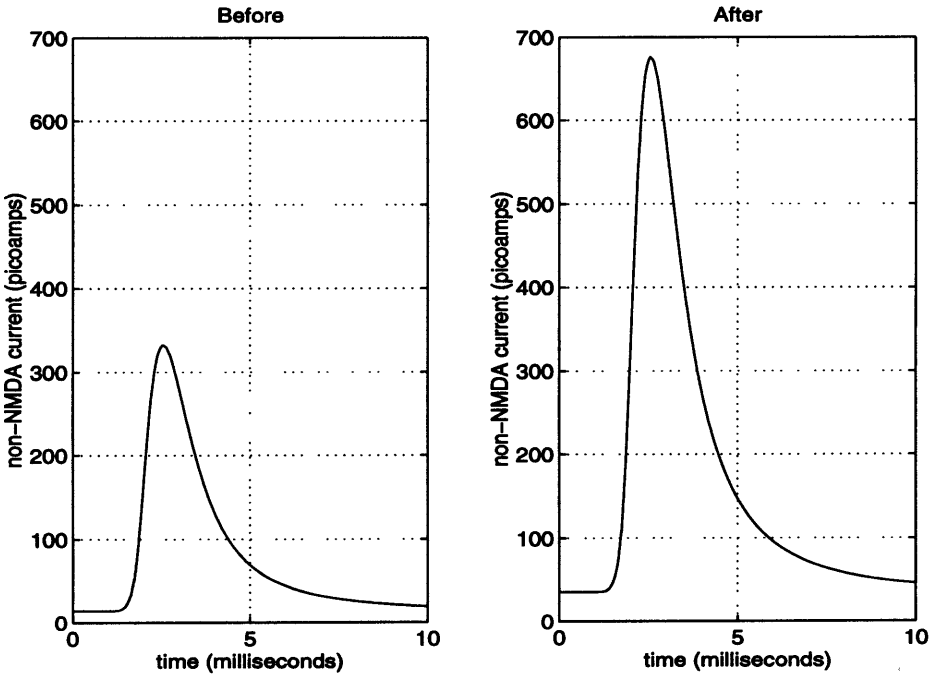

Figure 7.6: Change in non-NMDA conductance due to LTP

current is proportional to the conductance of the channel. Hence, after the sustained stimulation at 100 Hertz for 1 second, the conductance of the non-NMDA channel increases. The resulting change in postsynaptic potential (EPSP) is shown in Figures 7.7 and 7.8. Figure 7.7 depicts an EPSP before the induction of LTP, while Figure 7.8 depicts an EPSP after LTP has been induced.

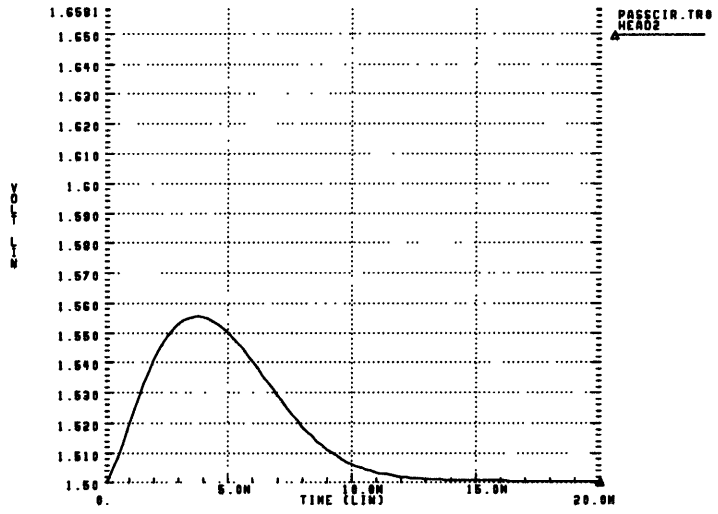

Figure 7.7: Postsynaptic membrane potential following a single presynaptic stimulus

