12-1999

Exploring the information processing capabilities of random dendritic neural nets

Tzusheng Pei

To the Graduate Council:

I am submitting herewith a dissertation written by Tzusheng Pei entitled "Exploring the information processing capabilities of random dendritic neural nets." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Computer Science.

Bruce MacLennan, Major Professor

We have read this dissertation and recommend its acceptance:

Jim Drake, Jim Hall, Michael Vose

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

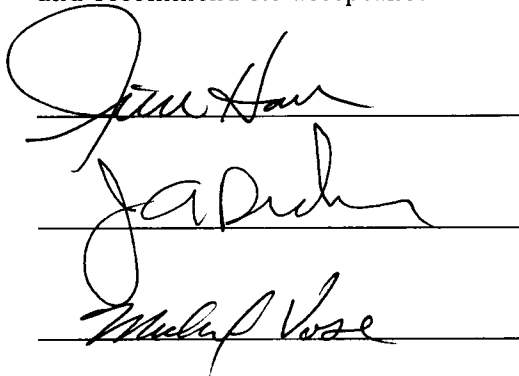(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Tzusheng Pei entitled "Exploring the Information Processing Capabilities of Random Dendritic Neural Nets." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Computer Science.

Bruce MacLennan, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor
and Dean of the Graduate School

Exploring the Information Processing Capabilities of Random Dendritic Neural Nets

A Dissertation

Presented for the

Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Tzusheng Pei

December 1999

Acknowledgments

I thank my advisor, Dr. Bruce MacLennan, for making this dissertation possible. In the past years I have learned a lot from him: I am now more aware of the limitations of discrete logic, which may be important for thinking about complex systems; I have also seen him working on undefined or very poorly defined fields, and he seems to be comfortable and confident about them. He serves as a great role model for me.

I am grateful to Dr. Jim Drake, Dr. Jim Hall, and Dr. Michael Vose for being on my dissertation committee. In particular, I thank Dr. Drake for his encouragement because he reinforced in me the knowledge that confidence is very important when one gets into the unknown. It has been a great pleasure to work with Dr. Hall, with whom I enjoyed discussing stimulating questions. I received insightful questions from Dr. Vose that were helpful when I got into trouble.

In the past years, Dr. Michael Hines of Yale University has been a great help in making his NEURON software available, and I appreciate his prompt response when I was having problems using it. Also I thank Jacob Hoehn of our lab staff for installing NEURON on our workstations for me, and Vivie Babb for her patient editing of this dissertation.

My deepest thanks go to my wife. Without her patience and her taking care of our children and myself, this work might not have been completed.

Abstract

The goal of this research is to investigate to what degree random artificial dendritic nets can differentiate between temporal patterns after modifying the synaptic weights of certain synapses according to a learning algorithm based on the Fourier transform.

A dendritic net is organized into subnets, which provide impulse responses to a function as a basis for Fourier decomposition of the input pattern. Each subnet is randomly generated. According to the simulations, randomly generated subnets with appropriate parameters are good enough to provide the impulse responses for the Fourier decomposition.

The electrical potential pattern across the membrane of the dendrites follows the cable equation. The simulations use a linear synapse model, which is an approximation to biologically realistic synapses. Both excitatory and inhibitory synapses are present in a dendritic net.

The simulations show that random dendritic nets with a small number of subnets can be modified to differentiate between electrical current patterns to a high degree when the membrane conductance of the dendrites is high, and they also show that the random structures are highly fault-tolerant. The performance of a random dendritic net does not change much after adding or deleting subnets.

Contents

List of Tables

List of Figures

Chapter 1

The Problem to Solve

1.1 The Problem

The objective of this research is to explore how random dendritic nets process information, represented as an electrical current pattern.

An electrical current pattern feeds into a dendritic net, and the network produces electrical potential output as the result of manipulation of the input pattern. In the learning step, ideally the dendritic net learns to recognize the input temporal electrical current pattern after modifying the dendritic net itself, i.e. the dendritic net is tuned to this particular input pattern. The simulations show to what degree artificial random dendritic nets can be modified to differentiate between electrical potential patterns, which represent information.

The simulations are not intended to model exactly any part of the nervous

system. The emphasis is on how the general dendritic structure can process information in a specific way.

1.2 What Are the Open Problems?

It appears that the most highly evolved information processing system is the human brain [Hameroff 87, p.35]. From the computing point of view, the human brain may be considered the most sophisticated computer.

Despite numerous publications of the research on the human brain reporting impressive progress, we still know very little about how the human brain works [Bullock 93, Preface]. Understanding how the brain works is not just accumulating data through experiments and local observations of brain tissues activities, which have been mistaken for understanding how the brain works [Bullock 93, Preface]. In order to understand how the brain works, higher-level theories and modeling about how various cognitive processes work are needed.

We do not know how brains learn, make judgments, or take action, all of which require theories of information processing. It means that we do not know enough about the brain, despite the fact that we have a huge amount of data about it.

Let's move down a little bit. The stomatogastric ganglion (mass of nerve cells) of the spiny lobster normally has 28 neurons. The network serves to "drive the muscles controlling the teeth of the gastric mill so that food can be ground up

for digestion" [Churchland 92, p.4]. The diagram of the circuit reveals impressive details [Churchland 92, p.5]. "What is not understood is how the neural cells interact to constitute a circuit that produces the rhythmic pattern" [Churchland 92, p.5]. Even for this simple network, the global behavior is hard to understand though we know an impressively large number of details about the constituent neurons. But still we don't know if it is detailed enough. This example shows that neuroscience in many areas is data rich, but theory poor [Churchland 92, p.16], even for a small system. This statement doesn't imply that neuroscience is theory poor in all aspects. It depends on the level of abstraction we are looking at. Clearly, at the very highest level, we have artificial neural nets to model massive connections, but they are quite remotely based on the neural system and hence are not realistic. Whether a model is realistic enough or not depends on what abstraction level the investigator looks for.

The central problem is that there is no accurate, comprehensive theory about how a neuron works with other neurons, though we have a large amount of data about neurons. This is an example in neuroscience which looks deceptively simple to us, but it is not trivial to construct a realistic simulation to model it. For this kind of accuracy in the rhythmic patterns, one may need to describe a very detailed pulse generation process in the neurons and to have an accurate synapse model in between neurons. Subneuronal processes, which include possible feedback loops, might also be involved. The whole network is a dynamical system.

1.3 Why Random Dendritic Nets ?

In this research, the goal is not so ambitious as to try to understand how the brain works, i.e. how it processes various information, or even how a neuron works with another neuron in general. Instead, the focus is on exploring the information processing capabilities of dendritic nets, which extend among neurons. This kind of information processing is a local interaction among neurons through spanning dendritic trees.

Axons provide long-range, nonlocal information transfer. For complex information processing, a network needs highly, often randomly, interconnected elements. The ubiquitous, innumerable local interactions, often dendrodendritic nets, provide high connectivity for computation [Pribram 91, p.5]. Local circuit neurons are found in many locations in the central nervous system. Computation is strongly influenced by local interactions that modify the postsynaptic dendritic process [Pribram 91, p.5 & p.10].

If synapses are allowed to grow stronger, weaker or to die out and the dendritic structure can then be arranged in a certain way, then this dendritic net may be able to adapt to the external stimulus and to come out with a slightly nonrandom structure to recognize the stimulus, an electrical current pattern. Another possibility to change a dendritic net structure is for dendrites to grow [Lund 78, p.169].

"The brains of highly developed animals contain more Golgi type II cells than Golgi type I cells" [Dowling 92, p.45]. The main difference between the mouse and the primate cerebral cortex is that there are more neurons with short or with no axon, i.e. Golgi type II neurons, in the primate brain than in the mouse brain [Dowling 92, p.45-46]. It is reasonable to suppose that these neurons with short or with no axon are not isolated, but have interactions with other neurons. Since they have short axons or are without axons, they can be involved in local interaction. This means that there is more local interaction among nearby neurons in the primate brain than in the mouse brain.

Dendrites are not exclusively for receiving input information and axons are not exclusively for providing output [Dowling 92, p.45]. For example, in Golgi type II neurons, which have short or, sometimes, no axons, dendrites and axon terminals are often both pre- and post-synaptic [Dowling 92, p.45]. Some dendrites and axons share the ability to transmit electrical signals, and in many neurons both information input and output can be on the same set of dendrite-like fine processes [Levitan 97, p.9].

Neurons with local interaction can be found in many locations of the sensory and central nervous system. They often form dendrodendritic connections [Pribram 91, p.5]. The reason may be that a considerable amount of important information processing happens in the subtle local neuronal interactions. In this study, the simulations concern dendrodendritic information processing.

This work may be considered a step toward understanding the working of the brain, which apparently shows superior intelligence not attainable by the current (digital) computer hardware and software systems. The information processing in dendrodendritic connections may be viewed as a computing model. Some particular distinctions of dendrodendritic processing include being parallel and analog.

There are many dendritic spines on dendrites for receiving analog information across synapses from other dendrites. The tree structure with synapses can do merging (convergence), branching (divergence), and other modifications on the information.

1.4   What's New About the Work in Computer Science ?

The main reason to explore the information processing capabilities of dendritic nets is that they may play an important role in the brain, particularly the human brain, which can be seen as the most sophisticated computing device. Note that even an insect's brain, which guides the highly intelligent and coordinated behavior of the insect, cannot be simulated satisfactorily and efficiently by any current hardware and software system on the digital computer. Essentially we want to see how dendritic nets can possibly work in the brain. We have tried to run realistic simulations as far as it was practical to do so.

Though artificial neural nets are far different from the biological networks

composed of real neurons, yet they capture one of the most important aspects of neurons: connections, which are usually complicated. A more realistic dendritic net, which also has complicated connections, can manipulate patterns in some ways that cannot be achieved in artificial neural networks. The following describes properties of dendritic nets and compares the more realistic dendritic net with the artificial neural network.

The common property of dendritic nets and artificial neural network is that they all have complicated connections, branching, and merging. These are some of the mechanisms to manipulate the input pattern.

A segment of the membrane of a dendrite can be described as a resistor and a capacitor in parallel. Normally an electrical potential pattern across the membrane propagating along a dendrite becomes smaller in amplitude along the dendrite, a phenomenon called passive spread [Levitan 97, p.54], due to electrical resistance across the membrane and along its axis. Longitudinal gradients of electrical potential dissipate with time.

A dendrite can change the shape of an input electrical potential pattern due to the capacitor across the membrane. The voltage that builds up across the circuit of resistor and capacitor rises and falls more slowly than the current impulse across the membrane [Dowling 92, p.429]. It is a way of manipulating the input pattern by changing its shape. The effect of capacitors and resistors on electrical patterns are not shown in artificial neural networks. Given any plausible combination

of G_m, membrane conductance, R_a, longitudinal resistance, and C_m, membrane conductance, there can be no sharp edges in any electrical potential pattern in a dendritic net. In particular, the smoothing effects of a dendrite suppress the high frequency components of the input electrical current pattern. This turns a segment of dendrite into a low pass filter, which loses part of the information of a pattern propagating along it.

We have stayed as close to real neurons as practically possible. At least our simulations have shown how well this kind of structure can process information with our learning algorithm.

In the multi-layer feed-forward artificial neural networks, there is no time delay of information propagation on two different paths from the input node to the output node. The flow of information from one layer to the next is synchronized. This kind of network does not need delay for its operation. The input is not temporal. The input to represent a complicated pattern, such as a two dimensional grey-scale character composed of pixels, is usually a discrete encoding of the pattern. The encoding is done off-line, not in the neural network.

Recurrent artificial neural networks use feedback, and layer-to-layer information passing is synchronized. They are also capable of processing temporal input patterns in which each pattern is represented as a time series of data.

There are two sources of time delay in a dendritic net. The major source is the time for an electrical pattern to get across a synapse. A minor source is the

8

time needed for the electrical potential pattern to travel along the dendrite. In our simulations, the delay of a synapse is randomly generated between [0, 2 msec] and the branch lengths of a dendritic tree are less than 60 microns.

In this study, a dendritic net is an analog system simulated by numerical computation. The input to a dendritic net is a temporal pattern. The information flow in a dendritic net is not synchronized. The times spent for information passing over two different paths from the input to any point in the network are different, depending on the delays of the synapses on the paths, the electrical properties of the dendrites in the two paths, and the length difference of the paths. With the unsynchronized mergings of patterns from many different paths, dendritic nets offer a very complicated manipulation of patterns.

An important issue was to explore learning by means of the modifications in synapses of the dendritic net in order to recognize an input pattern or to differentiate between patterns. The modification is local and simple for each synapse. The local modifications can lead to a global emergent property [MacLennan 92]. In our context, the local modifications will make the whole dendritic net capable of recognizing a temporal pattern or of differentiating between temporal patterns in an input set (defined in Section 4.2), which is an emergent global capability of the dendritic net.

One distinct structure of real dendritic nets is their seemingly random structure. Randomness or near-randomness is almost everywhere in nature. It is

significant if a near-random structure can self-modify itself locally into a more nonrandom structure in some way to recognize or differentiate between patterns. The former random structure seems to be unbiased and does not process information for any purpose, but the latter nonrandom structures can recognize a pattern or differentiate between the patterns in an input set. Our simulations show how well this can be done by a dendritic net.

In computer science, one of the goals is to build fast, fault-tolerant computers. As the simulations show, random dendritic nets are very fault-tolerant. This work may provide a basis for future massively parallel analog hardware with high fault-tolerance.

1.5   Related Work

As a result of advances in experimental techniques in the past two decades, a large amount of data about neurons has accumulated and has served as the empirical basis for the biologically realistic models [Hines 97]. The following list taken from [Hines 97] shows what has been modeled: the cellular mechanisms that generate and regulate chemical and electrical signals, drug effects on neuronal function, presynaptic and postsynaptic mechanisms underlying communication between neurons, integration of synaptic inputs, action potential initiation and conduction, cellular mechanisms of learning, cellular oscillations, thalamic net-

works of certain configurations to study spindle rhythmic oscillations [Destexhe 97], and neural information encoding.

[Bazhenov 98] studies the dynamical properties of synaptically coupled neurons and neurons coupled in lattice and chains. A coupled chain of neurons with reciprocal inhibition between neighboring neurons exhibited synchronous oscillations.

In [Bush 95], one column of layer V of the visual cortex using 80 pyramidal cells and 20 basket cells is simulated to study how networks of cortical neurons are organized to allow long-range contextual inputs to influence local processing.

According to the chapter called "Computational Overview" of [Churchland 92, p.61-p.140], higher-order models are what we generally find in books on artificial neural networks. The abstraction is at a high level, far from real neural networks. They do not model details about how two neurons interact. Our simulations are situated between realistic dendrodendritic nets and high level artificial neural networks.

A model for higher-level information processing in dendritic nets, one kind of interaction among neurons, has been formulated in [MacLennan 92]. It is a theoretical framework based on the generalized Fourier transform. It does not specify how detailed the implementation has to be. My simulations are of this kind.

Simulating large biological neural nets demands high computing power [Lytton

96] provides a method to reduce simulation time; its purpose is not to simulate any part of the nervous system. In [Lytton 96], a single lumped state-variable is used to represent a large number of converging synaptic inputs, because all of the synapses of a single type are doing identical, potentially redundant calculations at slightly different times. The Markov synapse model [Destexhe 94a, Destexhe 94b, Destexhe 95a, Destexhe 95b, Destexhe 95c, Destexhe 96], which preserves some major aspects of biophysically realistic synaptic processes, is used for simulations. The simulations were run on mutually-excitatory neurons receiving similar input spiking. Each individual synapse is not represented as a distinct entity. This arrangement maintains a single queue of spike arrival times instead of N queues for N synapses. "Typically, the time of a presynaptic activation is added to the appropriate synaptic delay and then stored on a queue." "When the calculated time is reached in the simulation, the item is removed from the queue and the postsynaptic element is activated." This approach greatly reduces the computing time spent on simulating neural networks composed of large numbers of neurons and synapses.

Note that Lytton's method is for synapses with similar spike pulse trains at the presynaptic side. This method cannot be applied to my simulations where the electrical potential travels along passive dendrites, i.e. conductors with electrical resistance and capacitance, and across synapses where complicated configurations, i.e. branching, merging, superposition, integration, and delay, happen along den-

drites and across synapses. The electrical potential patterns coming from different paths leading to the synapses will have different shapes, therefore, the synapses in my simulations will not have similar input spiking.

In my simulations, the patterns are not identical spike trains. Instead, the patterns are manipulated by the dendritic net and hence have an irregular shape. In this work, I am interested in the results of the manipulation of patterns by the network, i.e. in changing pattern shapes and amplitudes in dendritic structures of branching, superposition, and integration of temporal electrical potential patterns. The simulations show to what extent artificial random dendritic nets can be modified to differentiate between electrical potential patterns, which represent information. The emphasis is on how the dendritic structure can process information.

Chapter 2

Modeling Dendrites and Synapses

2.1 The NEURON Simulation Environment

All simulations were run under NEURON, a simulation language interpreter and environment for simulating neural configurations and activities developed by Michael Hines of Yale University and John W. Moore of Duke University Medical Centers.

Neuron processes, such as synapse models, can be compiled and incorporated into the NEURON simulation environment by user modeling programs to specify the activities of the neuron processes. NEURON provides an environment for simulating dendrites constructed by passive cables, axons with the active Hodgkin-Huxley property, and neuronal processes. The user writes modeling programs in the language provided in NEURON and runs them in the NEURON simulation environment.

For our simulations, a program in C/C^{++} was written to take a data file which specifies a dendritic net as the input and to generate a NEURON modeling program to be interpreted and run under the NEURON simulation environment.

2.2 Dendrites and Axons

Any segment of a dendrite is considered a cylinder whose electrical potential across the membrane satisfies the partial differential equation of linear cable theory, which is based on the dendrite's equivalent circuit.

Dendrites are assumed to be passive, which means that conductance is fixed. However, the membrane conductance depends on the electrical response: for a strong response, the conductance starts to change. Therefore, the cable theory is valid only when the response is about half the strength required for generating an action potential (spike) [Tuckwell 88a]. For simulations on NEURON [Hines 84][Hines 89][Hines 97], a parameter is used to specify whether a cable segment is passive or active.

An axon, which is an active cable (i.e. obeys the Hodgkin-Huxley equation), will preserve the shape of electrical spikes propagating along it. Whenever the electrical potential difference across the membrane is higher than a certain threshold, an action potential spike will be generated. The action potential will keep the shape of the spike along the axon and transmit information to another area of the

15

neural system.

2.3 Synapse Models

The electrical properties of all cells, including nerve cells, depend on the movement of small inorganic ions across the cell membrane [Levitan 97]. Ion channels are most important for nerve cells. Ions flow from one side of the membrane to the other side through ion channels, which are proteins spanning the membrane. Energy-driven pumps or carriers are proteins that carry an ion from one side of the cell membrane to the other side. For nerve cells, energy-driven pumps or carriers play a supporting role.

There are two kinds of synapses, electrical and chemical. Historically, we happen to know more about chemical synapses than electrical synapses. Both kinds of synapses are important. In the following, we restrict our discussion to chemical synapses, where synaptic currents are mediated by ion channels activated by neurotransmitters released from the presynaptic terminals.

The synaptic process of chemical synapses consists of a sequence of stages with associated quantities, i.e. presynaptic depolarization (electrical pulse), the amount of neurotransmitters released, the number of postsynaptic ion channels opened, and the generation of postsynaptic current and electrical potential. The quantity of each stage is related to the quantity of the previous stage. The relation

is approximately linear except for saturation and depletion of neurotransmitters.

We do not intend to discuss all existing synapse models. In the following sections, Markov and linear synapse models are described. Both models can be implemented and compiled into the NEURON simulation environment, which is then able to interpret synapses defined in simulation programs.

2.3.1 Synapse Model Based on the Markov Process

The Markov synapse model described in this section was proposed and implemented by Destexhe and his coworkers [Destexhe 94a, Destexhe 94c, Destexhe 95a, Destexhe 95b, Destexhe 95c, Destexhe 96]. The implementation can be incorporated into the NEURON environment. The following is a brief summary of this synapse model:

(a) Ion channels are proteins in different conformation states.

(b) The transition from one conformation state to another is a Markov process. The transition probability can be obtained by single-channel recording.

(c) If the concentration gradient is not extreme, the current through a synapse is proportional to the fraction of open channels.

(d) This model is accurate for describing synaptic transmission even with 2-3 states.

The Markov synapse model has been used by Destexhe and associates to demonstrate the input-output behavior on a single synapse.

Practically, it is difficult to simulate large dendritic nets with many (sixty to several hundred) synapses based on the Markov model when each synapse has its own parameters and each synapse receives a different electrical potential pattern at the presynaptic site. The problem lies in the difficulty of getting the parameters and input pattern to work in a controlled manner. In fact, it is a good model, which describes the synapse process realistically. For our purpose, we need to use a simplified model to make large dendritic net simulations feasible under the control of the person who does the simulation.

2.3.2   Linear Synapse Model

The following describes the linear synapse model, which captures the behavior of a synapse operating in the linear range without saturation or depletion of neurotransmitters. The implementation of the linear synapse model under NEURON is in Appendix A.

Let the resting potential be r, the presynaptic electrical potential at time t be pre_t, the postsynaptic potential at time t be $post_t$, and the multiplier (the synaptic weight) be m. A linear synapse is defined such that it satisfies the relation $post_{t+1} = r + m(pre_t - r)$. With reference to the resting potential, the value of the presynaptic electrical potential above or below the resting potential is multiplied by a multiplier (synaptic weight), which serves to amplify the input, to obtain the postsynaptic potential above or below the resting potential at the

18

next time slice.

The above formulation is for synapses without delay. For a synapse with delay, a first-in-first-out buffer is implemented to hold the presynaptic electrical potentials for a sequence of simulation time slices. At each time slice, at the end of the buffer an electrical potential value comes out to be the postsynaptic electrical potential, then the buffer is shifted one position toward the postsynaptic end, and then the presynaptic electrical potential feeds into the front end of the buffer. Note that it takes memory space in the computer to hold the buffer. The finer the time slices for numerical simulations, the larger the memory space needed.

This simple mechanism is intended to maintain the linear relationship between the presynaptic and postsynaptic electrical potentials relative to the resting potential. It is a reasonable approximation of a chemical synapse when conditions are not extreme. There is only one parameter m, the synaptic weight, to set for a linear synapse. This model is easy to work with for simulating large scale artificial dendritic nets when one's purpose is to use the linear range of the synapse.

2.3.3 Choosing A Synapse Model for Simulation

There are a number of factors involved in choosing a synapse model: biological realism, ease in controlling the simulation (simplicity), and computational feasibility.

Our purpose is to show how the dendritic structure can process information,

which is represented as an electrical current pattern. A dendritic net has to be able to differentiate between input patterns. Consequently we are interested in the linear range of the operation of a synapse, but not in the saturation range, where the output of the synapse is about the same regardless what pattern amplitude is currently feeding to the synapse from the presynaptic side. For this reason, we chose to use the simplified linear synapse model.

The Markov model is comprehensive and biologically realistic. It tries to model exactly how the biological chemical synapse operates, but for our simulations, it is difficult to use the Markov synapse model due to the large number of parameters involved, and providing different parameters for each Markov synapse in a large dendritic net is another difficult issue. Many times, adjustments are made to bring all synapses operating approximately in the linear range. This means that, for our purpose, the results should be about the same using either the Markov model or the linear model. In addition, it requires extra work to bring the Markov synapses in the linear range for each run. Also, higher computing power is required to simulate Markov synapses. This doesn't imply that the Markov model should not be used in simulating large dendritic nets. If one needs to look into the detailed, non-linear operation of synapses in a large dendritic net, the Markov model seems to be the only choice.

We chose the linear model for simulation because the linear synapse model implements the properties of the synapse we need and is easiest to use and takes

least time to simulate on a computer. In addition, the linear model fits exactly into the linear formulation of the learning algorithm for adjusting synaptic weights (strengths) to recognize or differentiate between input patterns.

Chapter 3

Simulating the Dendritic Net and the Learning Algorithm

3.1 Constructing the Dendritic Net

In our simulations, an artificial dendritic net is composed of a number of subnets, an input tree to feed into the subnets from the input side, and an output (summing) tree to combine the subnets from the output side after the synapses. Each subnet consists of a pair of random binary trees. There are synapses at the contact points of the two binary trees. The linear synapse model is used for all synapses in the dendritic structures in the simulations. Figure 3.1 shows the dendritic net structure. See Figure 3.2 for typical subnets. In each subnet, the two tree roots are the extremes at the top and the bottom. All tree branches are segments of
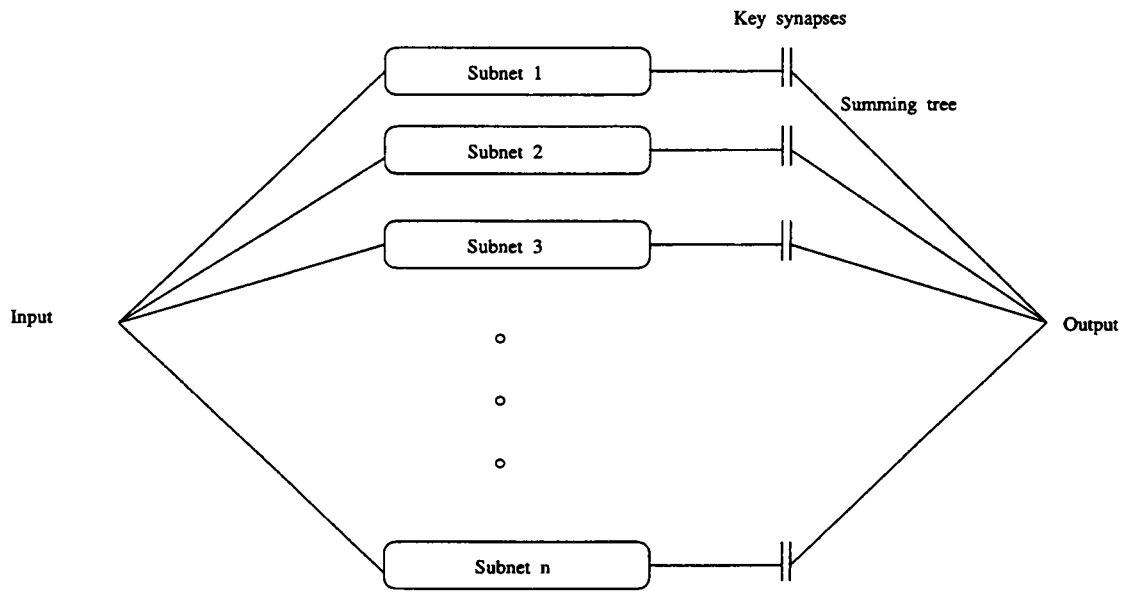
Figure 3.1: Dendritic net structure
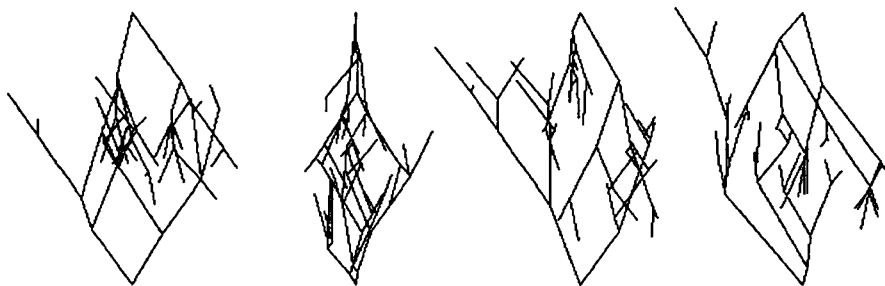
Figure 3.2: Subnets of a dendritic net

dendrites. Note that, for a dendritic net with n subnets, there are $(1 + n + 1)$ dendritic trees in Figure 3.1. The whole structure is to implement the components for the generalized Fourier transform. The subnets are grown randomly, and the summing tree is for convergent connections. The values of the parameters of the dendritic nets are shown in Table C.1 of Appendix C.