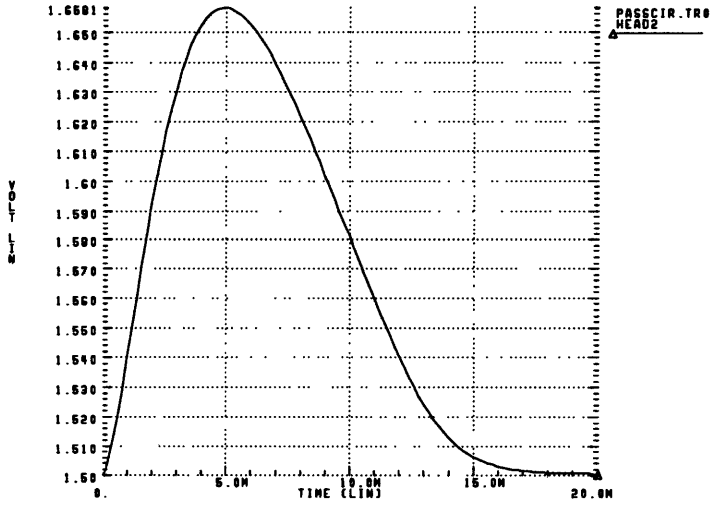

Figure 7.8: Postsynaptic membrane potential following a single presynaptic stimulus, after induction of LTP.

LTD can be induced by low-frequency presynaptic stimulation at 1 Hertz. Figure 7.9 depicts the non-NMDA current before and after LTP is induced. The channel conductance is decreased. The corresponding EPSP is shown if Figure 7.10.

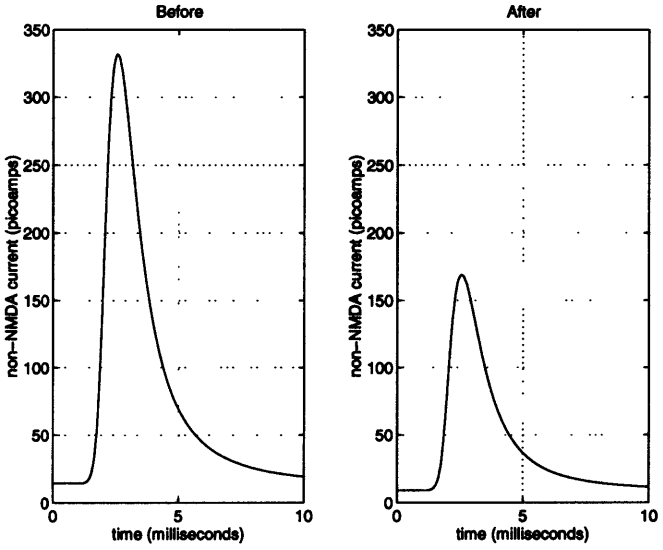

Figure 7.9: Change in non-NMDA conductance due to LTD

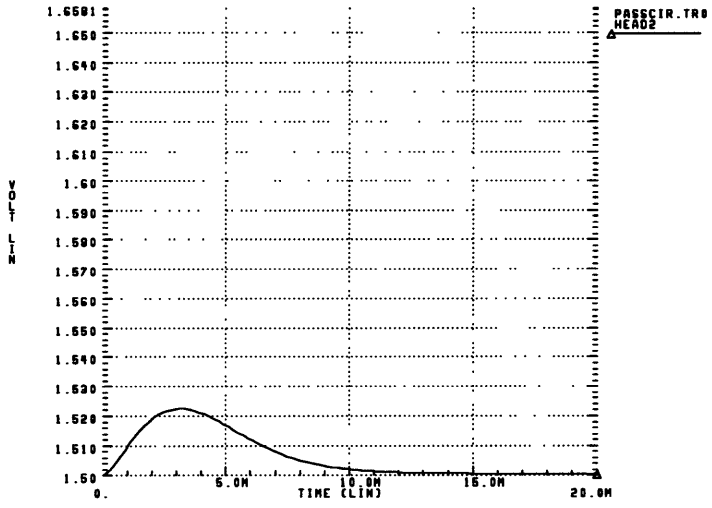

Figure 7.10: Postsynaptic membrane potential following a single presynaptic stimulus, following induction of LTD

Chapter 8

Layout and Fabrication

8.1 Fabrication Technology

The circuits discussed in the previous chapters were designed to be fabricated using the Orbit 2-micron double poly n-well process that is available for fabrication through the MOSIS service. The MOSIS service was selected as the fabrication service because of its relative cheapness and availability. Also, the MOSIS service does not require the customer to purchase a large quantity of finished parts. The minimum quantity is only 4.