For simplicity, all the structures are on a two-dimensional plane. The main operations of a dendritic net on the input pattern include branching, merging, superposition of overlapping electrical potential patterns, and then integration in the summing tree. Two-dimensional modeling shows all of these operations. Hence, it suffices for our purpose to have all structures on the same plane.

In simulations, each branch coming out of the tree root or out of a tree branching point until the next branching point or until the end of the branch is called a cable, with the parameters length, diameter, axial resistivity, membrane conductance, and membrane capacitance. For the purpose of numerical simulations, a cable is further divided into a sequence of one or more compartments. Each compartment has a corresponding equivalent circuit.

In constructing a random binary tree, the angle between two branches, the length of each branch, and the delay of a synapse are generated uniformly, randomly in their respective ranges. All the dendrite segments have the same diameter. There is no interaction between any two subnets. If the information processing being implemented doesn't require high precision, some weak interaction through
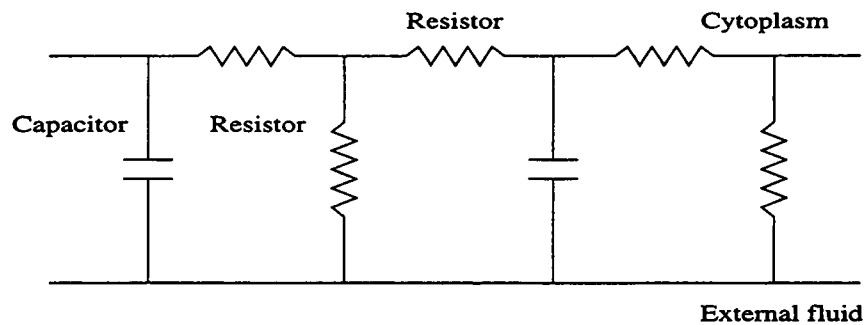
Figure 3.3: Equivalent circuit of membrane

synapses among subnets may be tolerable.

What our simulations do is to change the synaptic weights (strengths) of the key synapses connecting to the summing tree to enable the net to recognize or differentiate among the input patterns.

3.2 Modeling Dendrites with Cable Theory

The equivalent circuit of a neural membrane and the cable equation for describing a segment of dendrite can be found in [Hines 84, Hodgkin 52, Koch 89, MacGregor 93]. The following is a description of the equivalent circuit for a neural membrane and cable from [Hodgkin 52] and the formulation of cable theory from [Hines 84]. An equivalent circuit is shown in Figure 3.3. The lipid bilayer of the membrane determines the membrane capacitance, and the membrane protein for ion channels regulates the resistance of the membrane [Dowling 92, p.66]. Compared with the internal axial resistance, the external resistance can be ignored, because of its greater fluid volume [Kandel 81, p.48].

25

Current can be carried through the membrane by charging the membrane capacity and by the transmembrane currents. In the nervous system, there are four ion species involved: sodium ions (Na$^+$), potassium ions (K$^+$), calcium ions (Ca^{++}), and chloride ions (Cl$^-$) [Johnston 95, p.3]. There are two effects involved in determining each ion current: an electrical potential difference and a concentration gradient

A one-dimensional cable can be described by the following equation [Hines 84] based on the equivalent circuit:

$$\frac{1}{2\pi a}\frac{\partial}{\partial x}\left(\frac{\pi a^2}{R_a}\frac{\partial V}{\partial x}\right) = C_m\frac{\partial V}{\partial t} + I_{HH},$$

where a is the radius which is a function of the longitudinal coordinate x along the cable, V is the electrical potential across the membrane (mV), C_m is the membrane capacitance ($\mu F/cm^2$), R_a is the axial resistivity (Ωcm), and I_{HH} is the membrane current (amps/cm^2).

For numerical simulation in the NEURON environment, each cable branch itself can be a compartment or can be further divided into a sequence of compartments along the segment. The parameter nseg specifies the number of compartments. Each compartment has the parameters L (micron), the length of the compartment, D (micron), its diameter, e_pas (mV), the resting potential, g_pas (S/cm^2), the membrane conductance, r$_a$, axial resistance (Ω-cm), and cm

$(\mu F/\mathrm{cm}^2)$, the membrane capacitance. All the parameters are specified in the simulation program to run in the NEURON environment.

An efficient algorithm for solving the dynamic cable equation [Hines 84, Hines 89] on NEURON takes advantage of the acyclic tree structure of dendrites. It greatly reduces computing time compared with considering the whole structure as a general graph. A simulation program on NEURON specifies the dendritic properties and how the compartments are connected to form the dendritic tree. The user focuses on the structure and properties of the dendritic net. The NEURON environment interprets the simulation program and takes care of the numerical computation.

3.3 Linear System Theory

A linear system can be used as the framework to describe the electrochemical dynamics of the dendritic net, which is approximately linear within a certain range of its variables. The following briefly introduces basic linear system theory, on which the learning algorithm is based. More detailed descriptions of linear system theory can be found in [Faulkner 69, DeCarlo 89, Rugh 96]. For our purpose, we restrict the treatment of linear system theory to the real numbers **R** instead of to the complex numbers **C** .

Definition The unit step function $u(t)$ is defined in the domain of real numbers

except for zero, i.e. $t \in \mathbf{R} - \{0\}$. Its range is a subset of the real numbers $\mathbf{R}$ as follows:

$u(t) = 0$ for $t < 0$, $u(t) = 1$ for $t > 0$, and $u(t)$ undefined for $t = 0$. ♣

Definition The unit impulse function $\delta(t)$, $t \in R$, is the derivative of the unit step function $u(t)$ with respect to t as follows:

$\delta(t) = 0$ for $t \neq 0$ and $\int_{-\infty}^{+\infty} \delta(t)dt = 1$. ♣

The above unit impulse is an abstraction of a physical impulse of very short duration, with the area under the impulse being 1. It serves as a standard impulse for the system under study. The variable t is for time.

Suppose that we have a physical system with an electrical current pattern as its input. We inject a unit current impulse in a very brief time period and watch the output of the system. This output (impulse response) characterizes the system's response to an arbitrary input pattern. The following definition abstracts this input-output relation.

Definition The impulse response of a linear system S characterized by the linear operator ϕ, defined in the real numbers $\mathbf{R}$, is the response (output) of the system with input $x\delta(t')$, where x is a positive real amplitude and t' is a real-valued time variable. Then the normalized impulse response $h(t)$ is

$h(t) = \frac{\phi[x\delta(t')]}{x} = \phi[\delta(t')].$

In general, for the input impulse centered at $t' = \tau$, the normalized impulse response is

$$h(t - \tau) = \frac{\phi[x\delta(t' - \tau)]}{x} = \phi[\delta(t' - \tau)].$$

In the above equation, the variable t in h and t' in ϕ both refer to time, but they vary independently. The linear operator ϕ, which is not a function, generates a function of t for the impulse response. ♣

Once we know the normalized impulse response $h(t - \tau)$ of the system S with the input impulse $x\delta(t' - \tau)$, we can predict the response of the system S to an arbitrary input function. This means that the linear operator ϕ of the system S is completely characterized by its impulse response $h(t - \tau)$. Note that the impulse response (output) of a system can be measured for a physical or simulated system by injecting an approximation of the (input) impulse $x\delta(t' - \tau)$ into the system at the input site at time τ.

A dendritic net can work on both electrical current and electrical potential impulses. Figure 3.4 shows a typical electrical potential impulse (at the input site) generated by injecting a current square of short duration at the input site. For a dendritic net working on current input patterns, the current square is considered the impulse, and for a dendritic net working on electrical potential input patterns, the electrical potential pattern as shown in Figure 3.4 is the impulse. Figure 3.5 shows the impulse response of a dendritic net's subnet to an impulse in Figure 3.4 or to the square current impulse of short duration.

Definition The convolution $h(x) \star \phi(x)$ of real functions $h(x)$ and $\phi(x)$, $x \in \mathbf{R}$ and $h(x) = \phi(x) = 0$ for $x < 0$, is the following integral:
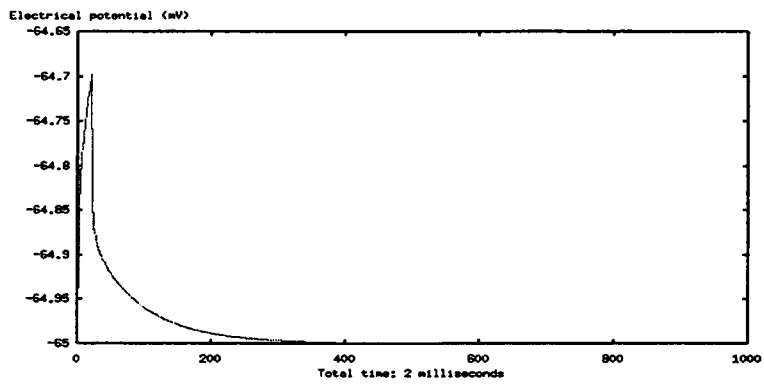
Figure 3.4: An impulse to a dendritic net
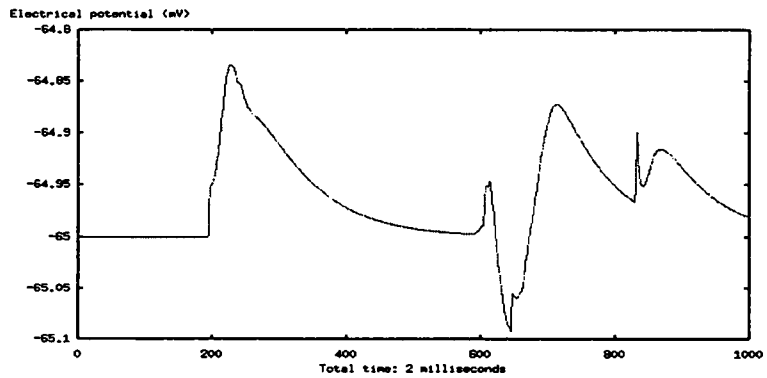
Figure 3.5: The response of a subnet to an impulse

30

$$h(x) \star \phi(x) = \int_0^\infty h(x-t)\phi(t)dt. \clubsuit$$

In a causal linear system, all relevant functions are zero for $x < 0$. Therefore, the above definition limits the convolution to nonnegative x. The following theorem shows what response a system will have with an arbitrary input function. It can be proved by dividing the input pattern into a succession of an infinite number of rectangular pulses and then applying the linear operator to these rectangular pulses. A sketch of the proof is on p.43.

Theorem The response (output) $y(t)$ of a linear system S, characterized by the linear operator ϕ with the impulse response $h(t)$ for $t \geq 0$, to an arbitrary input function $x(t)$, $t \in [0, T] \subset \mathbf{R}$, is the convolution of the input function $x(t)$ and the impulse response $h(t)$ of the system S, i.e. $y(t) = x(t) \star h(t)$. $\clubsuit$

Note that in the above formulation there is one input site and one output site. It can be easily extended to an input space and output space which are subsets of an n-dimensional Euclidean space ($\mathbf{R}^n$ for $n \in \mathbf{I}^{>0}$). For each pair of input site from the input space (a subset of $\mathbf{R}^n$) and output site from the output space (a subset of $\mathbf{R}^n$), there is an impulse response associated with them.

3.4 Solving the Equivalent Circuit of the Dendrite

3.4.1 Formulation and Solution in the Time Domain

Definition Let the real function $f(t)$ be piecewise continuous and

$\mid \int_0^\infty f(t)e^{-\sigma t}dt \mid < \infty$ for some $\sigma \in \mathbf{R}$. Then the Laplace transform [Mohanty 87]

of $f(t)$ is

$$F(s) = L(f(t)) = \int_{-\infty}^\infty f(t)e^{-st}dt,$$

where $s \in \mathbf{C}$ and the real part of s is less than σ.

The inverse Laplace transform [Mohanty 87] is defined as

$$f(t) = L^{-1}(F(s)) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} F(s)e^{st}ds,$$

where $j = \sqrt{-1}$, $c > \sigma$, and $s = \sigma + j\omega$ for $\omega \in \mathbf{R}$. ♣

Practically, a physical system runs in $t \geq 0$, and the system is determined by

the input and the initial condition in $t \geq 0$. The Laplace transform can then be

restricted to $t \geq 0$ as follows:

$$F(s) = L(f(t)) = \int_0^\infty f(t)e^{-st}dt.$$

Consider the electrical circuit of a small dendrite segment as shown in Figure

3.6. Its Laplace transform equivalence is shown in Figure 3.7.

By summing current at V_1 in Figure 3.7, we have the following equation:
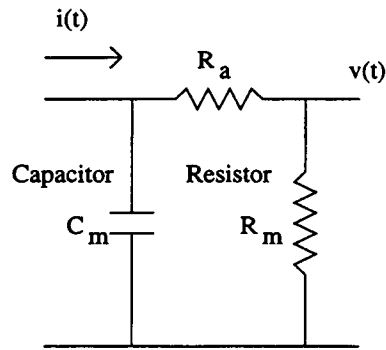
Figure 3.6: Equivalent circuit in the time domain for a small dendrite segment
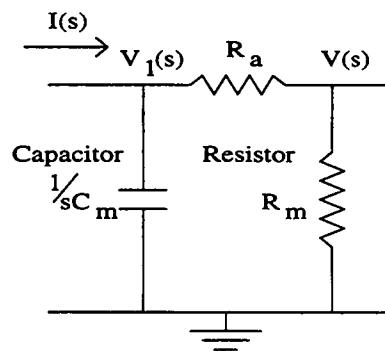
Figure 3.7: Equivalent circuit in the Laplace transform domain for the small dendrite segment

$$-I(s) + sC_m V_1(s) + \frac{V_1(s) - V(s)}{R_a} = 0.$$

The same can be done for the point V in Figure 3.7. Then we have the following equation:

$$\frac{V(s) - V_1(s)}{R_a} + \frac{V(s)}{R_m} = 0.$$

The relation between $I(s)$ and $V(s)$ can be established from the two equations.

Theorem Let a small segment of dendrite be described by the following two simultaneous equations:

$$\begin{cases} -I(s) + sC_m V_1(s) + \frac{V_1(s) - V(s)}{R_a} = 0, \\ \frac{V(s) - V_1(s)}{R_a} + \frac{V(s)}{R_m} = 0, \end{cases}$$

where $C_m, R_a, R_m \in \mathbf{R}^{\geq 0}$.

Then $V(s) = \frac{I(s)}{sC(1 + \frac{R_a}{R_m}) + \frac{1}{R_m}}$.

Proof

By eliminating $V_1(s)$ in the two equations, we get the expression for $V(s)$ in terms of $I(s)$. ♣

The following defines the transfer function of a system.

Definition The transfer function $H(s)$ of a system is the coefficient of the

relation between the input $I(s)$ and the output $O(s)$ in the Laplace transform domain, i.e. $O(s) = H(s)I(s)$. ♣

The following theorem links the transfer function and the impulse response of a system.

Theorem Let the transfer function of a system be $H(s)$, and let the impulse response of the system be $h(t)$. Then $h(t)$ is the inverse Laplace transform of $H(s)$, i.e.

$$h(t) = L^{-1}H(s).$$

Proof

Since the impulse response is the output of the system when the input is the Dirac-delta function, we have

$$
\begin{aligned}
I(s) &= L(\delta(t)) \\
&= \int_{-\infty}^{\infty} \delta(t)e^{-st}dt \\
&= e^{-st}\,|_{t=0} \\
&= 1.
\end{aligned}
$$

Then

$$O(s) = H(s)I(s) = H(s) \cdot 1 = H(s).$$

By taking the inverse Laplace transform of the above equation, we have

$$o(t) = L^{-1}(O(s)) = L^{-1}(H(s)).$$

$o(t)$ is the impulse response of the system when the input is $\delta(t)$, i.e.

$$h(t) = L^{-1}(H(s)).\clubsuit$$

Definition The transfer function $H(s)$ of a small dendrite segment is the coefficient of the relation between the input $I(s)$ and the output $V(s)$ in the Laplace transform domain, i.e. $V(s) = H(s)I(s)$. $\clubsuit$

Corollary The transfer function of the small dendrite segment is

$$H(s) = \frac{V(s)}{I(s)} = \frac{1}{sC_m(1 + \frac{R_a}{R_m}) + \frac{1}{R_m}} = \frac{R_m}{sC_m(R_m + R_a) + 1}.\clubsuit$$

The equation to relate input $I(s)$ and output $V(s)$ in the Laplace transform domain can be transformed back to the time domain when $I(s)$ is known. Then the relation between the input current $i(t)$ and the output voltage $v(t)$ in the time domain can be established.

In the simulations, an input pattern is a sequence of current steps clamped

36

to a dendrite. The following gives a detailed definition for an input pattern to a small dendrite segment.

Definition An input pattern $P(t)$ is a piece-wise continuous function defined in a time interval $[0, T]$, where $T \in \mathbf{R}^{>0}$. Let $t_0, t_1, t_2, ..., t_n$ be a monotonely increasing sequence of numbers in $\mathbf{R}$ such that $t_0 = 0$ and $t_n = T$. For each subinterval $[t_{i-1}, t_i)$, $i \neq n$, there is an amplitude $I_i \in R$ associated with this interval such that $P(t) = I_i$ if $t \in [t_{i-1}, t_i)$. For $t = T$, $P(T) = P(t_n) = I_n$. ♣

In our simulations, $n = 50$, $[0, T] = [0, 2 \text{ msec}]$, and the length of each time subinterval is a constant ($\frac{2}{50}$msec). The input site of a small dendrite segment is at one end of the segment, and the output site is at the other end of the segment.

Since the equation $V(s) = H(s)I(s)$, the Laplace transform, and the inverse Laplace transform are all linear, each step among n steps in the input pattern can be considered a separate input to one end of the small dendrite segment. We need to calculate only one step in the input pattern as the input. An example of input pattern is shown on p.78.

Definition A square input $s_i(t, I_i)$ of current amplitude I_i, $i \in \{1, 2, 3, ..., n \mid n \in \mathbf{I}^{>0}\}$ and $I_i \in \mathbf{R}$, in the time interval $[t_{i-1}, t_i)$, $t_i > t_{i-1} \geq 0$, is defined as follows:

$$s_i(t, I_i) = \begin{cases} I_i, & t \in [t_{i-1}, t_i) \text{ for } i \neq n \text{ or } t \in [t_{i-1}, t_i] \text{ for } i = n, \\ 0, & \text{otherwise. ♣} \end{cases}$$

37

One way to solve a circuit equation is to transform the input pattern in the time domain to the Laplace transform domain, then multiply the input in the Laplace transform domain by the transfer function to get the output in the Laplace transform domain, and then transform the whole equation back to the time domain to get the output in the time domain.

The Laplace transform of the square input $s_i(t, I_i)$ is

$$
\begin{aligned}
L[s_i(t, I_i)] &= \int_0^{+\infty} s_i(t, I_i) e^{-st} dt \\
&= \int_{t_{i-1}}^{t_i} I_i e^{-st} dt \\
&= \frac{I_i e^{-st}}{-s} \Big|_{t_{i-1}}^{t_i} \\
&= \frac{I_i}{s} (e^{-st_{i-1}} - e^{-st_i}).
\end{aligned}
$$

Lemma The Laplace transform of the square input $s_i(t, I_i)$ is $\frac{I_i}{s}(e^{-st_{i-1}} - e^{-st_i})$.♣

Theorem Let the current input to one end of a small dendrite segment be $s_i(t, I_i)$ and let the time constant τ be $C_m(R_m + R_a)$. Then the output electrical potential at the other end of the dendrite segment is

$$
v_i(t, I_i) = \begin{cases}
0, & \text{if } t < t_{i-1}, \\
I_i R_m (1 - e^{-\frac{t-t_{i-1}}{\tau}}), & \text{if } t \in [t_{i-1}, t_i), \\
I_i R_m e^{-\frac{t-t_i}{\tau}} (1 - e^{-\frac{t_i-t_{i-1}}{\tau}}), & \text{if } t \geq t_i.
\end{cases}
$$

38

Proof

The output $V_i(s, I_i)$ with the unit block $s_i(t, I_i)$ as the input in the Laplace transform domain is as follows:

$$
\begin{aligned}
V_i(s, I_i) &= \frac{1}{[sC_m(1 + \frac{R_a}{R_m}) + \frac{1}{R_m}]} \frac{I_i}{s}[e^{-st_{i-1}} - e^{-st_i}] \\[2mm]
&= \frac{I_i}{C_m(1 + \frac{R_a}{R_m})[s + \frac{\frac{1}{R_m}}{C_m(1+\frac{R_a}{R_m})}]s}(e^{-st_{i-1}} - e^{-st_i}) \\[2mm]
&= \frac{I_i}{C_m(1 + \frac{R_a}{R_m})} \frac{1}{s[s + \frac{1}{C_m(R_m+R_a)}]}(e^{-st_{i-1}} - e^{-st_i}) \\[2mm]
&= \frac{I_i}{C_m(\frac{R_m+R_a}{R_m})} \frac{1}{s[s + \frac{1}{C_m(R_m+R_a)}]}(e^{-st_{i-1}} - e^{-st_i}) \\[2mm]
&= \frac{I_i R_m}{C_m(R_m + R_a)}C_m(R_m + R_a)(\frac{1}{s} - \frac{1}{s + \frac{1}{C_m(R_m+R_a)}})(e^{-st_{i-1}} - e^{-st_i}). \\[2mm]
&= I_i R_m(\frac{1}{s} - \frac{1}{s + \frac{1}{C_m(R_m+R_a)}})(e^{-st_{i-1}} - e^{-st_i}) \\[2mm]
&= I_i R_m(\frac{e^{-st_{i-1}}}{s} - \frac{e^{-st_i}}{s} - \frac{e^{-st_{i-1}}}{s + \frac{1}{C_m(R_m+R_a)}} + \frac{e^{-st_i}}{s + \frac{1}{C_m(R_m+R_a)}}).
\end{aligned}
$$

Note that $L[u(t - t_i)] = \frac{e^{-st_i}}{s}$ and $L[u(t - t_i)e^{-a(t-t_i)}] = \frac{e^{-st_i}}{s+a}$, where the unit step function is defined as follows:

$$
u(t - t_i) = \begin{cases} 0, & t < t_i, \\ 1, & t \geq t_i. \end{cases}
$$

The two Laplace transforms can be proved by substituting the functions $u(t - t_i)$ and $u(t - t_i)e^{-a(t-t_i)}$ in the definition of the Laplace transform.

By taking the inverse Laplace transform on the equation for $V_i(s, I_i)$, we have

the electrical potential output in the time domain:

$$v_i(t, I_i) = I_i R_m \{ [u(t - t_{i-1}) - u(t - t_i)] - u(t - t_{i-1}) e^{-\frac{t - t_{i-1}}{C_m(R_m + R_a)}} $$

$$+ u(t - t_i) e^{-\frac{t - t_i}{C_m(R_m + R_a)}} \}.$$

Since the time constant $\tau = C_m(R_m + R_a)$, the above is equivalent to the following expression.

$$v_i(t, I_i) = \begin{cases} 0, & \text{if } t < t_{i-1}, \\ I_i R_m (1 - e^{-\frac{t - t_{i-1}}{\tau}}), & \text{if } t \in [t_{i-1}, t_i), \\ I_i R_m e^{-\frac{t - t_i}{\tau}} (1 - e^{-\frac{t_i - t_{i-1}}{\tau}}), & \text{if } t \geq t_i. \clubsuit \end{cases}$$

Observe that, in the time interval $[t_{i-1}, t_i)$ while the input $s_i(t, I_i)$ is applied, the output $v_i(t, I_i)$ does not respond right away. Instead, it increases but gradually slows down on the way up from 0 until it reaches $R_m(1 - e^{-\frac{t_i - t_{i-1}}{\tau}})$ at time t_i. This process is an exponential saturation.

After t_i, the input returns to 0. The output $v_i(t, I_i)$ does not immediately respond accordingly. Instead, it decays exponentially.

The most important parameter is the membrane resistance R_m, which is the least upper bound to which $\frac{v_i(t, I_i)}{I_i}$ will go if the input is a step function $u(t_0)$, $t_0 \geq 0$, of infinite length. In this case, when the time t approaches ∞, $\frac{v_i(t, I_i)}{I_i}$ approaches R_m. Here, we assume that $I_i \neq 0$. When $I_i = 0$, $v_i(t, I_i) = 0$.

How fast and how high $v_i(t, I_i)$ goes up for $t \in [t_{i-1}, t_i]$ and how fast it decays for $t > t_i$ depends on the time constant $\tau = C_m(R_m + R_a)$. For a smaller time constant τ, the response $v_i(t, I_i)$ to the input will be faster. The exponential saturation process will be faster in $[t_{i-1}, t_i]$ and the exponential decay in $t > t_i$ will be faster too.

A dendrite can be considered a linear system as described by the input-output relation in the Laplace transform domain. Each pattern in the input set is a sequence of 50 current steps. The output with an input pattern taken from the input set is a summation over all the outputs with all pattern's constituent current steps as the inputs.

Corollary Let $P(t)$ be a dendrite's input pattern with n steps in $[0, T]$ for $T \in \mathbf{R}^{>0}$ and $P(t) = s_i(t, I_i) = I_i \in R$ if $t \in [t_{i-1}, t_i) \subset [0, T]$, $t_0 = 0$ and $t_n = T$, and $P(T) = P(t_n) = s_n(t, I_n) = I_n$. Let the output be $D(P(t))$. Then

$$D(P(t)) = \Sigma_{i=0}^{n} D(s_i(t, I_i)) = \Sigma_{i=0}^{n} v_i(t, I_i).$$

The above applies to linear systems in general. ♣

Now from the time domain we explain why, compared with a square current source at one end of a small dendrite segment, there is a distortion in the electrical potential at the other end. Figure 3.8 shows the electrical potential $v(t, I)(mV)$ for $R_m = 0.6k\Omega$, $I = 1000nA$, $C_m = 1\mu F$, and $R_a = 0.1\Omega$, and Figure 3.9 shows
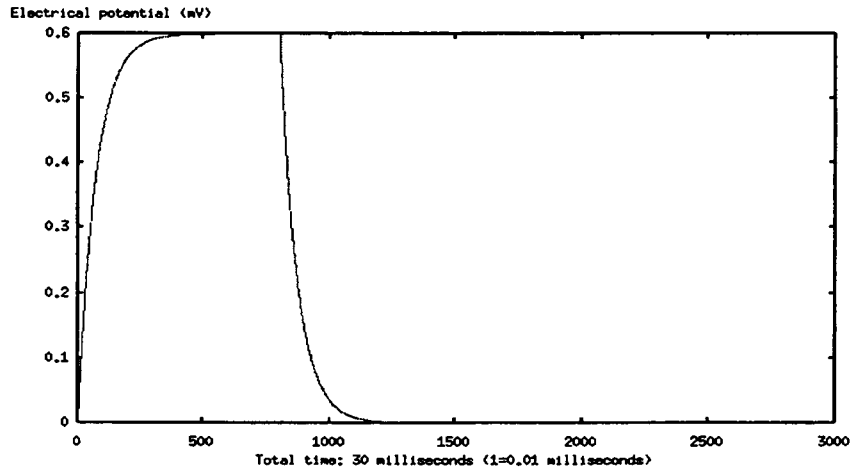
Figure 3.8: Electrical potential $v_i(t, I_i)(mV)$ at $R_m = 0.6k\Omega$, $I_i = 1000nA$, $C_m = 1\mu F$, and $R_a = 0.1\Omega$.
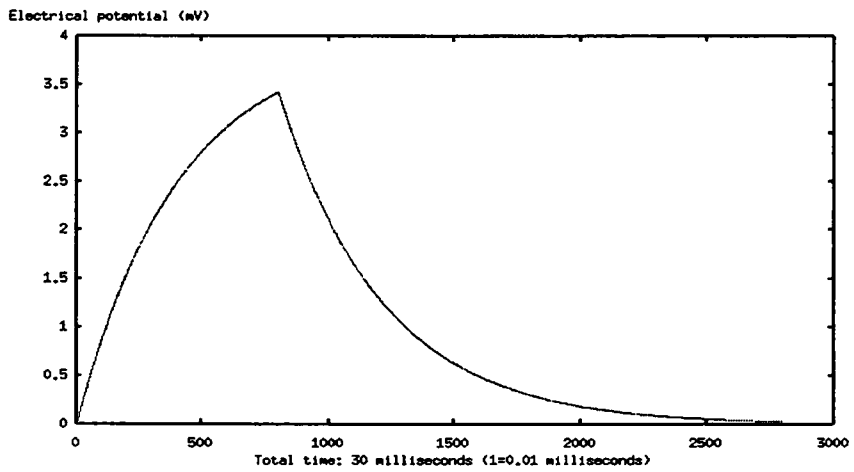
Figure 3.9: Electrical potential $v_i(t, I_i)(mV)$ at $R_m = 4k\Omega$, $I_i = 1000nA$, $C_m = 1\mu F$, and $R_a = 0.1\Omega$.

42

$v(t, I)$ with $R_m = 4k\Omega$ and the same other parameters. The source current is a square as follows:

$$s(t) = \begin{cases} 1000nA, & t \in [0 \text{ msec}, 8 \text{ msec}], \\ 0, & \text{otherwise.} \end{cases}$$

As an example, comparing Figure 3.8 with Figure 3.9, the higher the time constant τ (the higher the membrane resistance R_m or the lower the membrane conductance for this example), the higher the distortion, compared with the square input current, and the lower the time constant τ (the lower the membrane resistance R_m or the higher the membrane conductance), the higher the similarity with the square input current. Note that membrane resistance is the inverse of membrane conductance. The above is a description for the shape change. As for the magnitude of the electrical potential pattern, the higher the membrane resistance R_m, the higher the magnitude of the electrical potential.

The shape distortion is systematic. The higher the time constant τ, the longer the time for the electrical potential to rise exponentially to the maximum (exponential saturation) while the current square is being applied. Figure 3.9 shows that the front part (for low time t) of electrical potential has been chopped off a little compared with the square current input when the time constant τ is high. When the current square is over, the electrical potential starts to fall. How fast it falls depends on the time constant τ also. The higher the time constant τ,

43

the longer the time for the electrical potential to decay an amount and hence the higher the trail (for high t) of the electrical potential. All the distortions described here are the result of the lagging in the response to the input, both in rising and falling, determined by the time constant.

In summary, the membrane resistance R_m (and hence the membrane conductance) dominates how high the electrical potential will go. The maximum potential also depends on the time constant τ, particularly when the input current square is not long. How much the front part is chopped off and how much the tail is raised depend on the time constant τ. The parameters R_m and τ completely determine the lag in the response to the rise and fall of the input current pattern.

We have $D(P(t)) = \Sigma_{i=0}^{n} D(s_i(t, I_i)) = \Sigma_{i=0}^{n} v_i(t, I_i)$. This can be extended to the limit when the number of current steps $n \rightarrow \infty$. Any piecewise continuous function can be approximated to any degree by a sequence of steps of the same length, when the number of steps is sufficiently large. Now we can calculate the response of a linear system to an arbitrary piecewise continuous input function $f(t)$ if the impulse response of the system $h(t)$ is known.

The piecewise continuous input function $f(t)$ can be split up into a succession of rectangular impulses, each of width $\Delta\tau$ [Faulkner 69, p.15]. The area of each impulse is $f(\tau)\Delta\tau$. As $\Delta\tau \rightarrow 0$, this impulse generates a response of $(f(\tau)\Delta\tau)h(t-\tau)$. Let the number of rectangular impulses be n. Then the summation of the responses to the n impulses is the total response to all the rectangular

impulses together. When $n \to \infty$, this total response $g(t)$ is the convolution of $f(t)$ and $h(t)$, i.e.

$$g(t) = \int_{-\infty}^{\infty} f(\tau)h(t - \tau)d\tau.$$

By the definition of integration, the total response $g(t)$ of the system to the input $f(t)$ can be approximated within an arbitrary error $\epsilon(\in \mathbf{R})$ by the responses of the system to the n rectangular impulses when $n \geq N$ for a sufficiently large N. Note that the piecewise continuous input function $f(t)$ can also be approximated within an arbitrary error $\epsilon'(\in \mathbf{R})$ when $n \geq N'$, where N' is sufficiently large. We can pick $n = \max(N, N')$. Then we can approximate the response of the system to any input with the error we desire with n rectangular impulses. Consequently, it is sufficient to work with inputs composed of finite steps in the time domain. The result will be as close to the continuous counterpart to any degree as we desire. Note that no matter how small the constituent current steps are, the lagging of the response is still present in each step, and the total response is the summation of the responses to the individual current steps.

3.4.2 Properties of the Solution in the Frequency Domain

There are two parts in the frequency response: amplitude response and phase response.

The amplitude response is a measure of the amplification shown in the output

of an input frequency component, and phase response is a measure of a frequency component's delay in terms of phase shift. The actual time delay of a frequency component is defined as the phase delay.

Definition Let a system's transfer function be $H(s)$. The amplitude response [Rorabaugh 99, p.61] of the system is $|H(j\omega)|$, where $\omega \in \mathbf{R}^{\geq 0}$ is the frequency. For any frequency component with frequency ω in the input pattern, the amplitude of this frequency component is multiplied by $|H(j\omega)|$ in the output pattern.

Theorem The amplitude response of a small dendrite segment is

$$|H(j\omega)| = \frac{R_m}{\sqrt{\tau^2 \omega^2 + 1}},$$

where $\tau = C_m(R_m + R_a)$.

Proof

We have

$$
\begin{aligned}
|H(j\omega)| &= \left| \frac{V(j\omega)}{I(j\omega)} \right| \\
&= \left| \frac{1}{j\omega C_m(1 + \frac{R_a}{R_m}) + \frac{1}{R_m}} \right| \\
&= \left| \frac{R_m}{j\omega C_m(R_m + R_a) + 1} \right| \\
&= R_m \left| \frac{1 - j\omega C_m(R_m + R_a)}{[\omega C_m(R_m + R_a)]^2 + 1} \right| \\
&= \frac{R_m}{[\omega C_m(R_m + R_a)]^2 + 1} \mid 1 - j\omega C_m(R_m + R_a) \mid
\end{aligned}
$$

46

$$= \frac{R_m}{[\omega C_m(R_m + R_a)]^2 + 1}\sqrt{(1 + [\omega C_m(R_m + R_a)]^2)}$$

$$= \frac{R_m}{\sqrt{[\omega C_m(R_m + R_a)]^2 + 1}}$$

Let $\tau = C_m(R_m + R_a)$, then

$$|H(j\omega)| = \frac{R_m}{\sqrt{\tau^2\omega^2 + 1}}.\clubsuit$$

Theorem Some properties of the amplitude response $|H(j\omega)| = \frac{R_m}{\sqrt{\tau^2\omega^2+1}} = \frac{R_m}{\sqrt{[\omega C_m(R_m+R_a)]^2+1}}$ of a small dendrite segment are as follows:

(a) $|H(j\omega)|$ is a monotonely decreasing function of ω for $\omega \in \mathbf{R}^{\geq 0}$ and $|H(j\omega)| \to 0$ when $\omega \to \infty$,

(b) $\max(|H(j\omega)|) = R_m$ when $\omega = 0$,

Proof

(a) For any two $\omega_1, \omega_2 \in \mathbf{R}^{\geq 0}$ and $\omega_1 < \omega_2$, $|H(j\omega_1)| = \frac{R_m}{\sqrt{\tau^2\omega_1^2+1}} > \frac{R_m}{\sqrt{\tau^2\omega_2^2+1}} = |H(j\omega_2)|$ for $\tau, R_m > 0$ since $\sqrt{\tau^2\omega_1^2 + 1} < \sqrt{\tau^2\omega_2^2 + 1}$. We have $\omega_1 < \omega_2 \Rightarrow |H(j\omega_1)| > |H(j\omega_2)|$. Hence $|H(j\omega)|$ is monotonely decreasing for $\omega \in R^{\geq 0}$. It is clear that $|H(j\omega)| = \frac{R_m}{\sqrt{\tau^2\omega^2+1}} \to 0$ when $\omega \to \infty$.

(b) For a monotonely decreasing function of ω, the maximum is at the minimum of its domain, i.e. at 0. Then $\max(|H(j\omega)|) = |H(j\omega)|_{\omega=0} = |H(0)| = R_m$.

$\clubsuit$

As an example, comparing Figure 3.10 and Figure 3.11, we can see how membrane resistance affects the amplitude response. In Figure 3.10, the membrane resistance is 50Ω, and in Figure 3.11, the membrane resistance is 80Ω. When the membrane resistance increases from a lower value (50Ω) to a higher value (80Ω), the lower frequency components have higher amplitude response changes and the higher frequency components have lower amplitude response changes.

Definition The phase response $\theta(\omega)$ [Rorabaugh 99, p.61] is defined as

$$\theta(\omega) = \tan^{-1}\left(\frac{\Im[H(j\omega)]}{\Re[H(j\omega)]}\right),$$

and the phase delay $\tau_p(\omega)$ [Rorabaugh 99, p.61] is defined as

$$\tau_p(\omega) = \frac{-\theta(\omega)}{\omega}.\clubsuit$$

Theorem For a small dendrite segment, $\theta(\omega) = \tan^{-1}(-\omega\tau)$ and $\tau_p(\omega) = \frac{\tan^{-1}(\omega\tau)}{\omega}$, where $\tau = C_m(R_m + R_a)$.

Proof

We have the transfer function of a small dendrite segment

$$H(s) = \frac{V(s)}{I(s)} = \frac{1}{sC_m(1 + \frac{R_a}{R_m}) + \frac{1}{R_m}}.$$
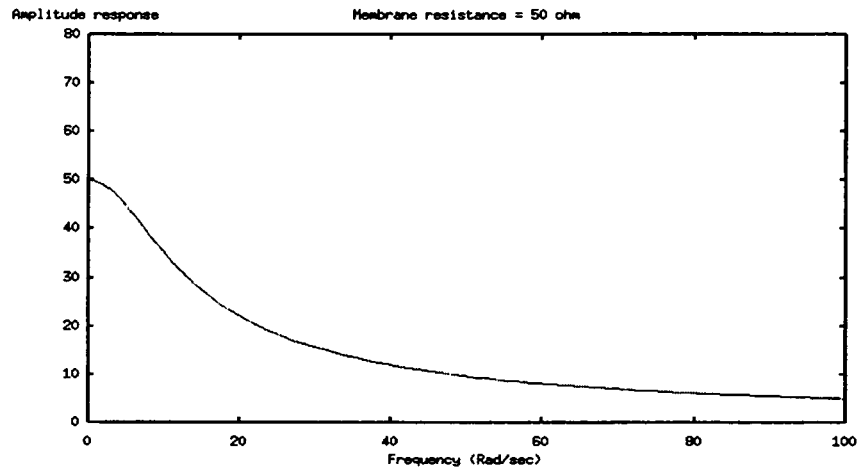
Figure 3.10: Amplitude response (Ω) of a small dendrite segment with $R_m = 50\Omega$, $C_m = 0.002F$, and $R_a = 1\Omega$.
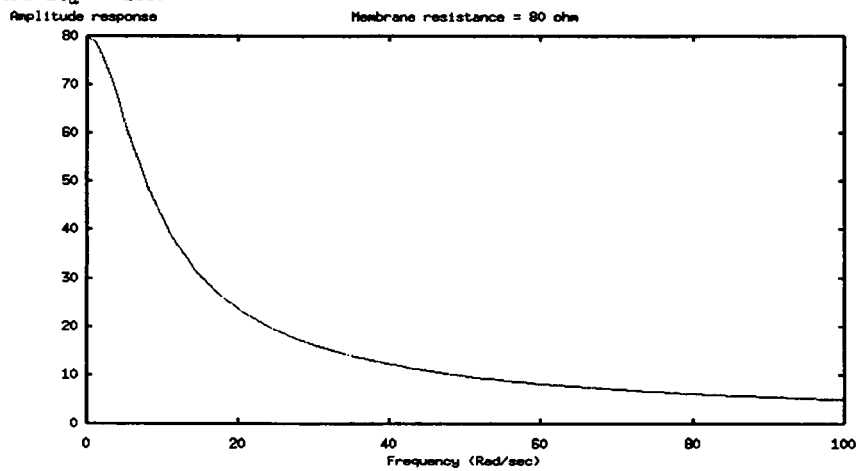
Figure 3.11: Amplitude response (Ω) of a small dendrite segment at $R_m = 80\Omega$, $C_m = 0.002F$, and $R_a = 1\Omega$.

$\Longrightarrow$

$$
\begin{aligned}
H(j\omega) &= \frac{1}{j\omega C_m\left(1 + \frac{R_a}{R_m}\right) + \frac{1}{R_m}} \\
&= \frac{R_m}{j\omega C_m(R_m + R_a) + 1} \\
&= \frac{R_m(1 - j\omega C_m(R_m + R_a))}{1 + [\omega C_m(R_m + R_a)]^2}.
\end{aligned}
$$

$\Longrightarrow$

$$
\left\{
\begin{aligned}
\Im[H(j\omega)] &= \frac{R_m[-\omega C_m(R_m+R_a)]}{1+[\omega C_m(R_m+R_a)]^2}, \\
\Re[H(j\omega)] &= \frac{R_m}{1+[\omega C_m(R_m+R_a)]^2}.
\end{aligned}
\right.
$$

$\Longrightarrow$

$$
\begin{aligned}
\theta(\omega) &= \tan^{-1}\left(\frac{\Im[H(j\omega)]}{\Re[H(j\omega)]}\right) \\
&= \tan^{-1}[-\omega C_m(R_m + R_a)] \\
&= \tan^{-1}(-\omega\tau). \\
\tau_p(\omega) &= \frac{-\theta(\omega)}{\omega} \\
&= \frac{-\tan^{-1}(-\omega\tau)}{\omega} \\
&= \frac{\tan^{-1}(\omega\tau)}{\omega}.\clubsuit
\end{aligned}
$$

Definition The accumulated delay for a linear system S from the frequency ω_p

to the frequency ω_q is defined as

$$D_S(\omega_p, \omega_q) = \int_{\omega_p}^{\omega_q} \tau_p(\omega)d\omega.$$

The total delay for the linear system S is defined as

$$D_S(0, \infty) = D_S(\omega_p, \omega_q) \mid_{\omega_p=0, \omega_q \to \infty} = \int_0^\infty \tau_p(\omega)d\omega. \clubsuit$$

According to [Beyer 78, p.440],

$$\int_0^\infty \frac{\tan^{-1}(ax) - \tan^{-1}(bx)}{x} dx = \frac{\pi}{2} \log \frac{a}{b},$$

where $a, b \in \mathbf{R}$ and $a, b > 0$.

For a small dendrite segment, we have the total delay

$$D_S(0, \infty) = \int_0^\infty \frac{\tan^{-1}(\omega\tau)}{\omega} d\omega.$$

Suppose that we have two small dendrite segments S_1 and S_2 with the time constants τ_1 and τ_2, respectively. The difference of the two total delays can be calculated as follows:

$$D_{S_1}(0, \infty) - D_{S_2}(0, \infty) = \int_0^\infty \frac{\tan^{-1}(\omega\tau_1) - \tan^{-1}(\omega\tau_2)}{\omega} d\omega$$

$$= \frac{\pi}{2} \log \frac{\tau_1}{\tau_2}.$$

Corollary Let τ_1 and τ_2 be the time constants for the two small dendrite segments S_1 and S_2, respectively, and $\tau_1, \tau_2 > 0$. The difference of the total delays of S_1 and S_2 is

$$D_{S_1}(0, \infty) - D_{S_2}(0, \infty) = \frac{\pi}{2} \log \frac{\tau_1}{\tau_2}. \clubsuit$$

In the above, when $\frac{\tau_1}{\tau_2} > 1$ (or $\tau_1 > \tau_2$), the difference of the two total delays $\frac{\pi}{2} \log \frac{\tau_1}{\tau_2} > 0$. This means that there is more total delay in a dendrite segment S_1 than in a dendrite segment S_2 when the time constant τ_1 of S_1 is larger than the time constant τ_2 of S_2.

3.5 The Learning Algorithm for the Dendritic Net

The learning algorithm is based on the generalized Fourier transform [Maclennan 92, Maclennan 94] of the impulse response of the linear system.

Definition The inner product of any two functions $\zeta(t)$ and $\phi(t)$, $t \in [0, T] \subset \mathbf{R}$, is $\langle \zeta(t), \phi(t) \rangle = \int_0^T \zeta(t) \phi(t) dt.$ $\clubsuit$

Definition An *input set I* of a linear system S defined in the interval $[0, T] \in \mathbf{R}$ of the system's operation range is any finite set of piecewise continuous functions

(patterns) defined in the range $[0, T]$.

If $\int_0^T p(t)p(t)dt = $ constant (independent of $p(t)$) for all functions $p(t)$, $t \in [0, T]$, in the input set I, then I is a *normalized* input set.

The *degree of similarity* of any two patterns (functions) in a normalized input set is defined as the inner product of the two patterns. If the inner product of two patterns is larger, the similarity is higher. ♣

For normalized patterns $\zeta(t)$ and $\phi(t)$ in an input set I defined in $[0, T]$, the inner product $\langle \zeta(t), \phi(t) \rangle$ is maximized when $\zeta(t) = \phi(t)$, $t \in [0, T]$.

Lemma Suppose that the piecewise continuous functions $\zeta(t)$, $\phi(t)$ are defined in $\mathbf{R}$ and $\zeta(t) = \phi(t) = 0$ for $t \notin [0, T]$. Then

$$\zeta(T - t) \star \phi(t) = \int_0^T \zeta((T - t) + p)\phi(p)dp.$$

Proof

Let $\zeta(T - t) = \xi(t)$. Then we have

$$
\begin{aligned}
\zeta(T - t) \star \phi(t) &= \xi(t) \star \phi(t) \\
&= \int_0^T \xi(t - p)\phi(p)dp \\
&= \int_0^T \zeta(T - (t - p))\phi(p)dp \\
&= \int_0^T \zeta((T - t) + p)\phi(p)dp. \text{ ♣}
\end{aligned}
$$

53

Theorem

Suppose that the piecewise continuous functions $\zeta(t)$, $\phi(t)$ are defined in $\mathbf{R}$ and $\zeta(t) = \phi(t) = 0$ for $t \notin [0, T]$. Then

$$\langle \zeta(t), \phi(t) \rangle = \zeta(T - t) \star \phi(t) \mid_{t=T}, t \in [0, T].$$

Proof

We already have

$$\zeta(T - t) \star \phi(t) = \int_0^T \zeta((T - t) + p)\phi(p)dp.$$

Then we have

$$
\begin{aligned}
\zeta(T - t) \star \phi(t) \mid_{t=T} &= \int_0^T \zeta((T - t) + p)\phi(p)dp \mid_{t=T} \\
&= \int_0^T \zeta(p)\phi(p)dp \\
&= \langle \zeta(t), \phi(t) \rangle.
\end{aligned}
$$

Hence $\langle \zeta(t), \phi(t) \rangle = \zeta(T - t) \star \phi(t) \mid_{t=T}$. ♣

Note that either one of the two functions $\zeta(T - t)$ or $\phi(t)$ can be interpreted as an input pattern and the other as the impulse response of a system. One possible interpretation is that $\zeta(T - t)$ is the impulse response of a system and $\phi(t)$ is

the input pattern. The equation in this theorem implies that if a linear system has an impulse response that is the time reverse of a particular pattern $\phi(t)$ in a normalized input set I, i.e. $\zeta(T-t) = \phi(T-t)$, then the linear system is a pattern matcher for $\phi(t)$. The output of the linear system is the final value (at time T) of the convolution of the input pattern $\phi(t) \in I$ with the system's impulse response $\phi(T-t)$, i.e. the time reverse of the input pattern $\phi(t)$.

Our purpose is to have a dendritic net which can differentiate between patterns. Assume that a dendritic net is approximately linear. Then the convolution of the impulse response of the dendritic net with the input pattern roughly describes the operation of the dendritic net.

Theorem Suppose that the two piecewise continuous functions $\zeta(t)$, $\phi(t)$ are defined in $\mathbf{R}$ and $\zeta(t) = \phi(t) = 0$ for $t \notin [0, T]$ and that there is a complete orthogonal basis of functions $\varrho_k(t)$, $k \in \{0, 1, 2, ...\}$, defined in the interval $[0, T]$, $T \in \mathbf{R}^{>0}$. Then we have

$$\langle \zeta(t), \phi(t) \rangle = \sum_{k=0}^{\infty} \frac{(\zeta(t) \star \varrho_k(t) \,|_{t=T})(\varrho_k(t) \star \phi(t) \,|_{t=T})}{\int_0^T \varrho_k{}^2(t)dt}.$$

Let $\mathrm{Norm}_k = \int_0^T \varrho_k{}^2(t)dt$ for $k \in \{0, 1, 2, ...\}$ be the normalization factors for the subnets' impulse responses. Then

$$\langle \zeta(t), \phi(t) \rangle = \sum_{k=0}^{\infty} \frac{(\zeta(t) \star \varrho_k(t) \,|_{t=T})(\varrho_k(t) \star \phi(t) \,|_{t=T})}{\mathrm{Norm}_k}.$$

Let $W_k = \frac{(\zeta(t) \star \varrho_k(t)|_{t=T})}{\text{Norm}_k}$ be the connection weight of the key synapse for subnet

k (see Figure 3.1), and let $O_k = \varrho_k(t) \star \phi(t) \mid_{t=T}$ be the final value of the output

pattern of the subnet k at the presynaptic site of the key synapse. Then

$$\langle \zeta(t), \phi(t) \rangle = \sum_{k=0}^{\infty} W_k \cdot O_k.$$

Proof

We already have

$$\langle \zeta(t), \phi(t) \rangle = \zeta(T - t) \star \phi(t) \mid_{t=T} \; .$$

We can write $\zeta(T-t)$, which can be considered the impulse response of a linear

system, in terms of the complete orthogonal basis:

$$\zeta(T - t) = \Sigma_{k=0}^{\infty} c_k \varrho_k(t),$$

where

$$c_k = \frac{\int_0^T \zeta(T - t) \varrho_k(t) dt}{\int_0^T \varrho_k{}^2(t) dt}.$$

By the definition of convolution, we have

$$c_k = \frac{\zeta(t) \star \varrho_k(t) \mid_{t=T}}{\int_0^T \varrho_k{}^2(t) dt}.$$

Then

$$\zeta(T - t) = \Sigma_{k=0}^{\infty} \frac{\zeta(t) \star \varrho_k(t) \mid_{t=T}}{\int_0^{\infty} \varrho_k{}^2(t)dt} \varrho_k(t).$$

Then the inner product can be written

$$\langle \zeta(t), \phi(t) \rangle = \left\{ \Sigma_{k=0}^{\infty} \frac{\zeta(t) \star \varrho_k(t) \mid_{t=T}}{\int_0^T \varrho_k{}^2(t)dt} \varrho_k(t) \right\} \star \phi(t) \mid_{t=T}.$$

Then

$$\langle \zeta(t), \phi(t) \rangle = \sum_{k=0}^{\infty} \frac{(\zeta(t) \star \varrho_k(t) \mid_{t=T})(\varrho_k(t) \star \phi(t) \mid_{t=T})}{\int_0^T \varrho_k{}^2(t)dt}.$$

The rest of the proof is just by substitution. ♣

Suppose that we have a dendritic net organized according to Figure 3.1 in Section 3.1, $\varrho_k(t)$ for $k = 1, 2, 3, \ldots$ is an orthogonal set, and the impulse response of subnet k is $\varrho_k(t)$. Note that this is an ideal case. In a simulation, the impulse responses of subnets may not be necessarily orthogonal.

Then the first factor $\zeta(t) \star \varrho_k(t) \mid_{t=T}$ is the value at the output site of subnet k at time T, i.e. the value at the presynaptic side before going across the key synapse to the summing tree (see Figure 3.1), when the input is the pattern $\zeta(t)$, to which the whole network is tuned.

For learning, this value divided by Norm_k will be used to set the synaptic weight (the multiplier in the linear synapse model) of the key synapse connecting

to the summing tree.

The second factor $\varrho_k(t) \star \phi(t) \mid_{t=T}$ is the output of subnet k at time T, i.e. the value at the presynaptic side of the key synapse connecting to the summing tree, with an arbitrary input pattern $\phi(t)$.

In summary, the first factor divided by the normalization factor for the subnet connecting to the summing tree will be used as the multiplier (the connection weight) for the key synapse, and the second factor is passed over from the presynaptic side to the postsynaptic side.

These two values are multiplied together to yield the postsynaptic value of the key synapse connecting to the summing tree from the subnet k. This postsynaptic value for subnet k will be summed with other values coming out from all the other subnets to get the output of the whole network.

Note that the summing tree does not exactly sum over all outputs from all subnets through the key synapses. Instead, the electrical potential at the root of the summing tree is roughly the superposition of all patterns coming from all the branches of the summing tree. The superposition is the summation divided by the number of branches. In addition, there are the longitudinal dissipation and the smoothing effects of a pattern propagating along a dendrite (see Section 4.3). Note that all branches of the summing tree have the same properties.

This paragraph summarizes the learning process. Before learning a pattern, a pulse of brief duration is injected into the dendritic net at its input site so that the

output patterns (functions) of subnets are approximately the impulse responses of all respective subnets. Then, the normalization factors for all subnets are determined. For learning a pattern, the pattern feeds into the whole network as the input. In this pass, the value of the first factor $\zeta(t) \star \varrho_k(t) \mid_{t=T}$ for subnet k is determined. Then the value of the first factor is divided by the normalization factor for this subnet, to set the synaptic weight of the key synapse of subnet k connecting to the summing tree. This process runs through all subnets. In this way, the network is intended to learn the specific input pattern and to become tuned to this particular input pattern. If this pattern is fed into the dendritic net again, the output will have the largest amplitude at time T compared to outputs of the other input patterns taken from the normalized input set.