The layout was conducted using the program MAGIC, a part of the U.C.-Berkeley CAD software series. It was selected because of its availability, cost(free), and simplicity. It was also chosen because the first round of fabrication would only include a small number of "neuron" circuits, intending to demonstrate their basic functionality. The layout was done using the scalable CMOS (SCMOS) design rules with lambda of 1 micron.

8.2 Layout

The layout that was submitted to MOSIS for fabrication contained one complete nerve cell membrane circuit, all of the synaptic circuits with the exception of the digital control circuitry, and the horizontal resistor circuit. The digital control circuitry of the synaptic circuit was to be implemented off chip using field programmable gate arrays. The threshold circuitry of the synapses consisted of four CMOS buffers, each with different switching thresholds. The scaling circuitry of the dendritic circuit was implemented off-chip using discrete components. All of the bias voltages were left to be set off-chip. In other words, there were no on-chip voltage references. This was done to increase the flexibility in testing. It would also be beneficial to be able to alter many of the reference voltages, which set

conductances and time constants, in an experimental setting.

Most of the circuits discussed in this thesis consist of transconductance amplifiers, current mirrors, differential pairs, and capacitors. In general, this building blocks were laid out and then used, with appropriate scaling when appropriate. The layouts of some of these building blocks are shown in the appendix.

Care was taken to ensure that the transistors in differential pairs and current mirrors were as matched as possible, using techniques such as interleave fingers. But in general, slight mismatches are tolerable, and also make the circuits more life-like.

The capacitors were laid out using the two layers of polysilicon that were available in the Orbit 2 micron process. Poly-poly capacitors have capacitances that are nearly independent of voltage, and thus can be operated over the full five volt voltage range. In contrast, poly-diffusion capacitors are highly nonlinear with respect to voltage. It was not critical to lay out especially precise capacitors, as they generally served to form a time constant with a conductance whose value could be externally set. Furthermore, there was no need to match any capacitors, so this was not an issue.

Care was also taken to protect against CMOS latchup, which results from the intrinsic bipolar transistors in CMOS technology. To avoid this problem, the resistances between the power supply and the wells and substrates should be minimized. To achieve this, numerous well and substrate contacts must be placed close to the source connections of both the NMOS and PMOS transistors[7].

8.3 Pin-out/Bonding Diagram

The pad frame used in the layout was the 40pc22x22_analog frame (tinychip) available from the MOSIS service. Its cavity is 2.2 mm. square. It is designed to be packaged in a 40 pin DIP (dual inline package) package. The frame contains the bonding pads for each pin

and electrostatic discharge protection circuitry. This size frame was chosen since the test circuits to be fabricated would not take up that much silicon area, and it was the cheapest fabrication size available. The pinout and bonding diagram are shown in the appendix.

8.4 Testing

The fabricated parts received from MOSIS were tested using analog test equipment. Unfortunately, the testing did not prove successful. There seemingly was a problem in packaging or bonding, or a fatal flaw in the fabrication process. Pins on the part that were supposed to be connected to gates of MOS transistors drew currents on the order of milli-amps when the chip was powered up. No useful or recognizable signals were able to be observed in the testing of the chip.