In the above formulation, the presynaptic activity, i.e. the electrical potential at the output site of the kth subnet, and the subnet's normalization factor determine the synaptic efficacy of the corresponding kth subnet's key synapse, i.e. the synaptic connection weight being set to c_k/Norm_k.

In the formulation, the state of a dendritic net is represented by a vector of synaptic weights of the key synapses.

Definition Let S be a dendritic net with n subnets which implement the impulse responses $\varrho_k(t)$, $k \in \{0, 1, 2, ..., n \mid n \in \mathbf{N}\}$, and let $I = \{\phi_k(t) \mid k \in \{1, 2, 3, ..., m \mid m \in \mathbf{N}\}$ and $t \in [0, T], T \in \mathbf{R}^{>0}\}$ be a normalized input set.

The *state* of the dendritic net S with n subnets is a weight vector

$[W_1, W_2, ..., W_n]$, where $W_k \in \mathbf{R}$. W_k is the synaptic weight of the key synapse connecting subnet k to the summing tree. ♣

According to the simulations, subnets generated randomly do not have orthogonal impulse responses. In the above and the following definitions, the dendritic net S is not required to have subnets with orthogonal impulse responses. Hence, $\varrho_k(t)$, $k \in \{0, 1, 2, ..., n \mid n \in \mathbf{N}\}$, are not necessarily orthogonal in a simulation.

Definition Let S be a dendritic net organized into subnets which implement the impulse responses $\varrho_k(t)$, $k \in \{0, 1, 2, ..., n \mid n \in N\}$. S *recognizes* a particular pattern $\phi_i(t)$ in the normalized input set $I = \{\phi_i(t) \mid i \in \{1, 2, ..., m\}, m \in \mathbf{N}, \text{ and } t \in [0, T], T \in \mathbf{R}^{>0}\}$ *by the final value* if and only if there is a weight vector $[W_1, W_2, ..., W_n]$, where $W_j \in \mathbf{R}$ and $j \in \{0, 1, 2, ..., n \mid n \in \mathbf{N}\}$, such that the output at time T of the dendritic net S presented with $\phi_i(t)$ as the input: $S(\phi_i(t)) = \sum_{k=1}^{n} W_k O_k$, where $O_k = \varrho_k(t) \star \phi_i(t) \mid_{t=T}$, is the maximum among all outputs at time T: $\{S(\phi_j(t)) \mid j \in \{1, 2, 3, ..., m \mid m \in \mathbf{N}\}\}$, with inputs $\phi_j(t)$ taken from the normalized input set I. ♣

The above definition recognizes an input pattern by the final value of the dendritic net's output pattern. All the simulations use recognizing a pattern by the final value.

Suppose that an input pattern runs from time 0 to time T. There is a pattern coming out at the output site (see Figure 3.1). The *output* of the dendritic net is defined as *the value of the output pattern at time T*, which is the final value of

60

the output pattern, since the dendritic net runs from time 0 to time T. Note that this output is a scalar value, not the output pattern.

There is another definition of a dendritic net recognizing an input pattern. In neural processes, usually we are concerned about whether the maximal electrical potential of the output pattern in time interval $[0, T]$ is larger than the electrical potential threshold for initiating an action potential. In this process, it is the maximal value, not the final value of the output pattern that does the job. The maximum of the output pattern is not necessarily the value of the output pattern at time T. Recognizing a pattern by the maximal value of the output pattern can be similarly defined. It will not be formally defined here.

This does not mean that recognizing patterns with final value cannot possibly exist in the nervous system, since the nervous system is so complex. We cannot rule out the possibility that recognizing patterns with final value works with other processes at a critical time when the recognizing process just finishes at time T. For example, a conjunctive synapse could sample the output at time T. Cooperation among processes in the neural system seems to happen everywhere.

Some restrictions on the input set may result in a higher probability of having the two definitions of recognizing patterns be consistent.

In the following, we derive the condition of the input set under which the two definitions of recognizing patterns are close or the same.

Lemma Suppose that the function $\phi(p)$ is defined in $\mathbf{R}$, $\phi(p) = 0$ if $p \notin [0, T]$.

Then

$\int_0^T \phi((T-t)+p)\phi(p)dp \leq \int_0^T \phi^2(p)dp$ for $t \in [0,T]$.

Proof

Let $\zeta(p,t)$ be a function with variables $p, t \in \mathbf{R}$ and $\zeta(p,t) = \phi((T-t)+p)$. Note that for any fixed value of t the function ζ reduces to a function of a single variable and this resulting function is a shift in the function ϕ of its variable p in the amount $T-t$.

Since $\phi(p) = 0$ if $p \notin [0,T]$, then $\zeta(p,t) = 0$ if $p \notin [t-T,t]$, where $t \in [0,T]$. Note that $t \geq 0$ and $t - T \leq 0$.

For any t in $[0,T]$, consider the range $[t-T,T]$, which is a superset of both $[t-T,t]$ and $[0,T]$. Since we have $\phi(p) = 0$ if $p \notin [0,T]$ and $\zeta(p,t) = 0$ if $p \notin [t-T,t]$, then $\phi(p) = \zeta(p,t) = 0$ for $p \notin [t-T,T]$. By the definition of $\zeta(p,t)$ and the fact that the integration of a function ($\phi(p)$ in this case) is equal to the integration of its shifted function ($\zeta(p,t)$ for any t in this case) in their respective nonzero domains, we have

$\int_{t-T}^t \zeta^2(p,t)dp = \int_0^T \phi^2(p)dp$.

The ranges of the integrations in the above equation can be expanded. Note that $\zeta(p,t) = 0$ for p in $(t,T]$ and $\phi(p) = 0$ for p in $[t-T,0)$. Then we have

(1) $\int_{t-T}^T \zeta^2(p,t)dp = \int_{t-T}^T \phi^2(p)dp$.

Observe that

$\int_{t-T}^T (\phi(p) - \zeta(p,t))^2 dp \geq 0$

62

$\Rightarrow (2) \int_{t-T}^T \phi^2(p)dp + \int_{t-T}^T \zeta^2(p,t)dp \geq 2\int_{t-T}^T \phi(p)\zeta(p,t)dp.$

Combining (1) and (2), we have

$\Rightarrow \int_{t-T}^T \phi^2(p)dp \geq \int_{t-T}^T \phi(p)\zeta(p,t)dp.$

Since $\phi(p)$ is 0 in $[t-T, 0)$, we have

$\Rightarrow \int_0^T \phi^2(p)dt \geq \int_0^T \phi(p)\zeta(p,t)dt.$

Finally, by the definition of ζ

$\int_0^T \phi((T-t)+p)\phi(p)dp \leq \int_0^T \phi^2(p)dp$ for $t \in [0,T]$. ♣

Note that the left-hand side of the above inequality is $\phi(T-t) \star \phi(t), t \in [0,T]$

and the right-hand side is an instance of the left hand side with $t = T$. Then we

have the following result.

Corollary The maximum of the following expression

$\phi(T-t) \star \phi(t), t \in [0,T]$

happens at $t = T$, and the maximum value is $\int_0^T \phi^2(p)dp$ ♣

Suppose that we have a dendritic net with impulse response $\phi(T-t)$ and

the dendritic net is tuned to the pattern $\phi(t)$. Note that $\phi(T-t) \star \phi(t)$ is the

amplitude of the output pattern at time t. When the input to the dendritic net

is $\phi(t)$, the maximum of the output pattern happens at $t = T$ (by the Corollary).

This maximum amplitude of the output pattern is the inner product of the impulse

response and the input pattern.

Theorem Suppose that two arbitrary piecewise continuous functions $\zeta(t), \phi(p)$

are defined in **R**, $\zeta(t) = \phi(t) = 0$ if $t \notin [0,T]$ for some $T > 0$.

If $\zeta(t) \neq \phi(t)$, the maximum value of $\zeta(T-t) \star \phi(t), t \in [0,T]$, does not necessarily happen at $t = T$.

Proof

It suffices to prove that there exists two piecewise continuous functions $\zeta(t)$ and $\phi(t)$ for some $t' \in [0,T)$ such that $\zeta(T-t) \star \phi(t) \mid_{t=t'} \geq \zeta(T-t) \star \phi(t) \mid_{t=T}$.

The above requirement is equivalent to

$$\int_0^T \zeta((T-t') + p)\phi(p)dp \geq \int_0^T \zeta(p)\phi(p)dp$$

$$\Longleftrightarrow$$

$$(1) \int_0^T \{\zeta[(T-t') + p] - \zeta(p)\}\phi(p)dp \geq 0.$$

Note that $0 < T - t' \leq T$.

The following presents an example of two functions and a specific value of t' to satisfy requirement (1).

Let

$$\zeta(t) = \begin{cases} 1, & t \in [0, T/2], \\ 2, & t \in (T/2, T], \\ 0, & t \in (T, \infty) \text{ or } t \in (-\infty, 0). \end{cases}$$

Then

$$\zeta(T - t' + p) = \begin{cases} 1, & p \in [-(T - t'), T/2 - (T - t')], \\ 2, & p \in (T/2 - (T - t'), T - (T - t')], \\ 0, & p \in (T - (T - t'), \infty) \text{ or } p \in (-\infty, -(T - t')). \end{cases}$$

We define the function $\phi(t)$ in terms of the function φ of two variables. For a fixed t',

$$\phi(t) \equiv \varphi(t, T - t') \equiv \begin{cases} 1, & t \in [0, T - (T - t')], \\ 0, & t \in (T - (T - t'), \infty) \text{ or } t \in (-\infty, 0). \end{cases}$$

Then, for this instance, (1) becomes

$\int_{T/2-(T-t')}^{T/2} 1 dt = T - t' \geq 0$, for any $t' \in [0, T)$. ♣

This is not a surprise since $\zeta(t)$ and $\phi(t)$ are arbitrary functions. In general, the maximum of the output pattern of an arbitrary dendritic net with an arbitrary pattern as the input is not necessarily the final value of the output pattern.

According to the Proof of the Theorem, we have

$$\zeta(T - t) \star \phi(t) \mid_{t=t'} \geq \zeta(T - t) \star \phi(t) \mid_{t=T} \Longleftrightarrow \int_0^T \{\zeta[(T - t') + p] - \zeta(p)\}\phi(p) dp \geq 0$$

for some $t' \in [0, T)$.

The above is equivalent to the following:

$$\zeta(T-t) \star \phi(t) \mid_{t=T} > \zeta(T-t) \star \phi(t) \mid_{t=t'} \Longleftrightarrow \int_0^T \{\zeta[(T-t')+p] - \zeta(p)\} \phi(p) dp < 0$$

for some $t' \in [0, T)$.

Corollary Suppose that the functions $\zeta(t), \phi(p)$ are defined in **R**, $\zeta(t) = \phi(t) = 0$ if $t \notin [0, T]$ for some $T > 0$.

$\zeta(T-t) \star \phi(t) \mid_{t=T} > \zeta(T-t) \star \phi(t) \mid_{t=t'}$ for $\zeta(t) \neq \phi(t)$ and $t' \in [0, T)$ if and only if $\int_0^T \{\zeta[(T-t')+p] - \zeta(p)\} \phi(p) dp < 0$. ♣

The Corollary above suggests at least one way to pick the input set to make the definition of recognizing patterns by the final value of the output pattern the same as recognizing patterns by the maximum value of the output pattern.

Theorem Suppose that there are only positive piece-wise continuous monotonely decreasing patterns (functions) in the normalized input set for a system. Then recognizing a pattern by the final value and recognizing a pattern by the maximal value are the same.

Proof

Let $\zeta(t)$ and $\phi(p)$, defined in $[0, T]$, be any two patterns in the normalized input set. Consequently, they are both positive and monotonely decreasing.

Let's examine the integration $\int_0^T \{\zeta[(T-t')+p] - \zeta(p)\} \phi(p) dp$ in the Corollary. Note that $T-t' > 0$ since $t' \in [0, T)$ and $\zeta[(T-t')+p] - \zeta(p)$ is negative since $\zeta(p)$ is

66

monotonely decreasing. Then the whole integration $\int_0^T (\zeta((T-t')+p) - \zeta(p))\phi(p)dp$ is negative since $\phi(p)$ is positive.

By the Corollary, this integration holds true if and only if $\zeta(T-t) \star \phi(t) \mid_{t=T} > \zeta(T-t) \star \phi(t) \mid_{t=t'}$.

In this inequality, $\zeta(T-t)$ can be interpreted as the impulse response of the dendritic net which is tuned to the pattern $\zeta(t)$. The term $\zeta(T-t) \star \phi(t) \mid_{t=T}$ is the value of the output pattern at time T, and the term $\zeta(T-t) \star \phi(t) \mid_{t=t'}$ is the value of the output pattern at time t'. This inequality is equivalent to the statement that the value of the output pattern at time T is larger than the value of the output pattern at any time starting from time 0 to any time less then T. This means that the final value (at time T) of the output pattern is also the largest value of the output pattern. Hence the two definitions for recognizing patterns are the same under this condition. ♣

The condition that the patterns in the input set are positive and monotonely decreasing is not the only way to pick an input set to meet the requirement in the Corollary as long as the input set satisfies the inequality $\int_0^T \{\zeta[(T-t')+p] - \zeta(p)\}\phi(p)dp < 0$ in the Corollary.

A pattern which rises sharply to its maximum and then falls gradually until it settles around the dendrite's resting potential roughly meets the requirement of the theorem. A normalized input set with elements of this kind roughly meets the requirement for making the two definitions about the same. Many electrical

potential patterns of various time constants in the neural system roughly satisfy this requirement.

We have already defined the similarity of any two patterns in a normalized input set as the inner product of the two patterns. Now we define a measure of the difference between two output patterns. For two very different patterns, we expect the dendritic net to differentiate between them better than between two similar patterns.

Definition Let $\phi_1(t)$ and $\phi_2(t)$ be two patterns in a normalized input set for a dendritic net S operating in the time range $[0, T]$, and let the output pattern of S with input $\phi(t)$ be $S(\phi(t))$.

The measure of the difference when the dendritic net S takes the input pattern $\phi_1(t)$ and when it takes the input pattern $\phi_2(t)$ is $\max\{S(\phi_1(t)) \mid_{t \in [0,T]}\} - \max\{S(\phi_2(t)) \mid_{t \in [0,T]}\}$ when using the maximum values, and the measure is $[S(\phi_1(t)) - S(\phi_2(t))] \mid_{t=T}$ when using the final values. ♣

3.6 Correlation and the Performance of a Dendritic Net for Differentiating between Patterns

Random subnets do not necessarily implement the learning algorithm (see Section 3.5) accurately. For the purpose of measuring the performance of a dendritic net, correlation is introduced to measure the association or the linear relationship

between two random variables. It is used to relate inner products among input patterns and the operation of dendritic nets on input patterns. Note that, according to the learning algorithm, ideally the operation of a dendritic net is an inner product process on the input pattern and the pattern this dendritic net has been tuned to.

This section is not intended to review fundamental statistics. Only relevant definitions are presented here. How well a dendritic net performs is defined in terms of correlation.

Definition Let X and Y be two random variables. The correlation between X and Y can be estimated by the sample correlation r defined as follows:

$$r = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\Sigma_{i=1}^{n}(Y_i - \bar{Y})^2}},$$

where the n pairs, $n \in \mathbf{I}^{>0}$, of (X_i, Y_i), $i \in \{1, 2, 3, ..., n\}$, for $X_i, Y_i \in \mathbf{R}$ are a sample of size n from the bivariate population $\langle X, Y \rangle$.

For the ease of calculation, the sample correlation r is equivalent to

$$r = \frac{\Sigma_{i=1}^{n}X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{\Sigma_{i=1}^{n}X_i^2 - n(\bar{X})^2}\sqrt{\Sigma_{i=1}^{n}Y_i^2 - n(\bar{Y})^2}}.$$

In the above, $r \in [-1, +1]$, where $r = 1$ indicates perfect positive correlation, $r = 0$ shows no correlation, and $r = -1$ is for perfect negative correlation. ♣

Definition Let the normalized input set for a dendritic net D be $I = \{p_i(t) \mid$

$t \in [0, T], T \in \mathbf{R}^{>0}$, and $i \in \{1, 2, 3, ..., n \mid n \in \mathbf{N}\}\}$. For each pattern $p_i(t)$ in I, there is a sequence of inner products:

$$\langle p_i(t), p_1(t) \rangle, \langle p_i(t), p_2(t) \rangle, \langle p_i(t), p_3(t) \rangle, ..., \langle p_i(t), p_n(t) \rangle.$$

Let the dendritic net be tuned to pattern $p_i(t)$ by setting the connection weights of the key synapses according to the learning algorithm (see Section 3.5 for details), and let D_i denote the dendritic net D being tuned to the pattern $p_i(t)$. Then there is a corresponding sequence of outputs (final values of the output pattern) of the dendritic net:

$$D_i(p_1(t)), D_i(p_2(t)), D_i(p_3(t)), ..., D_i(p_n(t)).$$

Now, there are n ordered pairs in the following for calculating the correlation:

$$\{(\langle p_i(t), p_k(t) \rangle, D_i(p_k(t))) \mid k \in \{1, 2, 3, ..., n\}, n \in \mathbf{I}^{>0}\}.$$

The performance measure of the dendritic net D is the correlation between the sequence of inner products and the sequence of the dendritic net's outputs (final values) as shown in the above n pairs. The higher the correlation, the better the dendritic net performs.♣

For an input set with n patterns, there are n data pairs for each pattern. Then

the total number of data pairs to calculate the correlation is n^2 for n patterns. In our simulations, we have 20 input patterns. Therefore, there are 400 pairs of data for calculating the correlation.

The following Central Limit Theorem is taken from [Hays 71].

The Central Limit Theorem If a population has a finite variance σ^2 and mean μ, then the distribution of sample means from samples of N independent observations approaches a normal distribution with variance σ^2/N and mean μ as sample N increases. When N is very large, the sampling distribution of means is approximately normal. ♣

Definition Let the sample size be n ($\in \mathbf{I}^{>0}$), let the ith observation of the sample of the random variable X be X_i, and let the sample mean of X be $\bar{X}$. Then the sample standard deviation S is

$$S = \sqrt{\frac{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}. ♣$$

One of the consequences of the Central Limit Theorem is to estimate the mean of a sample. The following formulation is taken from [Berenson 96].

Theorem The $(1 - \alpha) \times 100\%$ confidence interval estimate for the mean μ_x is

$$\bar{X} \pm t_{n-1} \frac{S}{\sqrt{n}}$$

71

or

$$\bar{X} - t_{n-1}\frac{S}{\sqrt{n}} \leq \mu_x \leq \bar{X} + t_{n-1}\frac{S}{\sqrt{n}},$$

where $\bar{X}$ is the sample average of n observations, S is the sample standard deviation, and t_{n-1} is the critical value of the t distribution with $n-1$ degrees of freedom for an area of $\alpha/2$ in the upper tail. ♣

3.7 Discussion on the Numerical Complexity of the Simulations

Simulating a large artificial dendritic net demands much computing power. Normally the larger the network, the longer it takes to run the simulation on a digital computer. A more accurate statement is that the more complex the manipulation of the input pattern, the more time it takes to simulate. For a complicated manipulation of the input pattern, small time slices may be needed for numerical computation. Since the synaptic delay is implemented by a buffer, a larger memory space for buffers is needed for smaller time slices.

Under NEURON, the numerical simulation of the dendrites includes two parts [Hines 84, Hines 89, Hines 97]. The first part is to properly label cable segments to take advantage of the acyclic tree structure of dendrites. The second part is to solve the dynamic equation for the properly labeled tree structures. For large dendritic nets, labeling could take a substantial portion of the total simulation

time. The result of properly labeling the cable segments makes the second part substantially faster than directly solving the dynamic equation for a general graph structure.

Numerical instability is a problem in simulating highly complex dendritic nets, especially when there are complicated branchings and mergings of complicated patterns coming together at some points in the dendritic net. The only way to avoid numerical instability is to make the time slices for simulation sufficiently small so that the output stabilizes.

Note that if the time slices are reduced by half, it takes twice the time for the second part (after labeling the cable segments) to simulate the same network with the same input pattern. The complexity of a dendritic net depends on the size of the network, the number and locations of branchings and mergings of the dendrites, the number and locations of synapses, etc.

Chapter 4

Results of Simulations and

Discussions

4.1   Goal of the Simulations

The goal of the simulations is to show how well random dendritic nets composed

of dendrites and linear synapses can be modified into a state which enables them

to differentiate between temporal electrical current patterns. We are interested in

the structure of dendritic nets for recognizing or differentiating between patterns.

4.2   The Input Set

The following describes the normalized input set I for the dendritic net simula-

tions. In this normalized input set, all patterns are defined over the time interval

[0, 2 msec]. See Section 3.5 for the definition of a normalized input set.

There are 20 patterns of 50 electrical current steps in the normalized input set. All the patterns are randomly generated. A pattern starts from time 0 at current 0, and comes back to 0 before reaching 2 milliseconds. In this process, the next random number will be generated in the interval $[-1, +1]$. The next current step value of the pattern will be the present current pattern value plus 0.2 multiplied by this random number. Now we have a raw pattern of 50 values. A moving averaging process of window size 6 is applied to these 50 values. Suppose that we have raw data points in a sequence: $r_1, r_2, r_3, ..., r_n$, and $n = 50$. The new data sequence will be $d_1, d_2, d_3, ..., d_n$, and $d_i = \frac{r_i + r_{i-1} + r_{i-2} + r_{i-3} + r_{i-4} + r_{i-5}}{6}$, where r_{i-k} ,$0 \leq k \leq 5$, is set to 0 if $i - k \leq 0$. Each pattern is a sequence of 50 current steps, and each step is $\frac{2}{50}$msec. After all patterns are generated, they are normalized. The resulting patterns are the elements of the normalized input set.

The first 15 patterns (from pattern 0 to pattern 14) are positive, i.e. the electrical current steps are all higher than 0. In the next 5 patterns (from pattern 15 to pattern 19), the electrical current steps may be less than 0. See Appendix B for the figures of all patterns in the normalized input set I.

The similarity measure of two patterns is their inner product. All the inner products for all pattern pairs taken from the normalized input set are shown in Appendix D.

Note that the input set contains electrical current patterns, but not electrical

75

potential patterns. In principle, the learning algorithm (see Section 3.5) does not restrict the input to be a certain physical quantity. The input can be an electrical current pattern or an electrical potential pattern.

4.3 Longitudinal Dissipation and Smoothing Effect of the Dendrite

A pattern becomes smaller in electrical potential while propagating along a dendrite. Suppose that a current electrode at site 0 produces the change in membrane electrical potential ΔV_0. Let the membrane potential change at a distance of x along the dendrite from the electrode site be $\Delta V_m(t)$. Then the longitudinal dissipation can be described by the following equation [Koester 81]:

$$\Delta V_m(t) = \Delta V_0 e^{-x/\lambda},$$

where the length constant $\lambda = \sqrt{\frac{R_m}{R_a}}$. R_m and R_a are the membrane and axial (longitudinal) resistances respectively. When axial resistance R_a increases, the exponential decay factor $e^{-x/\lambda}$ decreases. It means that the longitudinal dissipation goes faster when the axial resistance increases.

The amplitude response $|H(j\omega)| = \frac{R_m}{\sqrt{\tau^2 \omega^2 + 1}}$ of a dendrite segment is a monotonely decreasing function of the frequency ω, $|H(j\omega)| \to 0$ when $\omega \to \infty$. It

implies that the amplitude of a higher frequency component will get lower amplification than that of a lower frequency component, and the very high frequency components have amplitudes close to 0. Consequently, a pattern becomes smoother while propagating along a dendrite because the amplitudes of higher frequency components, which characterize the sharp details of the pattern shape, of the pattern become smaller and the amplitudes of the lower frequency components become larger. This means that large variations in patterns in higher frequency components become smaller variations and amplitudes in lower frequency components are amplified while propagating along a dendrite.

Suppose that we have two normalized patterns $f(t)$ and $g(t)$ which are quite similar in shape except in the sharp details. Consequently, they have the same lower frequency components but have very different higher frequency components. The similarity (as the inner product of $f(t)$ and $g(t)$) could be quite different as they differ in higher frequency components, but they are more similar on a dendrite since higher frequency components are depressed. A dendritic net, which is built from dendrites and operates as performing an inner product according to the learning algorithm, cannot recognize or differentiate between patterns very accurately if the input patterns differ only in the higher frequency components. It means that a dendritic net does not correlate well to a high accuracy on the similarity of pattern shapes when higher frequency components are present in the input patterns.
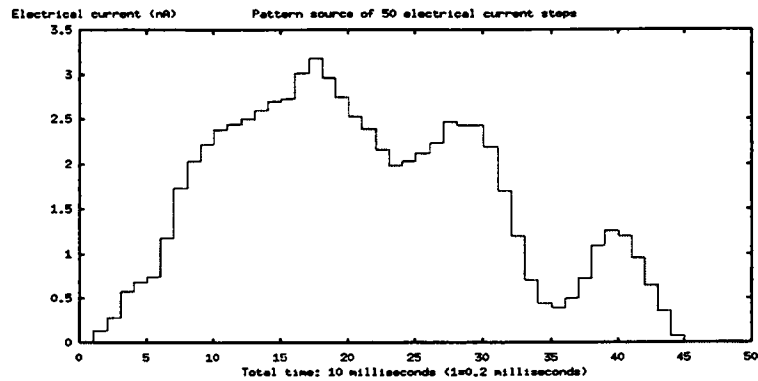
Figure 4.1: Electrical current steps to generate an electrical potential pattern

An example original electrical current pattern, which is a sequence of 50 steps of current amplitudes clamped to the dendrite, is shown in Figure 4.1. This current pattern generates Figures 4.2 - 4.5, which show the longitudinal dissipation and the smoothing effect of a dendrite. The membrane conductance is 0.001 S/cm^2, a classical value for the dendrite, and the membrane capacitance is $0.8\mu F/cm^2$. The total length of the dendrite is 300 microns.

Comparing Figure 4.1 with Figures 4.2 - 4.5, we can see clearly that sharp edges are gone, which reflects the smoothing effect of the dendrite. Figures 4.2 - 4.5 show the longitudinal dissipation: the longer the distance from the current source, the lower the amplitude of the pattern shape.

78
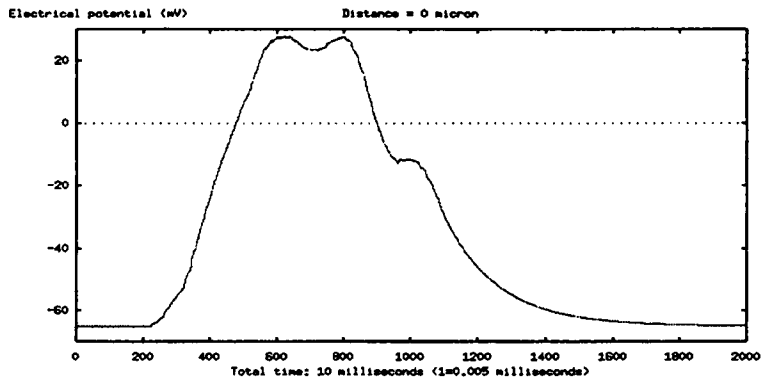
Figure 4.2: The electrical potential pattern at the site of the current source
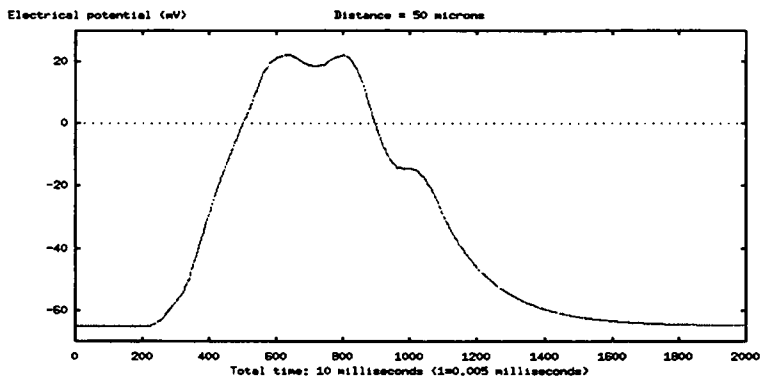
Figure 4.3: Longitudinal dissipation at 50 microns
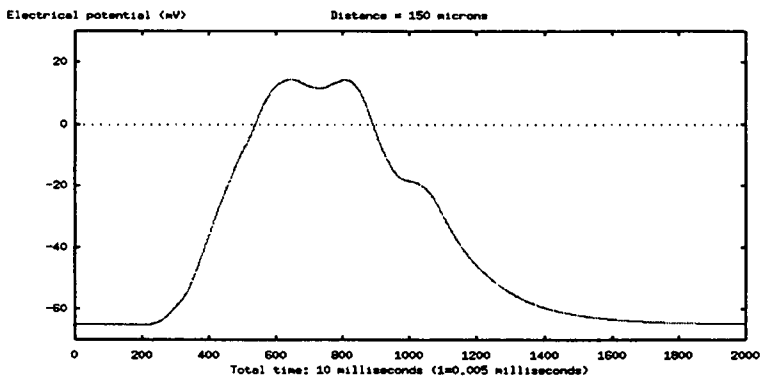
Figure 4.4: Longitudinal dissipation at 150 microns

79
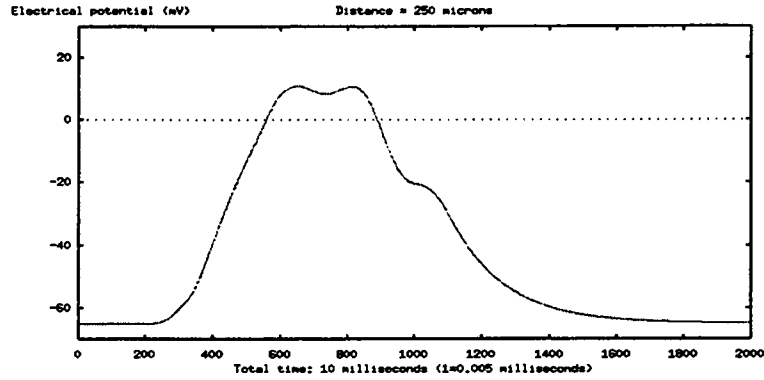
Figure 4.5: Longitudinal dissipation at 250 microns

4.4 Patterns Propagating under Various Membrane Conductances

The equivalent circuit and its solutions in Section 3.4 can be applied to a realistic dendrite, which is a cable segment with diameter and length.

The units of a dendrite cable segment's parameters are given here. The unit of the membrane conductance g_pas is S/cm^2, the unit of the diameter D is the micron, and the unit of the length L is the micron. In the following definitions, the numerical value and the unit of each parameter is written separately.

The conversion from the axial resistance r_a (Ω-cm) to resistance R_a in the equivalent circuit (see Figure 3.6) is defined as

$$
\begin{aligned}
R_a(\text{mega}\,\Omega) &= \frac{r_a(\Omega - \text{cm})L(\text{micron})}{\pi\left(\frac{D(\text{micron})}{2}\right)^2} \\
&= \frac{r_a(\Omega - m) \cdot 10^{-2} \cdot L(m) \cdot 10^{-6}}{\pi\left(\frac{D(m)}{2}\right)^2 \cdot (10^{-6})^2}
\end{aligned}
$$

80

$$= \frac{r_a L}{\pi(\frac{D}{2})^2} \cdot 10^4 \Omega$$

$$= \frac{r_a L}{\pi(\frac{D}{2})^2} \cdot 10^{-2} \text{mega } \Omega.$$

The membrane resistance R_m in the equivalent circuit can be defined in terms of the membrane conductance g_pas as follows:

$$R_m(\text{mega } \Omega) \equiv \frac{1}{g(\mu S)}$$

$$= \frac{1}{2\pi(\frac{D(\text{micron})}{2})L(\text{micron})\text{g\_pas}(S/cm^2)}$$

$$= \frac{1}{\pi(D(m) \cdot 10^{-6})L(m) \cdot 10^{-6} \cdot \text{g\_pas}(S/m^2) \cdot 10^4}$$

$$= \frac{1}{\pi \cdot D \cdot L \cdot \text{g\_pas} \cdot 10^{-6} \cdot S \cdot 10^{-2}}$$

$$= \frac{1}{\pi \cdot D \cdot L \cdot \text{g\_pas} \cdot \mu S \cdot 10^{-2}}$$

$$= \frac{1}{\pi \cdot D \cdot L \cdot \text{g\_pas} \cdot 10^{-2}}\text{mega } \Omega$$

The conversion from the membrane capacitance per unit area c_m ($\mu F/cm^2$) to the membrane capacitance C_m in the equivalent circuit is defined as

$$C_m(\mu F) = 2\pi(\frac{D(\text{micron})}{2})L(\text{micron})c_m(\mu F/cm^2)$$

$$= \pi(D(m) \cdot 10^{-6})L(m) \cdot 10^{-6} \cdot c_m(\mu F/m^2) \cdot 10^4$$

$$= \pi D \cdot L \cdot c_m \cdot 10^{-8}(\mu F).$$

Thus, we have the formulas for a dendrite cable segment's parameters C_m, R_m,

and R_a. From these we can calculate the amplitude response $|H(j\omega)| = \frac{R_m}{\sqrt{\tau^2\omega^2+1}}$ and the phase delay $\tau_p(\omega) = \frac{\tan^{-1}(\omega\tau)}{\omega}$, where $\tau = C_m(R_m + R_a)$, of the dendrite cable segment and compare them with the output pattern in the time domain. An example input pattern is shown in Figure 4.1. This comparison explains the difference between the shape of the input pattern at one end of the dendrite segment and the shape of the output pattern at the other end of the dendrite segment. There are three membrane conductance values, and for each membrane conductance value, there are three plots, one each for the output pattern, the amplitude response, and the phase delay of the dendrite cable segment.

In the NEURON runs, the parameters for the dendrite cable segment are as follows:

Parameter	Value	Unit
r_a	35.4	$\Omega - $ cm
c_m	0.8	$\mu F/$cm^2
L	300	microns
D	3.00	microns

Figures 4.6 - 4.8, Figures 4.9 - 4.11, and Figures 4.12 - 4.14 show the output pattern, the amplitude response, and the phase delay of the dendrite cable segment with membrane conductance g_pas $= 0.005$ $S/$cm^2, g_pas $= 0.001$ $S/$cm^2, and g_pas $= 0.0002$ $S/$cm^2 respectively.
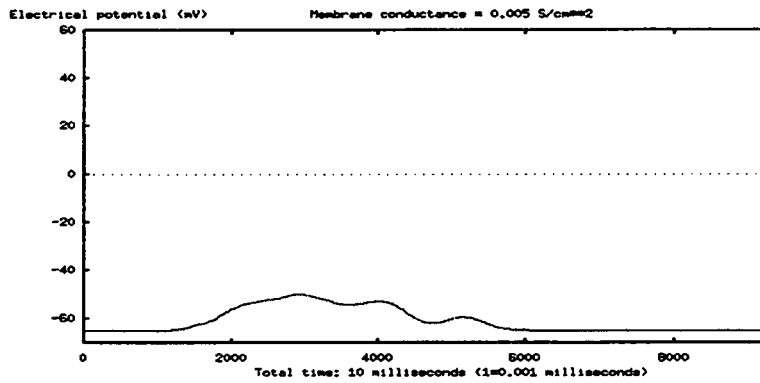
Figure 4.6: Pattern shape at the output end of the small dendrite segment with membrane conductance 0.005 S/cm^2
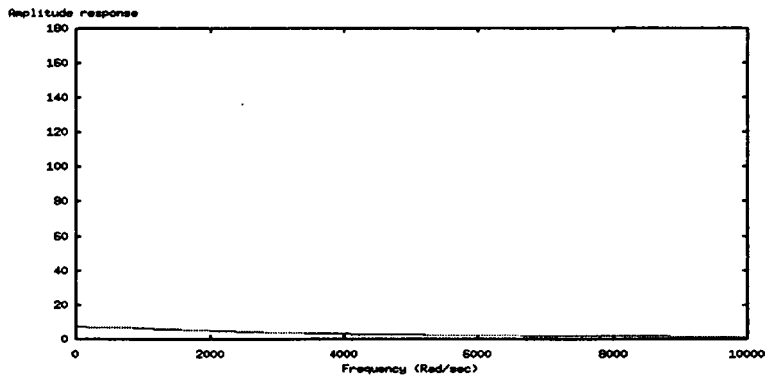
Figure 4.7: Amplitude response (Ω) of the small dendrite segment with membrane conductance 0.005 S/cm^2
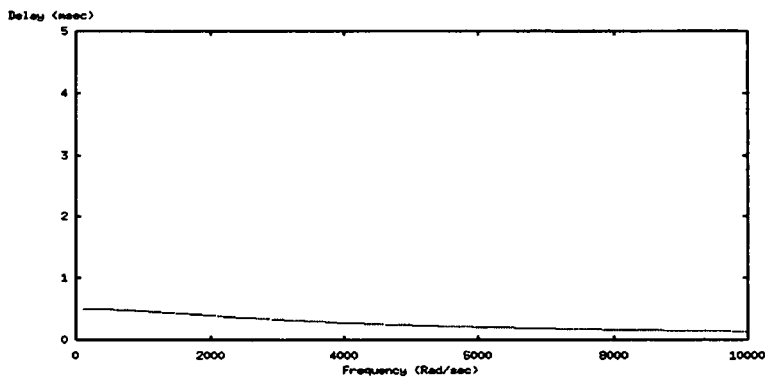
Figure 4.8: Phase delay of the small dendrite segment with membrane conductance 0.005 S/cm^2

83
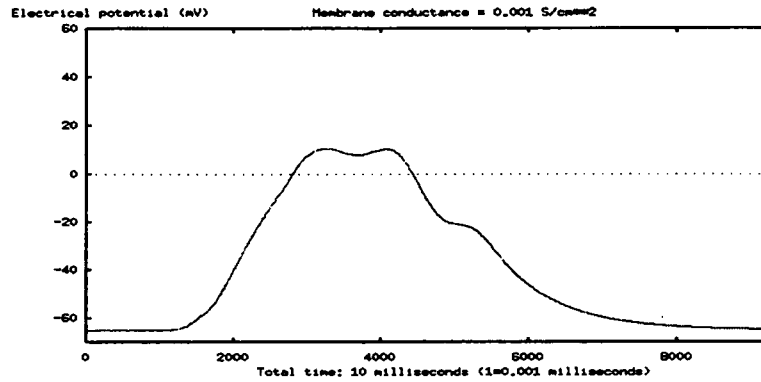
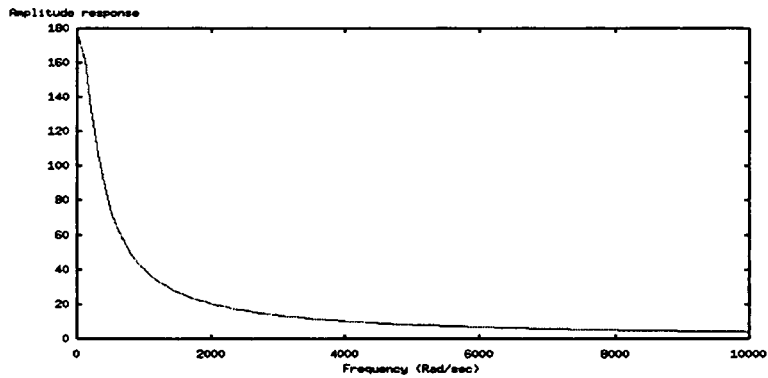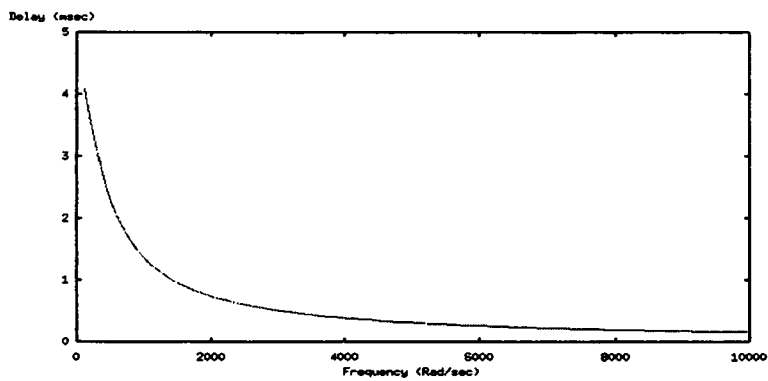Figure 4.9: Pattern shape at the output end of the small dendrite segment with membrane conductance 0.001 $S/$cm^2
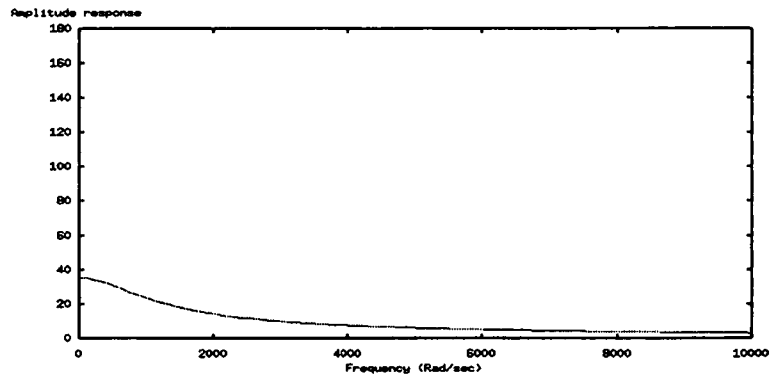
Figure 4.10: Amplitude response (Ω) of the small dendrite segment with membrane conductance 0.001 $S/$cm^2
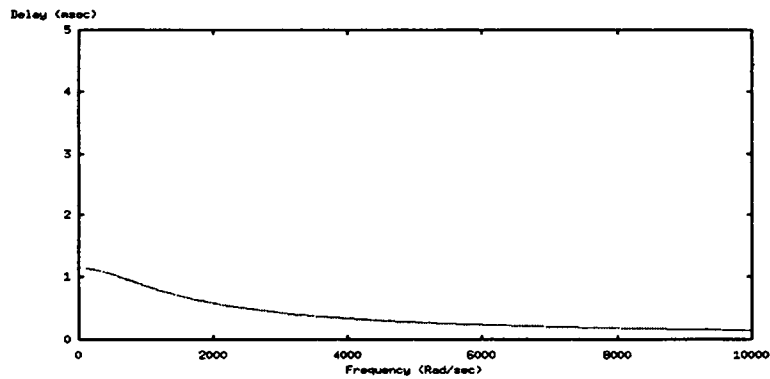
Figure 4.11: Phase delay of the small dendrite segment with membrane conductance 0.001 S/cm^2
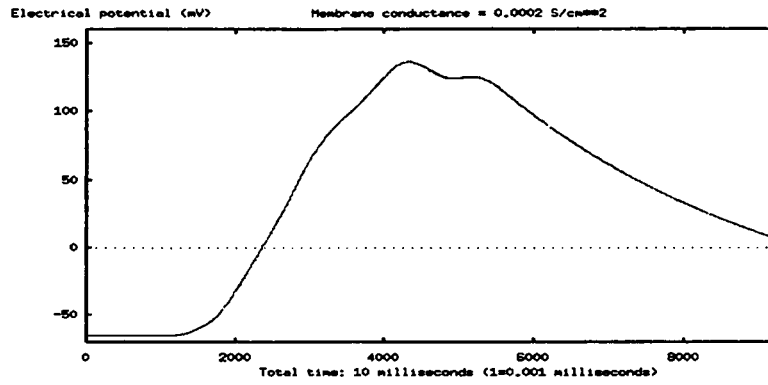
Figure 4.12: Pattern shape at the output end of the small dendrite segment with membrane conductance 0.0002 $S/$cm^2

Figure 4.13: Amplitude response (Ω) of the small dendrite segment with membrane conductance 0.0002 $S/$cm^2

Figure 4.14: Phase delay of the small dendrite segment with membrane conductance 0.0002 $S/$cm^2

Comparing the three groups of figures corresponding to the three membrane conductances, the lower the membrane conductance, the higher the amplitude response at a lower frequency. This means that the lower frequency components are magnified and the higher frequency components are suppressed when the membrane conductance value decreases. This is reflected in the output patterns in the time domain. We see a rounder pattern at a lower membrane conductance value. According to the plots of the phase delay, the lower the membrane conductance, the larger the phase delay at lower frequencies. We see a greater shift of the pattern to the right when the membrane conductance is smaller. This describes the distortion of the output pattern compared with the input pattern.

4.5 Dendritic Nets with Random Subnets for Differentiating Between Patterns

4.5.1 Introduction

The impulse responses of random subnets can be used as the components of the Fourier transform in the dendritic net learning algorithm (see Section 3.5). As the simulations show, the impulse responses of the random subnets do not have to be orthogonal to work. In our simulations, we use normalized impulse responses of the random subnets. Randomly generated subnets can do the job well when the time constant of the dendrite is small so that the distortion of the input

patterns on the dendrites is not severe. For a conductance equal to or larger than $0.001 \, S/\text{cm}^2$, and a membrane capacitance equal to $0.8 \mu F/\text{cm}^2$, random dendritic nets with 5, 10, and 15 subnets can differentiate to a high degree between input patterns of electrical current.

4.5.2 Parameters for Simulating Random Dendritic Nets

Table C.1 in Appendix C shows the parameters for generating the random dendritic nets in the simulations. Both excitatory and inhibitory synapses are present in all random dendritic nets. Each synapse in a subnet is equally likely to be excitatory or to be inhibitory. For a positive square input (electrical potential higher than the resting potential) at the presynaptic site, some synapses will have a postsynaptic electrical potential higher than the resting potential and all the other synapses will have a postsynaptic electrical potential lower than the resting potential. The parameter g_syn is for the linear synapse model (see Appendix A). The higher the g_syn, the closer the postsynaptic electrical potential is to the presynaptic electrical potential. Some of the parameters are range variables, which are uniformly, randomly generated in the indicated ranges. The branch length of the first level of a dendritic tree is generated in the range $[0, 60 \, \text{microns}]$. The next level is generated in the range $[0, 60 \times \text{(next level branch factor)} \, \text{microns}]$ and so on. Since patterns dissipate quite fast along a dendrite, an amplification factor of 3 is introduced for each synapse. A pattern becomes three times higher

across a synapse.

4.5.3 Differentiating Patterns by Random Dendritic Nets

A total of three membrane conductance values are used for generating random dendritic nets. The subnet size of a random dendritic net can be 5, 10, or 15 subnets. For each conductance value and subnet size, there are 8 random dendritic nets generated. Any two of the 8 random dendritic nets have different random subnets. In order not to have any association between any two random dendritic net structures, a given subnet in a dendritic net can not appear in any other dendritic net with the same membrane conductance value and the same subnet size. Each of the random dendritic nets takes input patterns from the same normalized input set.

The performance of each random dendritic net is measured in terms of statistical correlation (see Section 3.6) to show how well this dendritic net differentiates between patterns. The performances of the random dendritic nets with different conductance values and different subnet sizes are shown in Tables C.2 - C.8. Each table in C.2 - C.8 shows the performance of 8 different random dendritic nets with the same conductance value and the same number of subnets. The mean and the standard deviation of the performances of all the 8 random dendritic nets are calculated and listed in the table, except for the dendritic nets with low membrane conductance $(0.0002 S/\text{cm}^2)$, which show bad performance.

All simulations in this section are summarized in Table C.9. This table shows the confidence intervals for each set of 8 dendritic nets. According to the Central Limit Theorem (see Section 3.6), when the sample size is very large, the distribution of the sample means is approximately normal. In our case of 8 observations (simulations), it should be treated as an approximation to the normal distribution.

The conclusion from the above simulations is that dendritic nets with membrane conductance $0.001S/\text{cm}^2$ or higher, membrane capacitance $0.8\mu F/\text{cm}^2$, and subnet sizes $5-15$ can differentiate between electrical current patterns to a high degree as indicated by the high correlations. Randomly generating the subnets for the Fourier components is good enough for building competent dendritic nets for differentiating between patterns.

Note that randomly generating structures is one of the easiest processes in nature. The orthogonality of the impulse responses of the subnets does not seem to be important. For random subnets with membrane conductance $0.005\ S/\text{cm}^2$, the average of the absolute values of the inner products of the impulse responses of any two subnets is about 26% of the norm of a subnet's impulse response.

Random subnets work surprisingly well. Although a random process may be the easiest way for the neural system to build subnets, we still do not know if the real dendritic net works exactly the same way. At least we have simulated one possible way the dendritic net can be modified to differentiate between patterns. It would be interesting to set up a biological experiment to confirm the simulations.

4.5.4 A Detailed Example of Differentiating Between Patterns by a Random Dendritic Net

To get a feel for how well a dendritic net can differentiate between patterns, the detailed results of simulating one of the dendritic nets with 10 subnets and a conductance of $0.005S/\text{cm}^2$ are shown in Table C.10 and Appendix D. The figures in Appendix D show the relation between the cross inner products of all input patterns in the input set with the pattern the dendritic net is being tuned to. They also show the final values of dendritic net output patterns for different input patterns (see Section 3.6 for performance measure). Roughly, the final values of output patterns are strongly associated with the cross inner products of input patterns.

For each entry of Table C.10, "Pattern number" indicates what pattern the dendritic net is tuned to, and both "Rank for maximum value" and "Rank for final value" show that pattern's rank, among all the patterns in the input set, when it is fed into this dendritic net. For example, in the first entry of Table C.10, when the dendritic net is tuned to Pattern 0 and takes Pattern 0 as the input, the rank of the maximum value of the output pattern is 1 and the rank of the final value of the output pattern is 2 compared with the output patterns resulting from the input of all the patterns in the input set. Ideally, every entry under "Rank for maximum value" and "Rank for final value" should be 1. This

would mean the dendritic net perfectly recognized all the patterns in the input set, and that recognizing by the final value and recognizing by the maximal value coincided. The entries under "Rank for final value" in Table C.10 are mostly low digits, which indicate that the dendritic net performs well, though not perfectly, in ranking the patterns in the input set.

The relation between the final values and the pattern inner products is shown in Figure 4.15. The correlation (see Section 3.6 for definition) between the final values of output patterns and the inner products of input patterns is 0.9869. With 400 data and the high correlation value, the probability that they are not associated is extremely low. The plot in Figure 4.15 and the correlation value confirm that the two groups of figures, i.e. the final values of output patterns and the inner products of input patterns (in Appendix D), are correlated and visually similar.

In the presentation of the results, both ranking and statistical correlations are used. Ranking is a kind of measure on how well a dendritic net recognizes a pattern among all patterns in the input set. This measure is specific to the input set used for simulation. Statistical correlation more objectively describes how strong the input patterns and the final values of the output patterns relate. It is good for measuring how well a dendritic net differentiates between patterns. The patterns in the input set used in this study happen to be quite similar. As a result, high statistical correlation is required for accurate ranking for this input
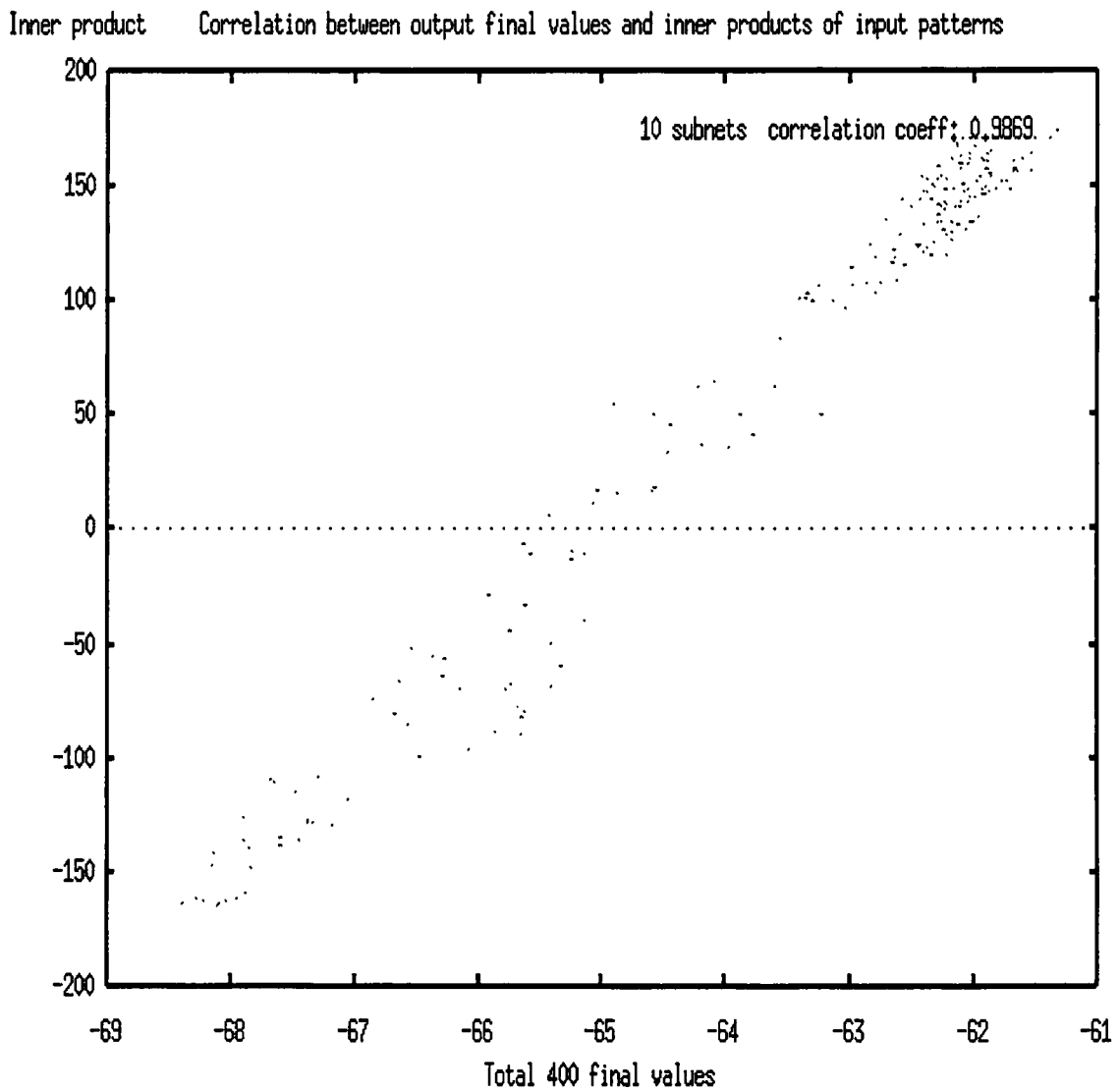
Figure 4.15: The correlation between the output final values and the input pattern inner products for a dendritic net with 10 subnets and membrane conductance $0.005S/\text{cm}^2$.

set. In dendritic nets where the patterns processed are not that close to each other, the dendritic net may not need to have high correlation to rank the input patterns accurately.