8.5 Estimated Layout Density

Roughly 15 silicon neuron-synapse pairs can be integrated on one MOSIS tinychip (die area 2.2 mm. x 2.2 mm.), with the 2 micron process, if the digital control circuitry/memory is implemented off-chip, using FPGAs. If this circuitry is included on-chip, then roughly 10 neuron-synapse pairs can be integrated on the MOSIS tinychip. Considering that the density of an integrated circuit is roughly inversely proportional to the square of the minimum feature size, a decrease in feature size to 0.25 or 0.35 micron (typical industrial process, then the number of silicon neuron-synapse pairs that can be integrated on a 2.2 mm. x 2.2 mm. die is increased by a factor of roughly 100. Thus, depending whether the digital control circuitry for the synapse is implemented off or on chip, approximately 1000 to 1500 silicon neuron-synapse pairs can be implemented on a die area of 2.2 mm. x 2.2 mm. Die areas can typically be 10-100 times larger than this. However, as one increases the die area, and pinout, the number of circuits capable of being integrated may not scale linearly, especially when most of the neuron-synapses pairs will be intercon-

nected with one another. The increase wiring complexity would consume a large amount of silicon area. A conservative estimate of the maximum number of silicon-synapse pairs that may be integrated onto a single integrated circuit would thus be on the order of 10000.

Chapter 9

Conclusion

9.1 Summary

CMOS circuits have been designed that emulate the electrical characteristics of actual neurons. In particular, these circuits emulate the nerve cell membrane, the dendritic structure, and a particular type of synapse called a Hebbian synapse. These circuits are more neuromorphic compared to most analog neural networks. The nerve cell membrane circuits contain sections that are represent the sodium and potassium ion channels in the membrane. The synapse circuits contains sections that correspond to different types of synaptic current channels. Circuits were developed that, through a feedback mechanism, modify the synaptic conductance, or strength.

The circuits have all been simulated with HSPICE. Also, simulations of a complete neuron, incorporating the nerve cell membrane, dendrite, and a single synapse circuit were conducted. In reality, a single neuron would have many synapses, but such simulations are difficult in HSPICE due to the total number of transistors. The circuits were also layed out in silicon and submitted for fabrication. The resulting fabricated device did not function, although this was not a design flaw.

9.2 Future Work

The first task that must be accomplished is to obtain a functioning fabricated device. Assuming this is functional, hardware simulations of small neural networks can be performed, with the ultimate goal of emulated a large biological neural network. Experiments based on physiological principles can be performed to test different learning algorithms.

At the transistor level, circuits can be designed to generate many of the voltage references that are needed. While some of these references are intended to be variable, in cer-

tain cases it may be desired that they are fixed. Also, circuits can be developed to emulate different types of synapses, and different learning techniques, including those involving sliding thresholds. Nonetheless, the circuits presented are an important first in building an artificial neural network that emulates the analog behavior of real neurons.

References

[1] Douglas, R. and Mahowald, M. "A Silicon Neuron." *Nature*, Vol. 354, pp. 515-518, Dec. 19-25, 1991.

[2] Jabri, M.A. and Schultz, S.R. "Analogue VLSI 'integrate-and-fire' neuron with frequency adaptation." *Electronics Letters*,Vol. 31, No. 16, pp.1357-1358, 1995.

[3] "Synaptic Plasticity:LTP and LTD."*Current Opinion in Neurobiology*, 4, pp. 389-99, 1994.

[4] Mead, Carver. *Analog VLSI and Neural Systems*.Reading, MA: Addison-Wesley, 1989.

[5] Douglas, R. and Mahowald, M. "A Construction Set for Silicon Neurons." A*n Introduction to Neural and Electronic Networks, 2nd Ed.* San Diego: Academic Press, 1995.

[6] Zador, A., Koch, C., and Brown, T. "Biophysical Model of a Hebbian Synapse." *Proc. Natl. Acad. Sci.*, Vol. 87, pp. 6718-22, Sept. 1990, Neurobiology.

[7] Rabaey, Jan M. *Digital Integrated Circuits: A Design Perspective*. Prentice-Hall, 1996.

[8] Grebene, Alan B. *Bipolar and MOS Analog Integrated Circuit Design*. New York: Wiley-Interscience, 1984.