4.5.5 Fault-tolerance of Random Dendritic Nets

For the purpose of studying the fault-tolerance of random dendritic nets, 5-subnet random dendritic nets are generated first, then 5 more random subnets are added to make the number of subnets 10, and then 5 more subnets are added to make 15-subnet random dendritic nets. They can be viewed in the other way. The 10-subnet random dendritic nets are the results of deleting 5 subnets from the 15-subnet random dendritic nets, and the 5-subnet random dendritic nets are the results of deleting 5 subnets from the 10-subnet random dendritic nets. There are 8 random dendritic nets generated in each group of 5-subnet, 10-subnet, or 15-subnet random dendritic nets.

The performances of all random dendritic nets are measured to see if there are significant increases or decreases in performance after deleting/adding random subnets. All performances are measured in terms of correlations (see Section 3.6). All the results of simulations are summarized in Tables E.1 - E.4 of Appendix E. The membrane conductance is $0.001S/\text{cm}^2$ and the membrane capacitance is $0.8\mu F/\text{cm}^2$ for the dendrites. The performances of 5-subnet random dendritic nets are shown in Table E.1, Table E.2 shows the performances of 10-subnet

93

random dendritic nets, and Table E.3 shows the performances of 15-subnet random dendritic nets. All the performances are summarized in Table E.4.

Comparing Table E.1 with Table E.2, the mean of performance increase by adding 5 subnets to the 5-subnet random dendritic nets is 0.0066 in correlation. Comparing Table E.2 with Table E.3, the mean performance increase from adding 5 subnets to the 10-subnet random dendritic nets is 0.0077 in correlation.

Comparing Table E.2 with Table E.1 again, the performances of the dendritic nets for Runs 1, 3, 4, 5, 8 increase as a result of adding 5 random subnets to the 5-subnet random dendritic nets, while the performances of the dendritic nets for Runs 2, 6, 7 decrease as a result of adding 5 more random subnets to the 5-subnet random dendritic nets.

Comparing Table E.3 with Table E.2 again, the performances of the dendritic nets for Runs 2, 4, 5, 6, 7 increase adding 5 random subnets to the 5-subnet random dendritic nets, while the performances of the dendritic nets for Runs 1, 3, 8 decrease adding 5 random subnets to the 10-subnet random dendritic nets.

As Table E.1-E.4 show, there is only small difference in the performance correlations of the dendritic nets with 5, 10, and 15 subnets. The conclusion is that adding or deleting some subnets to or from a random dendritic net does not change the performance of the dendritic net much. This implies that random dendritic structures are highly fault-tolerant. It would be interesting to look at all structures in the brain and to determine if random structures are the key to the high

fault-tolerance of the brain. The emphasis is on randomness.

An interesting case to investigate is when the number of subnets is reduced to 1. Table E-5 lists the results. In general, the performance is not good, but among the 8 dendritic nets generated, one shows good performance (correlation 0.9593). This means that a properly wired dendritic net with 1 subnet can do a good job.

4.6 Discussion on The Input Set of Electrical Current Patterns and Its Associated Input Set of Electrical Potential Patterns

Suppose that we have an input set I of electrical current patterns. Since an electrical current pattern generates an electrical potential pattern at the same site, the set of electrical potential patterns generated by the current patterns in the input set I can be considered an associated input set I'. Note that both I and I' will lead to the same weight vector for the key synapses for learning. For recognizing or differentiating between patterns, the learning algorithm can take an input set of either electrical current patterns or electrical potential patterns. The output also can be an electrical current pattern or an electrical potential pattern.

In the nervous system, the current coming from the presynaptic side of a synapse can be considered an input, and the output is often an electrical potential pattern, which determines whether the output will generate an action potential

or not.

If we have a normalized input set of current patterns, then the associated electrical potential patterns generated at the input site may not be necessarily normalized due to distortion. If the distortion is small, then it is not significant whether we use an input set of electrical current patterns or an input set of electrical potential patterns

Suppose that the electrical potential is the more relevant quantity at the input site and the output is an electrical potential pattern. If the input electrical current patterns in the normalized input set generate distorted electrical potential patterns when they are applied to a dendrite, then the dendritic net is taking a possibly unnormalized input set I' of electrical potential patterns instead of the original normalized input set I of electrical current patterns.

The shape change of the patterns due to distortion affects the normalization of the patterns in the time interval of the input set, and may change the order of the inner products among the patterns. Since the learning algorithm approximates the inner product on normalized patterns, the shape distortion due to high membrane resistance R_m has a great influence on the learning process when we work on a normalized input set of electrical current pattern. In order to preserve the shapes of the patterns to a high degree, the membrane conductance has to be high enough. How large the conductance should be depends on how well the dendritic net has to differentiate between input current patterns.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

Dendritic nets with high membrane conductance and with a small number (5-15) of randomly generated subnets can differentiate to a high degree between input electrical current patterns by the final value of the output electrical potential pattern after modifying the synaptic weights of the key synapses. As the simulations show, the statistical correlation coefficient between the final values of output patterns and the inner product of their corresponding input patterns with the pattern to which the dendritic net is tuned is very high for high membrane conductance. The modification of a dendritic net is small scale: for a random dendritic net with n subnets for $n \in \mathbf{I}^{>0}$, there are n synaptic weights to modify. It seems that random structures are the easiest to build.

The random dendritic structures are highly fault-tolerant. Adding or deleting some subnets to or from a random dendritic net does not change the performance much. This property is highly desirable for building computing devices.

5.2 Future Work

An important issue is to find out why dendritic nets with random subnets perform well even when the impulse responses of subnets are quite dependent. It is necessary to understand why the dependence of random subnets does not degrade the performance much. This may be important for understanding the building blocks of intelligence and the process of biological evolution.

We have shown that random dendritic structures are highly fault-tolerant. It would be interesting to further examine whether random structures are the reason why the brain is highly fault-tolerant.

In this work, we have investigated the capabilities of general dendritic structures for recognizing electrical current patterns. For further simulations, the input patterns may have to be more specific, and they could be obtained by recording patterns at certain locations of the nervous system. For example, repetitive patterns are common in the nervous system. Normalizing input patterns may be an important issue, since the amplitudes of patterns may vary widely.

Often, the nervous system operates on the maximal value of a pattern (e.g. to

trigger an action potential). It may be interesting to see if we can find important input patterns in some part of the nervous system for which maxima and final values of the output patterns coincide. Another direction is to devise an algorithm to work on the maximal values of output patterns.

Bibliography

Bibliography

[Arfken 95] George B. Arfken and Hans J. Weber (1995). *Mathematical Methods for Physicists*, fourth edition. Academic Press.

[Bazhenov 98] M. Bazhenov, R. Huerta, M.I. Rabinovich, and T.J. Sejnowski (1998). Cooperative Behavior of a Chain of Synaptically-coupled Neurons. *Physica D* **116**: 392-400.

[Berenson 96] Mark L. Berenson and David M. Levine (1996). *Basic Business Statistics*, sixth edition. Prentice Hall, Inc.

[Beyer 78] William H. Beyer, editor (1978). *CRC Handbook of Mathematical Sciences,* 5th edition. CRC Press, Inc.

[Bullock 93] Theodore Holmes Bullock (1993). *How Do Brains Work ? : Papers of a Comparative Neurophysiologist.* Boston: Birkhäuser.

[Bush 95] P. Bush and T.J. Sejnowski (1995). Models of Cortical Networks. In *The Cortical Neuron*, edited by I.Mody and M. Gutnik. Oxford University

Press.

[Churchland 92] Patricia S. Churchland and Terrence J. Sejnowski (1992). *The Computational Brain*. A Bradford Book, The MIT Press, Cambridge, Massachusetts.

[Cowan 85] C.F.N. Cowan and P.M. Grant, editors (1985). *Adaptive Filters*. Prentice-Hall, Inc.

[DeCarlo 89] Raymond A. DeCarlo (1989). *Linear Systems*. Prentice Hall.

[Destexhe 94a] A. Destexhe, Z. F. Mainen, and T. J. Sejnowski (1994). An Efficient Method for Computing Synaptic Conductances Based on a Kinetic Model of Receptor Binding. *Neural Computation* **6**: 14-18.

[Destexhe 94b] Alain Destexhe, Zachary F. Mainen, and Terrence J. Sejnowski (1994). Synthesis of Models for Excitable Membranes, Synaptic Transmission and Neuromodulation Using a Common Kinetic Formalism. *The Journal of Computational Neuroscience* **1**: 195-231.

[Destexhe 95a] Alain Destexhe, Zachary F. Mainen, and Terrence J. Sejnowski (1995). Synaptic Currents, Neuromodulation and Kinetic Models. In *The Handbook of Brain Theory and Neural Networks*, edited by Michael A. Arbib. MIT Press.

[Destexhe 95b] Alain Destexhe, Diego Contreras, Mircea Steriade, Terrence Sejnowski, and John Huguenard (1995). Computational Models Constrained by Voltage-Clamp Data for Investigating Dendritic Currents. In *Computational Neuroscience*, edited by J. Bower, pp. 53-58. Academic Press, New York.

[Destexhe 95c] Alain Destexhe, Zachary F. Mainen, and Terrence J. Sejnowski (1995). Fast Kinetic Models for Simulating $AMPA$, $NMDA$, $GABA_A$ and $GABA_B$ Receptors. In *The Neurobiology of Computation*, edited by J. Brower, pp. 9-14. Kluwer Academic Press, Norwell, MA.

[Destexhe 96] A. Destexhe, D. Contreras, M. Steriade, T.J. Sejnowski, and J.R. Huguenard(1996). In Vivo, in Vitro and Computational Analysis of Dendritic Calcium Currents in Thalamic Reticular Neurons. *Journal of Neuroscience* **16**: 169-185.

[Destexhe 97] Alain Destexhe and Terrence J. Sejnowski (1997). Synchronized Oscillations in Thalamic Networks: Insights from Modeling Studies. In *Thalamis*, Volume III, edited by M. Steriade, E.G. Jones, and D.A. McCormick, pp. 331-371. Elsevier, Amsterdam.

[Dowling 92] John E. Dowling (1992). *Neurons and Networks : An Introduction to Neuroscience.* Belknap Press of Harvard University Press.

[Faulkner 69] E. A. Faulkner (1969). *Introduction to the Theory of Linear Systems*. Chapman and Hall LTD.

[Gluck 90] Mark A. Gluck, Eric S. Reifsnider, and Richard F. Thompson (1990). Adaptive Signal Processing and the Cerebellum: Models of Classical Conditioning and VOR Adaptation. In *Neuroscience and Connectionist Theory*, edited by Mark A. Gluck and David E. Rumelhart, pp.131-185. Lawrence Erlbaum Associates, Inc.

[Hameroff 87] Stuart R. Hameroff (1987). *Ultimate Computing*. Elsevier Science Publishers B.V.

[Hays 71] William L. Hays and Robert L. Winkler (1971). *Statistics: Probability, Inference, and Decision*. Holt, Rinehart and Winston, Inc.

[Haykin 96] Simon Haykin (1996). *Adaptive Filter Theory*, third edition. Prentice-Hall, Inc.

[Hertz 91] John Hertz, Anders Krogh, and Richard G. Palmer (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley Publishing Company.

[Hildebrand 65] Francis B. Hildebrand (1965). *Methods of Applied Mathematics*. Prentice-Hall, Inc.

[Hines 84] Michael Hines (1984). Efficient Computation of Branched Nerve Equations. *Int. J. Biomed. Comp.* **15**: 69-76.

[Hines 89] Michael Hines (1989). A Program for Simulation of Nerve Equations with Branching Geometries. *Int. J. Biomed. Comp.* **24**: 55-68.

[Hines 97] M. Hines and N. T. Carnevale (1997). The NEURON Simulation Environment. *Neural Computation* **9**: 1179–1209.

[Hodgkin 52] A.L. Hodgkin and A.F. Huxley (1952). A Quantitative Description of Membrane Current and Its Applications to Conduction and Excitation. *J. of Physiol.* **117**: 500-544.

[Honig 84] Michael L. Honig and David G. Messerschmitt (1984). *Adaptive Filters: Structures, Algorithms, and Applications.* Kluwer Academic Publishers.

[Johnston 95] Daniel Johnston and Samuel Miao-Sin Wu (1995). *Foundations of Cellular Neurophysiology.* The MIT Press.

[Kandel 95] Eric R. Kandel, James H. Schwartz, and Thomas M. Jessell (1995). *Essentials of Neural Science and Behavior.* Apple & Lange.

[Kauffman 93] Stuart A. Kauffman(1993). *The origins of Order—Self-Organization and Selection Evolution.* Oxford University Press, Inc.

[Kelso 95] J.A. Scott Kelso (1995). *Dynamic Patterns: the Self-Organization of Brain and Behavior.* The MIT Press.

[Koch 89] Christof Koch and Idan Segev, editor (1989). *Methods in Neuronal Modeling: From Synapses to Networks.* MIT Press.

[Koester 81] John Koester (1981). Functional Consequences of Passive Electrical Properties of the Neuron. In *Principles of Neural Science*, edited by Eric R. Kandel and James H. Schwartz, pp.44-52. Elsevier North Holland, Inc.

[Kuno 95] Motoy Kuno (1995). *The Synapse: Function, Plasticity, and Neurotrophism.* Oxford University Press.

[Levitan 97] Irwin B. Levitan and Leonard K. Kaczmarek (1997). *The Neuron: Cell and Molecular Biology*, second edition. Oxford University Press.

[Levy 90] William B. Levy, Costa M. Colbert, and Nancy L. Desmond (1990). Elemental Adaptive Processes of Neurons. In *Neuroscience and Connectionist Theory*, edited by Mark A. Gluck and David E. Rumelhart, pp.187-235. Lawrence Erlbaum Associates, Inc.

[Lund 78] R.D. Lund (1978). *Development and Plasticity of the Brain.* Oxford University Press.

[Lytton 96] William W. Lytton (1996). Optimizing Synaptic Conductance Calculation for Network Simulations. *Neural Computation* 8:501-510.

[MacGregor 93] Ronald J. MacGregor (1993). *Theoretical Mechanics of Biological Neural Networks.* Academic Press, Inc.

[MacLennan 90] Bruce MacLennan (1990). Field Computation: A Theoretical Framework for Massively Parallel Analog Computation, Parts I-IV. *Technical*

Report CS-90-100, Department of Computer Science, University of Tennessee, Knoxville, TN 37996.

[MacLennan 91] Bruce MacLennan (1991). Gabor Representations of Spatiotemporal Visual Images. *Technical Report CS-91-144,* Department of Computer Science, University of Tennessee, Knoxville, TN 37996.

[MacLennan 92] Bruce MacLennan (1992). Information Processing in the Dendritic Net, invited contribution, *Rethinking Neural Networks: Quantum Fields and Biological Data,* edited by Karl Pribram, 1993, pp. 161-197. Hillsdale, NJ: Lawrence Erlbaum.

Also: *Technical Report CS-92-180,* Department of Computer Science, University of Tennessee, Knoxville, TN 37996.

[MacLennan 93] Bruce MacLennan (1993). Field Computation in the Brain, invited contribution, *Rethinking Neural Networks: Quantum Fields and Biological Data,* edited by Karl Pribram, pp. 199-232. Hillsdale, NJ: Lawrence Erlbaum.

Also: *Technical Report CS-92-174,* Department of Computer Science, University of Tennessee, Knoxville, TN 37996.

[MacLennan 94] Bruce MacLennan (1994). Continuous Computation and the Emergence of the Discrete, invited contribution, *Origins: Brain & Self-*

Organization, edited by Karl Pribram, pp. 121-151. Hillsdale, NJ: Lawrence Erlbaum.

[Mohanty 87] Nirode Mohanty (1987). *Signal Processing.* Van Nostrand Reinhold Company, Inc.

[Rorabaugh 96] C. Britton Rorabaugh (1999). *DSP Primer.* The McGraw-Hill Companies.

[Rugh 96] Wilson J. Rugh (1996). *Linear System Theory,* second edition. Prentice-Hall, Inc.

[Pribram 91] Karl Pribram (1991). *Brain and Perception : Holonomy and Structure in Figural Processing.* Lawrence Erlbaum Associates, Inc.

[Schwartz 90] Eric L. Schwartz, editor (1990). *Computational Neuroscience.* A Bradford Book, The MIT Press.

[Shepherd 94] Gordon M. Shepherd (1994). *Neurobiology,* third edition. Oxford University Press.

[Sz.-Nagy 65] Béla Sz.-Nagy (1965). *Introduction to Real Functions and Orthogonal Expansions.* Oxford University Press.

[Tuckwell 88a] Henry C. Tuckwell (1988). *Introduction to Theoretical Neurobiology, Volume I: Linear Cable Theory and Dendritic Structure.* Cambridge University Press.

[Tuckwell 88b] Henry C. Tuckwell (1988). *Introduction to Theoretical Neurobiology, Volume II: Nonlinear and Stochastic Theories.* Cambridge University Press.

[Widrow 85] Bernard Widrow and Samuel D. Stearms (1985). *Adaptive Signal Processing.* Prentice-Hall, Inc.

Appendices

Appendix A

Implementation of the Linear

Synapse Model

The following lists the program to implement the linear synapse model. It can be

compiled with NEURON to get a special version of executable file of NEURON.

```
TITLE Linear Synapse

COMMENT

This file implements the linear synapse model.

For each time slice dt

1. Subtract the resting potential from the presynaptic electrical

   potential and then multiply this resulting number by the

   multiplier.  Add the resting potential back to it to get a final

   value intended to be the postsynaptic electrical potential.
```

2. Copy the final value in the above to the FIFO ring delay buffer.

   At the same time the buffer is shifted one position.

   The length of the buffer multiplied by the time slice dt is the

   delay of the synapse.

3. The postsynaptic electrical potential is the output of the delay

   buffer just being shifted out at the time slice preceding the

   current time slice.

ENDCOMMENT

INDEPENDENT {t FROM 0 TO 20000 WITH 1 (ms)}

: The following two variables have to be the same.

DEFINE MAX_DELAY 50000

VERBATIM

#define MAX_DELAY_IN_C 50000

ENDVERBATIM

NEURON {

        POINT_PROCESS linear_synapse

        RANGE g_syn, e, i, multiplier, raw_delay

        POINTER pre

        NONSPECIFIC_CURRENT i

        : GLOBAL

}

```
UNITS {

        (nA)   = (nanoamp)

        (mV)   = (millivolt)

        (umho) = (micromho)

}

ASSIGNED {

        : pointer to presynaptic variable

        pre

        delay_v_buffer[MAX_DELAY] (mV)   : Keep the voltages in this

                                         : array over time. Then

                                         : feed into the postsynaptic

                                         : compartment.

        new_v (mV)

        v_syn (mV)

        current_ring_head_position       : Put new data at this

                                         : location.

        current_ring_tail_position       : Get data from this

                                         : location.

        i (nA)

}

PARAMETER {
```

```
        multiplier

        : Synapse conductance

        g_syn       (umho)


        : Delay of synapse.

        raw_delay  (ms)

        : Resting potential.

        e           (mV)

        : The voltage at the postsynaptic compartment.

        : v          (mV)

        : Length of time slice, a global variable.

        : dt         (ms)
}

BREAKPOINT

{

     SOLVE ring

     VERBATIM

       /*

         Will do nothing before data moves to the postsynaptic

         side.

       */
```

```
        if (t > raw_delay) { i = g_syn*(v - v_syn); }

    ENDVERBATIM

}

PROCEDURE ring()

{

VERBATIM

  if (t <= dt)

  {

    (int) current_ring_head_position = 0;

    (int) current_ring_tail_position = 0;

  }

  new_v = (pre-e)*multiplier+e;

  if (raw_delay > dt)

  {

    if (t <= raw_delay)

    {

      delay_v_buffer[(int) current_ring_head_position] = new_v;

      (int) current_ring_head_position

      = get_next_ring_position((int) current_ring_head_position);

    }

    else
```

```c
  {

    v_syn = delay_v_buffer[(int) current_ring_tail_position];

    delay_v_buffer[(int) current_ring_head_position] = new_v;

    (int) current_ring_head_position

    = get_next_ring_position((int) current_ring_head_position);

    (int) current_ring_tail_position

    = get_next_ring_position((int) current_ring_tail_position);

  }

}

else

{

  v_syn = new_v;

}

/*  Print out data on the ring buffer

printf("%i %f %i %f\n",

  (int) current_ring_head_position,

  delay_v_buffer[(int) current_ring_head_position],

  (int) current_ring_tail_position,

  delay_v_buffer[(int) current_ring_tail_position]);

*/
```

ENDVERBATIM

```
}

VERBATIM

int get_next_ring_position(int array_index)

{

  if (array_index < MAX_DELAY_IN_C)

  {

    return (array_index + 1);

  }

  else

  {

    return (1);

  }

}

ENDVERBATIM
```

Appendix B

Patterns in the Normalized Input Set

The following figures show the 20 patterns in the normalized input set. Each pattern consists of 50 steps of current segments. The total time of a pattern is 2 milliseconds. Each current step runs for 0.04 milliseconds.
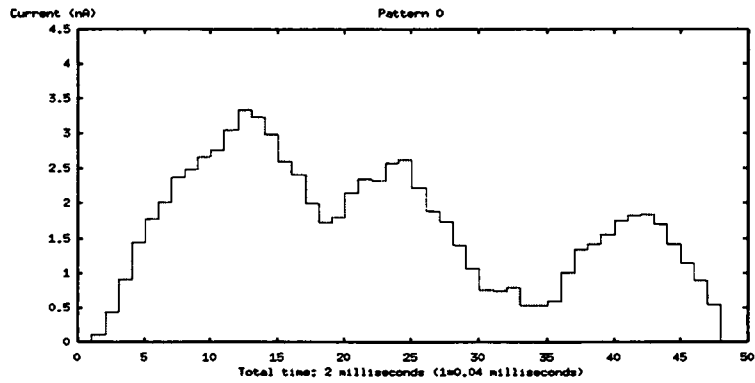
Figure B.1: Input pattern 0
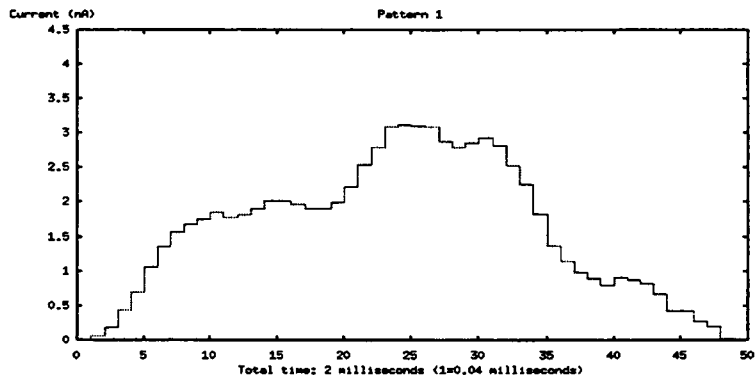
Figure B.2: Input pattern 1
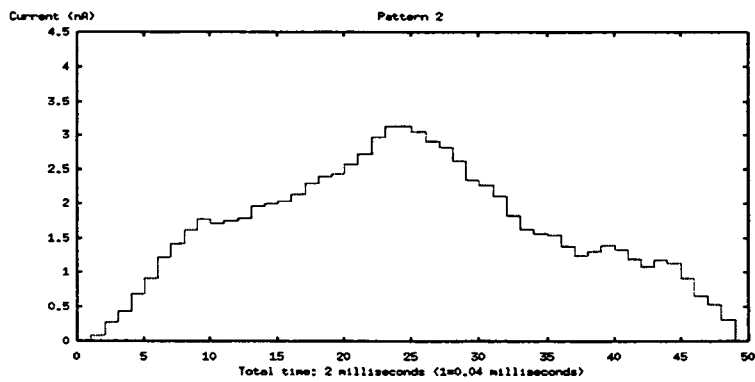
Figure B.3: Input pattern 2

119
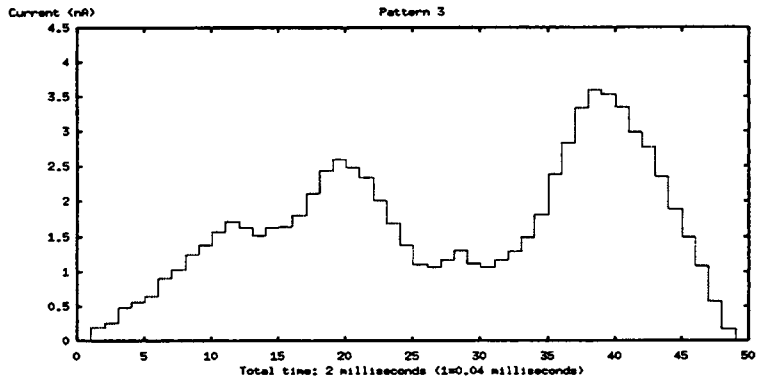
Figure B.4: Input pattern 3
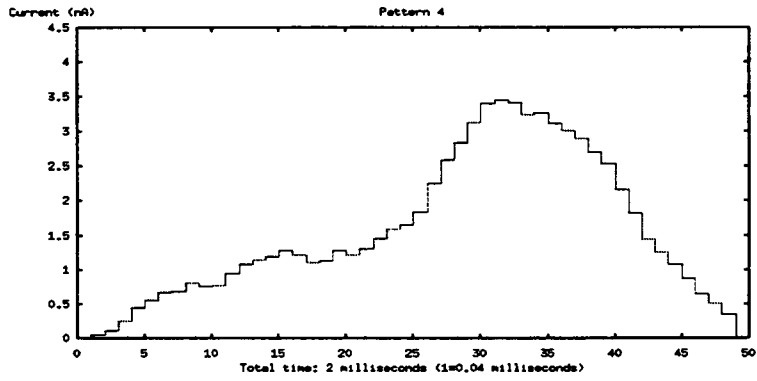
Figure B.5: Input pattern 4
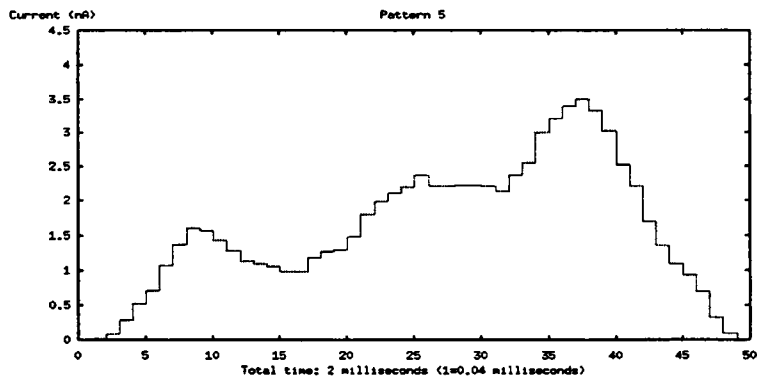
Figure B.6: Input pattern 5

120
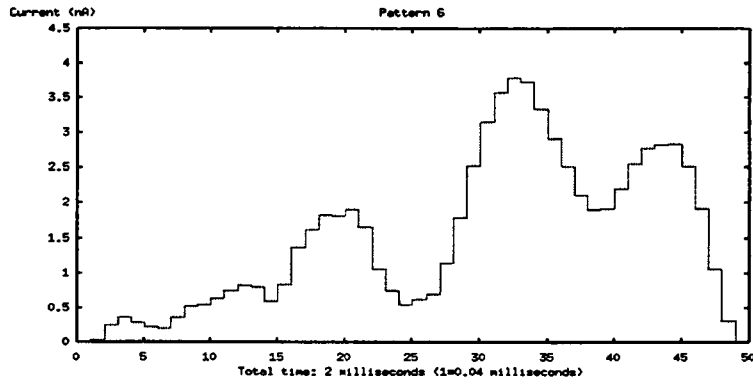
Figure B.7: Input pattern 6
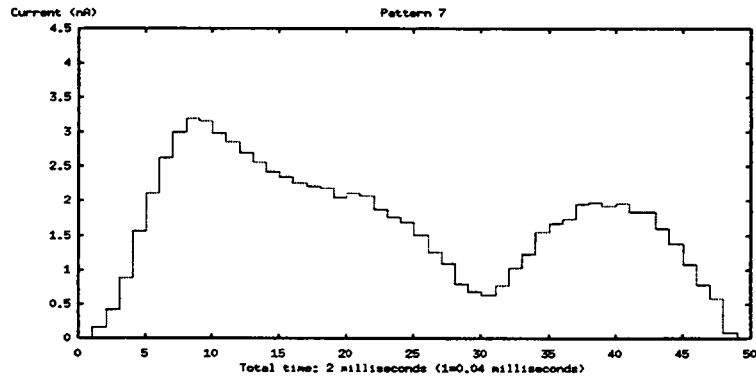
Figure B.8: Input pattern 7
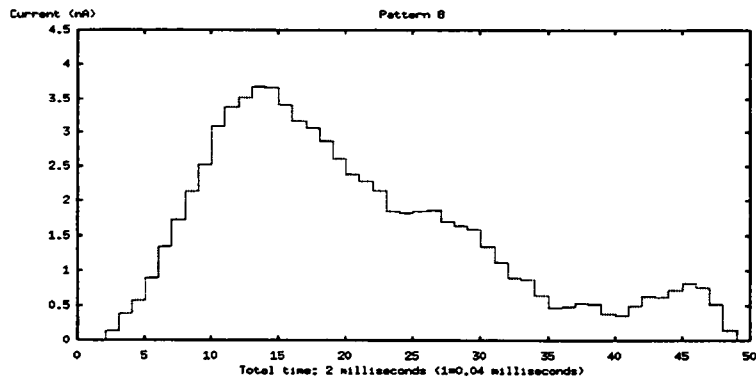
Figure B.9: Input pattern 8

121
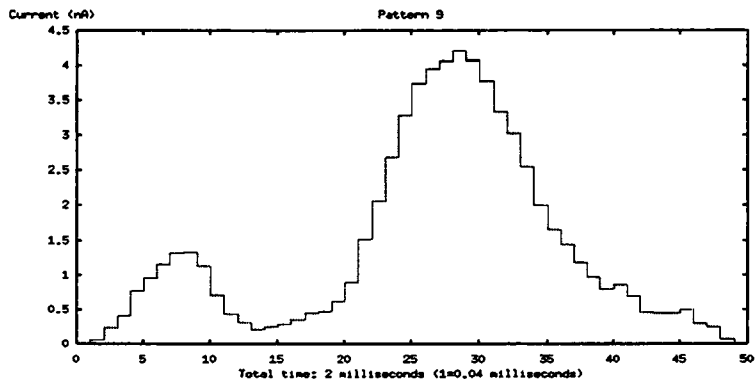
Figure B.10: Input pattern 9
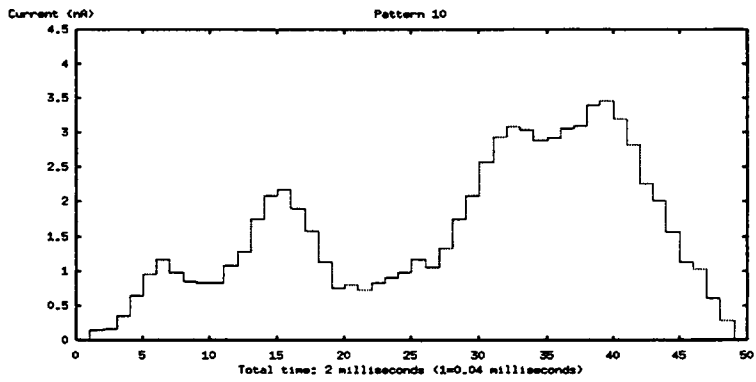
Figure B.11: Input pattern 10
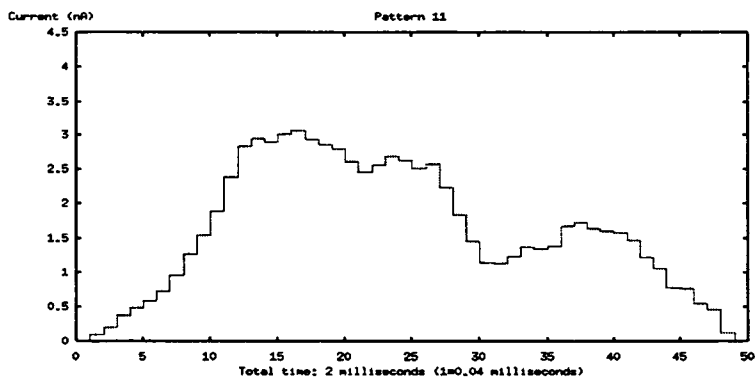
Figure B.12: Input pattern 11

122
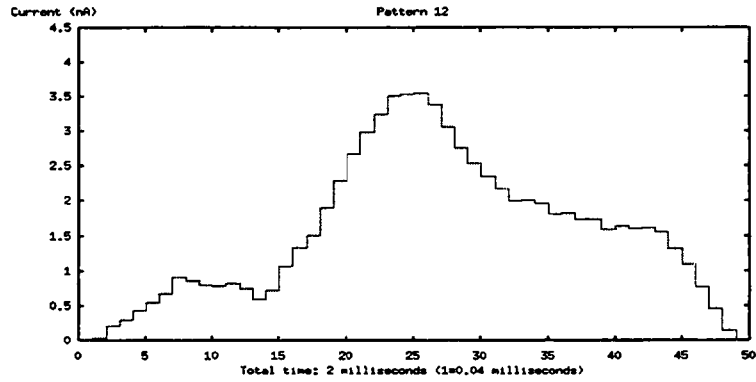
Figure B.13: Input pattern 12
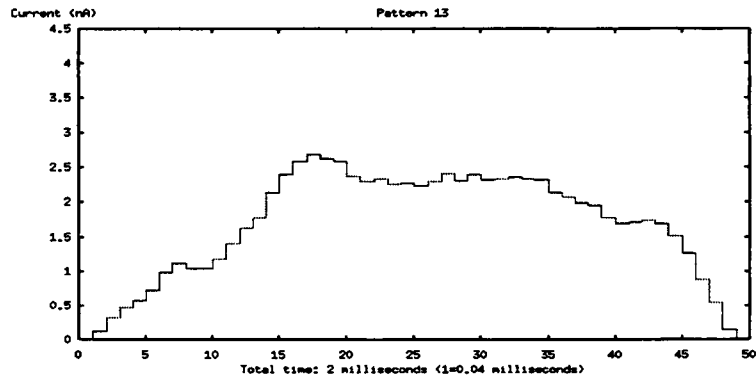
Figure B.14: Input pattern 13
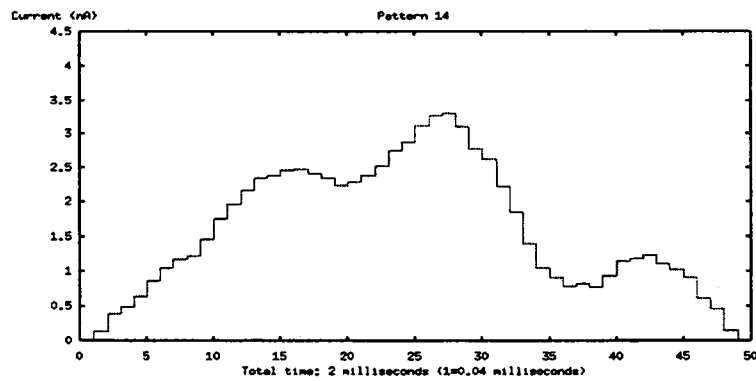
Figure B.15: Input pattern 14

123
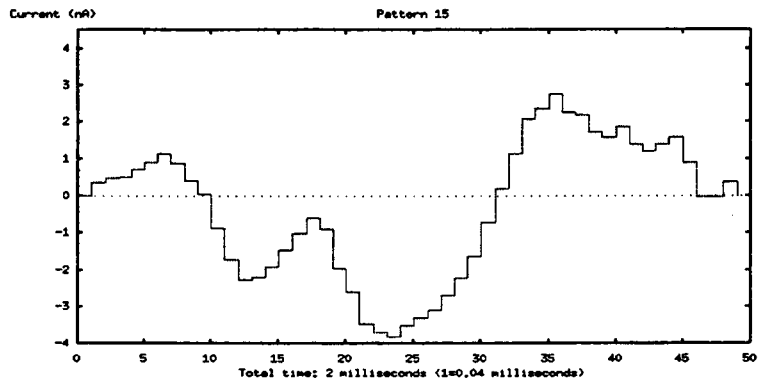
Figure B.16: Input pattern 15
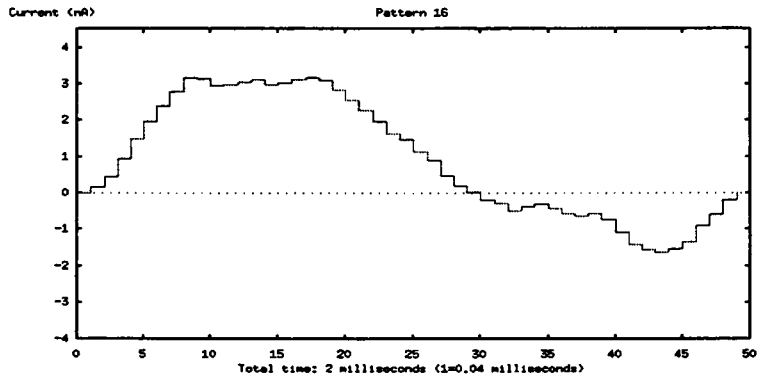
Figure B.17: Input pattern 16
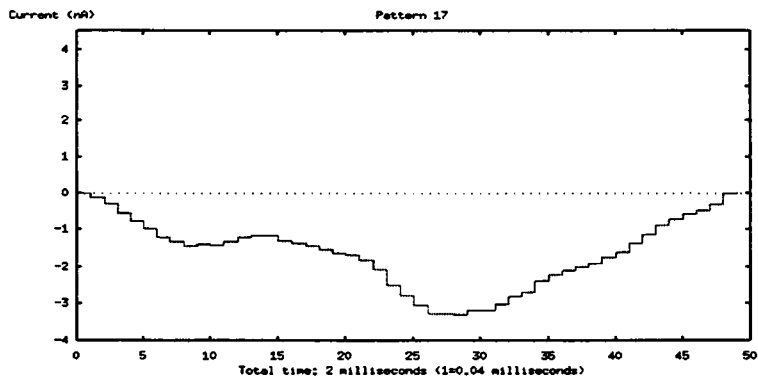
Figure B.18: Input pattern 17

124
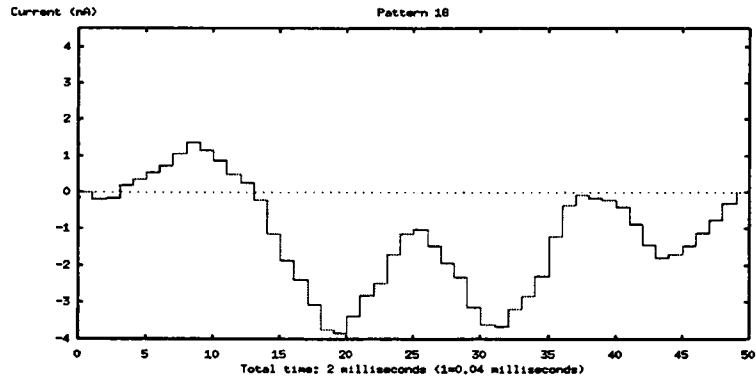
Figure B.19: Input pattern 18
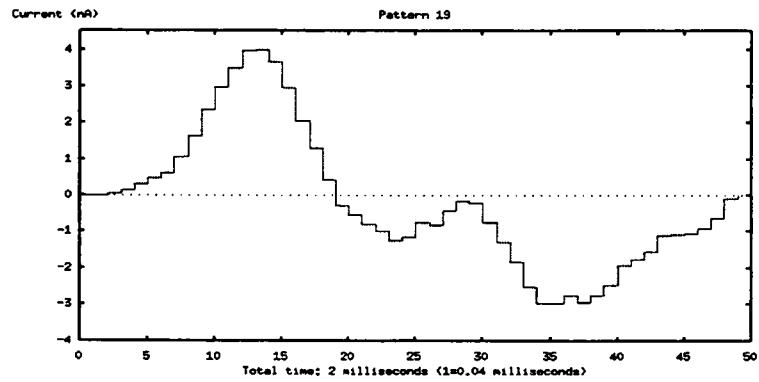
Figure B.20: Input pattern 19

125

Appendix C

Results of Differentiating Between Patterns by Random Dendritic Nets

The parameters used for dendritic net simulations (see Section 4.5.2) are listed in Table C.1, and the results of Section 4.5.3 are summarized in the Table C.2 - C.9. Table C.10 shows the ranking of the results in Section 4.5.4.

Table C.1: Parameters for random dendritic net simulations

Parameter	Value	Unit
Total time for simulations	2	$msec$
Time slice	0.0005	$msec$
Number of levels in the dendritic trees	4	
Branching angle of tree branches	$0 - 60$	$degree$
The number of synapses in a subnet	12-15	
Direction of synapses	forward	
First level branch length of dendritic trees	0-60	$micron$
Next level branch factor	.80	
Diameter of dendrites	3	$micron$
Resting potential	-65	mV
Axial resistance	35.4	$\Omega - cm$
Membrane capacitance	0.8	$\mu F/cm^2$
Synapse amplification factor	3	
Synapse delay	0-2	$msec$
g_syn for linear synapse	0.5	

Table C.2: Correlation for the random dendritic net with capacitance $= 0.8$ μF, conductance $= 0.001$ S/cm^2, and 5 subnets

Run	Correlation
1	0.7152
2	0.9788
3	0.9443
4	0.9452
5	0.9704
6	0.9489
7	0.9669
8	0.9696
Mean	0.9300
Standard deviation	0.0878

Table C.3: Correlation for the random dendritic net with capacitance = 0.8 μF, conductance = 0.001 S/cm^2, and 10 subnets

Run	*Correlation*
1	0.9802
2	0.9932
3	0.9516
4	0.9706
5	0.9507
6	0.6238
7	0.9717
8	0.9777
Mean	0.9274
Standard deviation	0.1234

Table C.4: Correlation for the random dendritic net with capacitance = 0.8 μF, conductance = 0.001 S/cm^2, and 15 subnets

Run	*Correlation*
1	0.9768
2	0.9621
3	0.9774
4	0.9681
5	0.9866
6	0.9846
7	0.9648
8	0.9684
Mean	0.9736
Standard deviation	0.0091

Table C.5: Correlation for the random dendritic net with capacitance $= 0.8~\mu F$, conductance $= 0.005~S/cm^2$, and 5 subnets

Run	Correlation
1	0.6755
2	0.8163
3	0.8966
4	0.9106
5	0.6755
6	0.8981
7	0.9800
8	0.9578
Mean	0.8513
Standard deviation	0.1187

Table C.6: Correlation for the random dendritic net with capacitance $= 0.8~\mu F$, conductance $= 0.005~S/cm^2$, and 10 subnets

Run	Correlation
1	0.9791
2	0.7201
3	0.9791
4	0.9618
5	0.9596
6	0.9728
7	0.8873
8	0.9721
Mean	0.9290
Standard deviation	0.0896

Table C.7: Correlation for the random dendritic net with capacitance = 0.8 μF, conductance = 0.005 S/cm^2, and 15 subnets

Run	Correlation
1	0.9781
2	0.9385
3	0.9814
4	0.8731
5	0.9780
6	0.9145
7	0.9145
8	0.9927
Mean	0.9463
Standard deviation	0.0428

Table C.8: Correlation for the random dendritic net with capacitance = 0.8 μF, conductance = 0.0002 S/cm^2, and 10 subnets

Run	Correlation
1	-0.4766
2	0.5321
3	0.3799
4	0.5528
5	0.4553
6	0.4430
7	0.1001
8	-0.5228

Table C.9: Summary of the performances of the random dendritic nets

Membrane conductance	Membrane capacitance	Number of subnets	Mean of correlations	Standard deviation	Confidence interval of 99% for the mean of any sample of 8 correlations from 8 random dendritic nets
$0.001S/\text{cm}^2$	$0.8\mu F$	5	0.9300	0.0878	$0.8214 \leq \mu \leq 1.0000$
$0.001S/\text{cm}^2$	$0.8\mu F$	10	0.9274	0.1235	$0.7746 \leq \mu \leq 1.0000$
$0.001S/\text{cm}^2$	$0.8\mu F$	15	0.9736	0.0091	$0.9623 \leq \mu \leq 0.9849$
$0.005S/\text{cm}^2$	$0.8\mu F$	5	0.8513	0.1187	$0.7044 \leq \mu \leq 0.9982$
$0.005S/\text{cm}^2$	$0.8\mu F$	10	0.9290	0.0896	$0.8181 \leq \mu \leq 1.0000$
$0.005S/\text{cm}^2$	$0.8\mu F$	15	0.9463	0.0428	$0.8933 \leq \mu \leq 0.9994$

Table C.10: Ranks of maximum values and final values for a dendritic net with 10 random subnets and membrane conductance 0.005 S/cm^2

Pattern number	Rank for maximum value	Rank for final value
0	1	2
1	3	4
2	3	4
3	16	1
4	12	1
5	13	2
6	13	1
7	4	1
8	5	2
9	11	2
10	15	2
11	7	1
12	7	5
13	8	2
14	1	8
15	18	1
16	1	1
17	1	1
18	2	2
19	1	1

Appendix D

Results of Simulating a Dendritic

Net with 10 Random Subnets

The following shows two groups of figures, where one group includes all the inner products among the patterns in the input set and the other group contains the final values of the output patterns of a dendritic net with 10 random subnets and membrane conductance 0.005 S/cm^2 when taking the patterns in the input set as the inputs (see section 4.5.4 for details). The inner product of two patterns is done by summing over all the multiplications of the 50 corresponding current steps of the two current patterns.

Each figure of inner products is followed by a figure of final values. These two corresponding figures look similar. This means that this random dendritic net performs well in differentiating between patterns in the input set (see Appendix
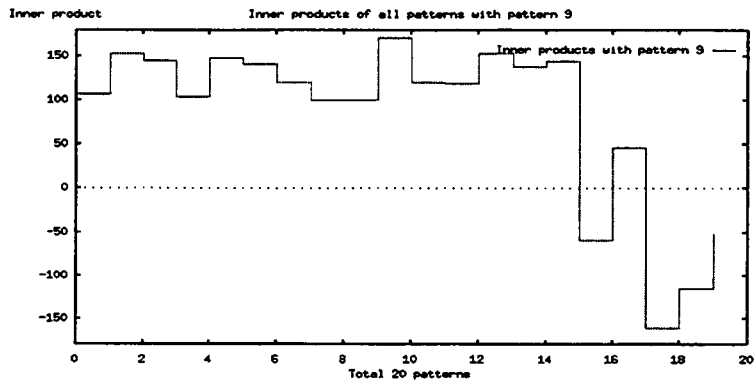
133

B).
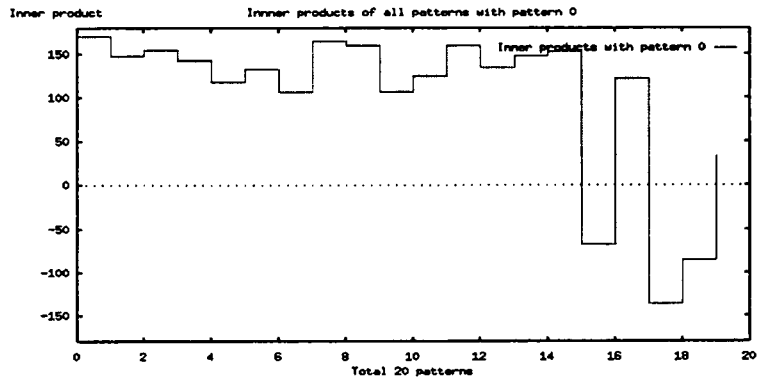
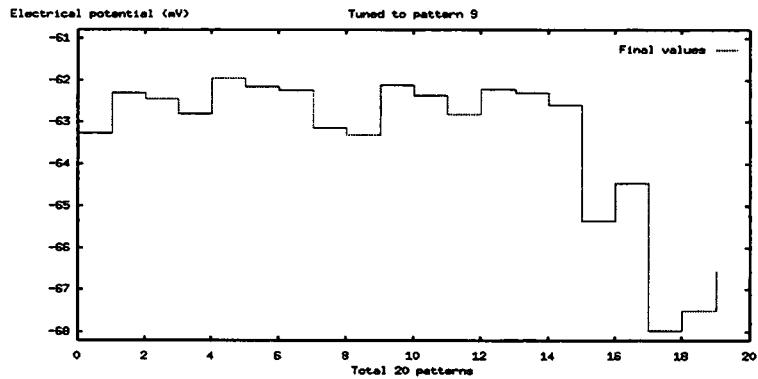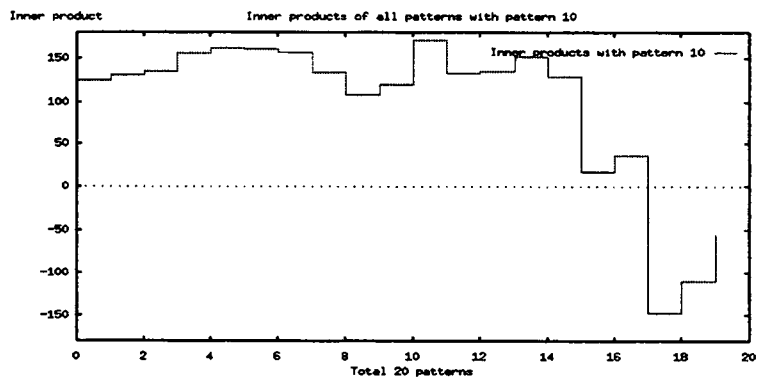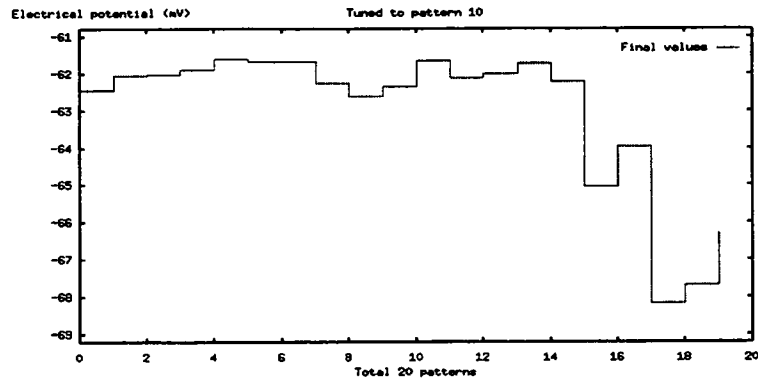Figure D.1: Cross inner products of all patterns with pattern 0
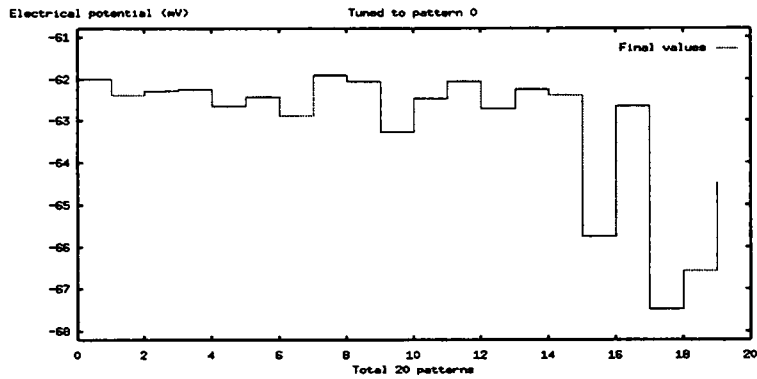
Figure D.2: Final values of the dendritic net tuned to pattern 0 for all input patterns
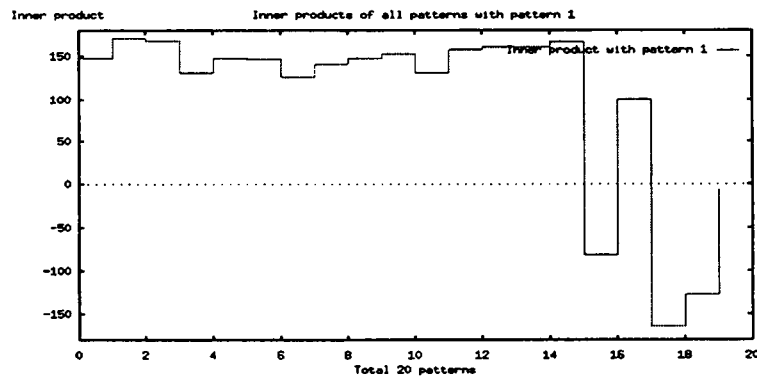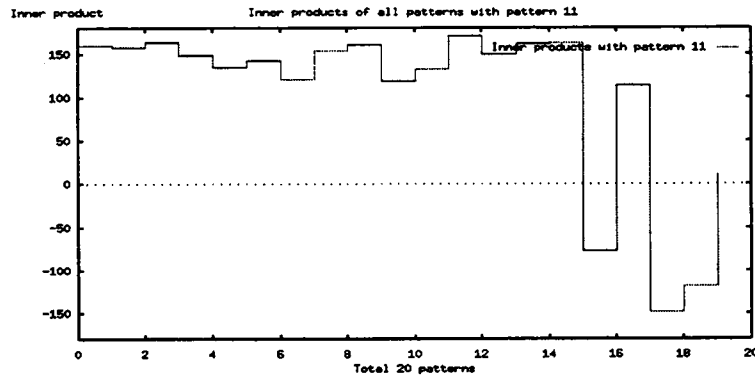

Figure D.3: Cross inner products of all patterns with pattern 1

135
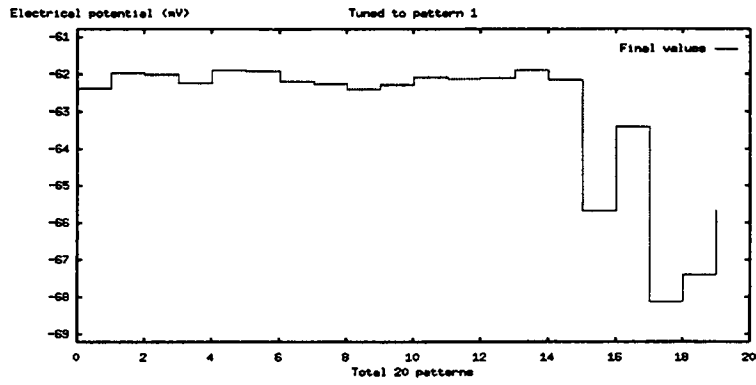
Figure D.4: Final values of the dendritic net tuned to pattern 1 for all input patterns
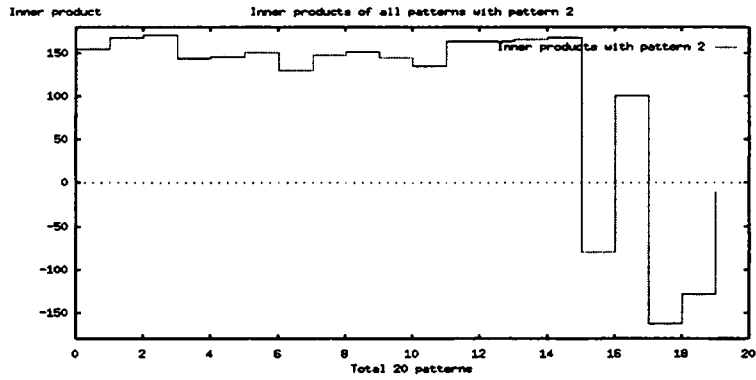
Figure D.5: Cross inner products of all patterns with pattern 2
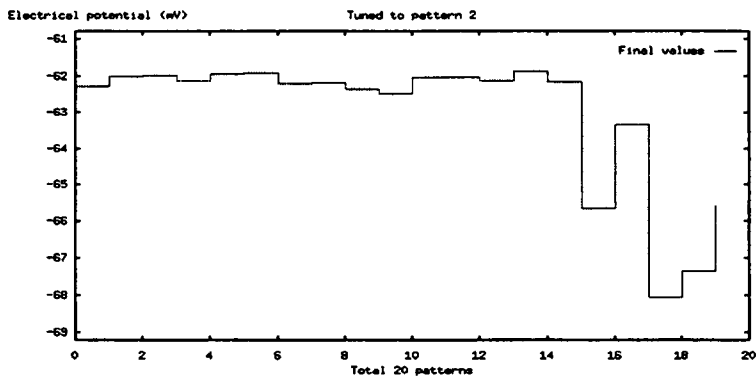
Figure D.6: Final values of the dendritic net tuned to pattern 2 for all input patterns

136
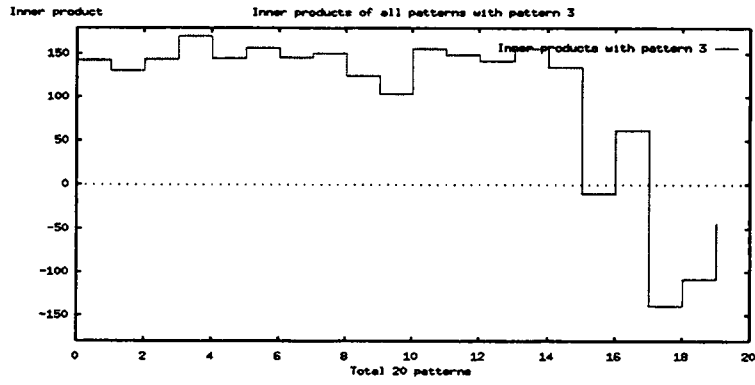
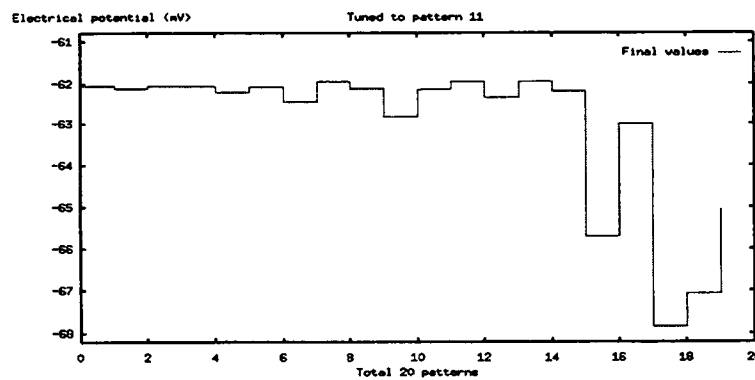Figure D.7: Cross inner products of all patterns with pattern 3
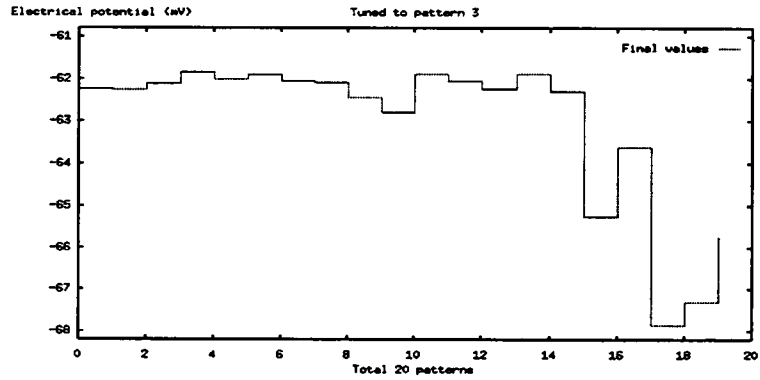
Figure D.8: Final values of the dendritic net tuned to pattern 3 for all input patterns
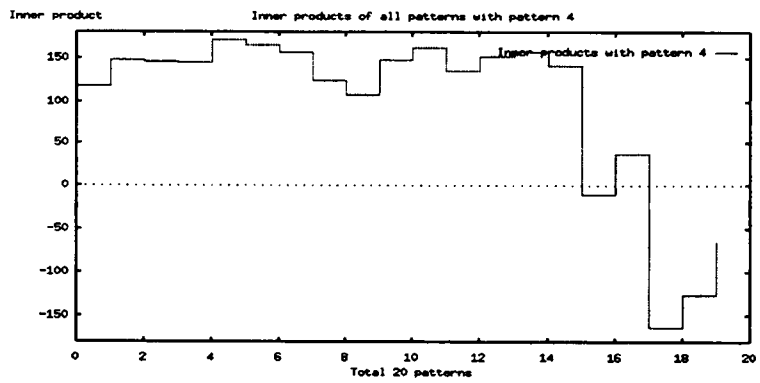
Figure D.9: Cross inner products of all patterns with pattern 4

137
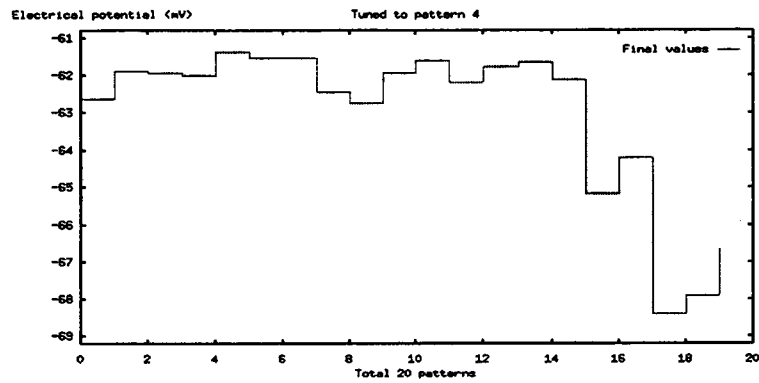
Figure D.10: Final values of the dendritic net tuned to pattern 4 for all input patterns
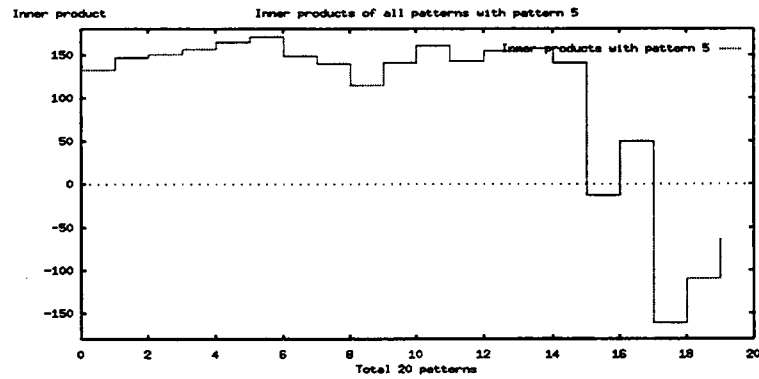
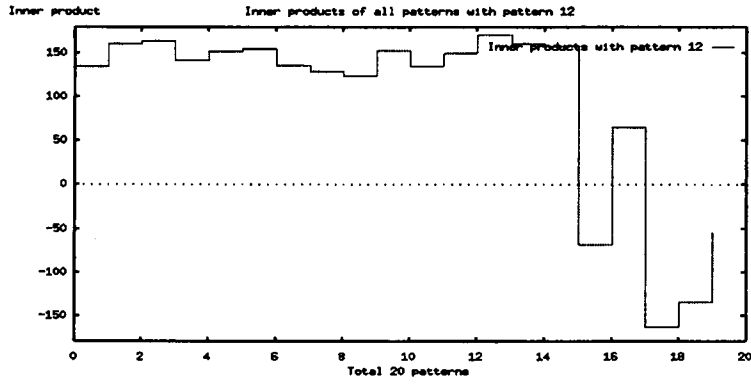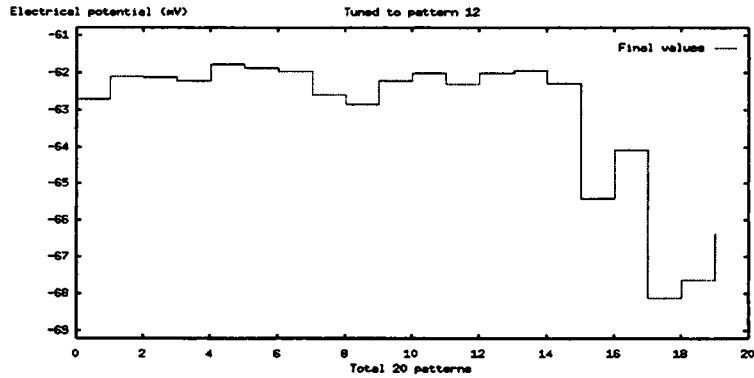Figure D.11: Cross inner products of all patterns with pattern 5
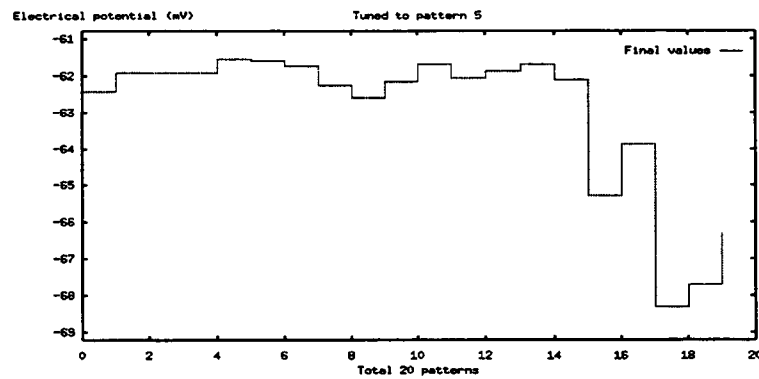
Figure D.12: Final values of the dendritic net tuned to pattern 5 for all input patterns

138
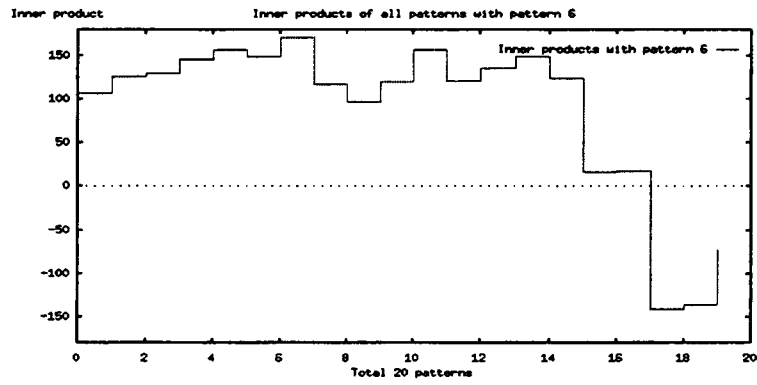
Figure D.13: Cross inner products of all patterns with pattern 6
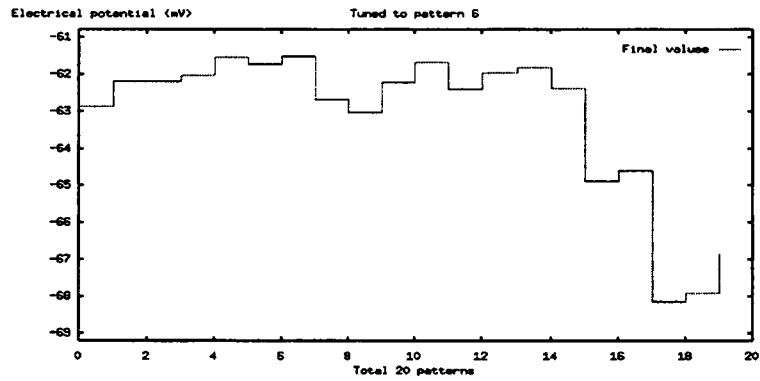
Figure D.14: Final values of the dendritic net tuned to pattern 6 for all input patterns
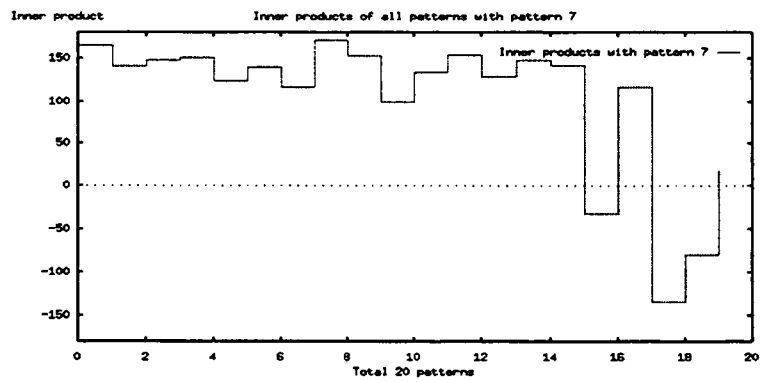
Figure D.15: Cross inner products of all patterns with pattern 7

139
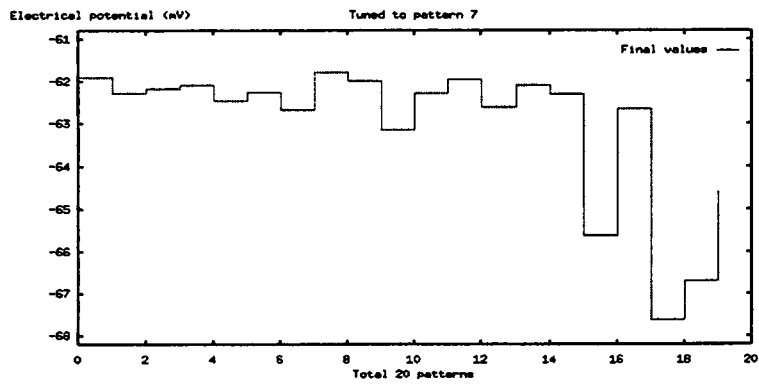
Figure D.16: Final values of the dendritic net tuned to pattern 7 for all input patterns
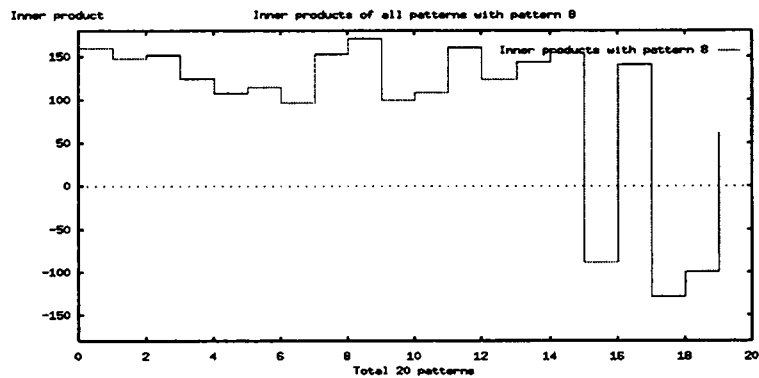
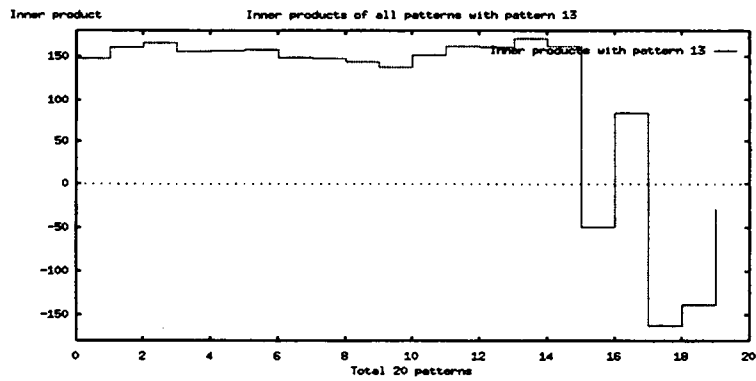

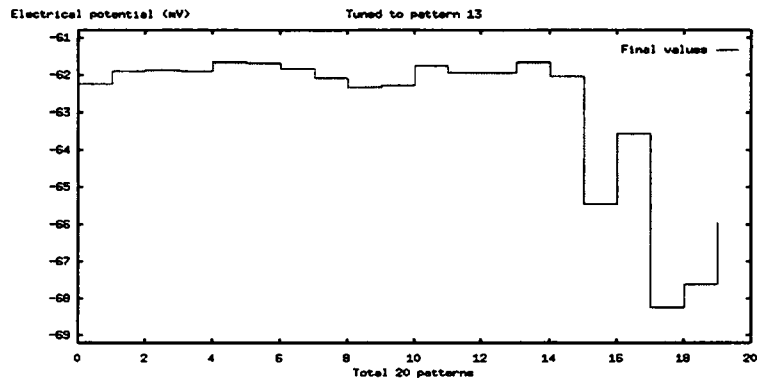Figure D.17: Cross inner products of all patterns with pattern 8
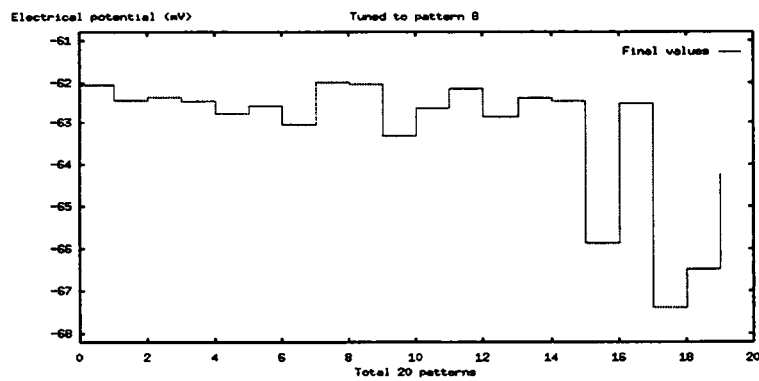
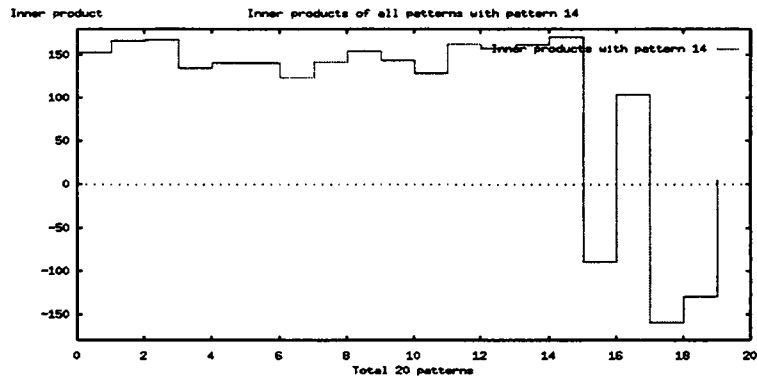Figure D.18: Final values of the dendritic net tuned to pattern 8 for all input patterns

140

Figure D.19: Cross inner products of all patterns with pattern 9

Figure D.20: Final values of the dendritic net tuned to pattern 9 for all input patterns

Figure D.21: Cross inner products of all patterns with pattern 10

141

Figure D.22: Final values of the dendritic net tuned to pattern 10 for all input patterns

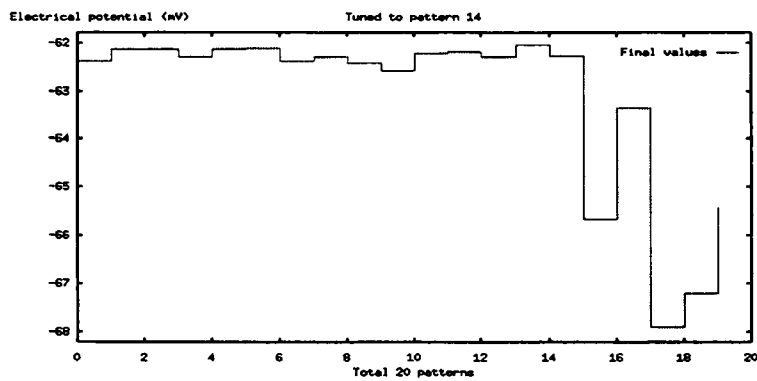Figure D.23: Cross inner products of all patterns with pattern 11

Figure D.24: Final values of the dendritic net tuned to pattern 11 for all input patterns

Figure D.25: Cross inner products of all patterns with pattern 12

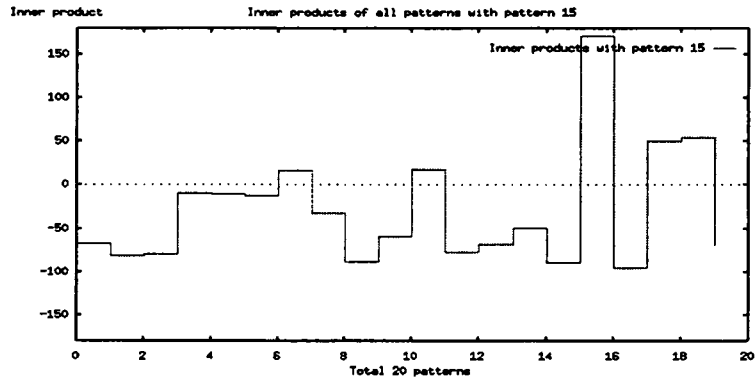Figure D.26: Final values of the dendritic net tuned to pattern 12 for all input patterns

Figure D.27: Cross inner products of all patterns with pattern 13
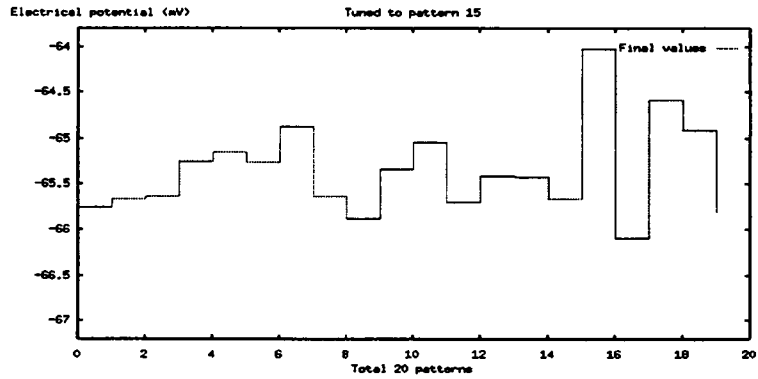
Figure D.28: Final values of the dendritic net tuned to pattern 13 for all input patterns

Figure D.29: Cross inner products of all patterns with pattern 14

Figure D.30: Final values of the dendritic net tuned to pattern 14 for all input patterns

144

Figure D.31: Cross inner products of all patterns with pattern 15

Figure D.32: Final values of the dendritic net tuned to pattern 15 for all input patterns
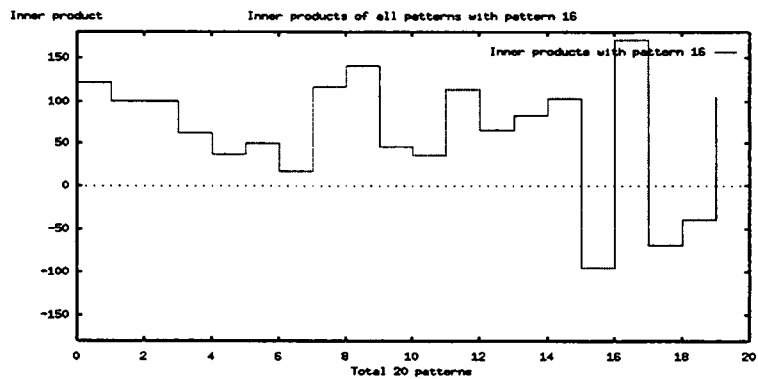

Figure D.33: Cross inner products of all patterns with pattern 16
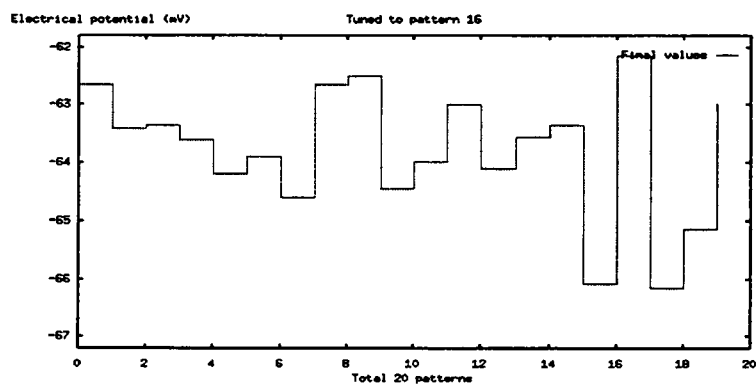
145

Figure D.34: Final values of the dendritic net tuned to pattern 16 for all input patterns

Figure D.35: Cross inner products of all patterns with pattern 17

Figure D.36: Final values of the dendritic net tuned to pattern 17 for all input patterns
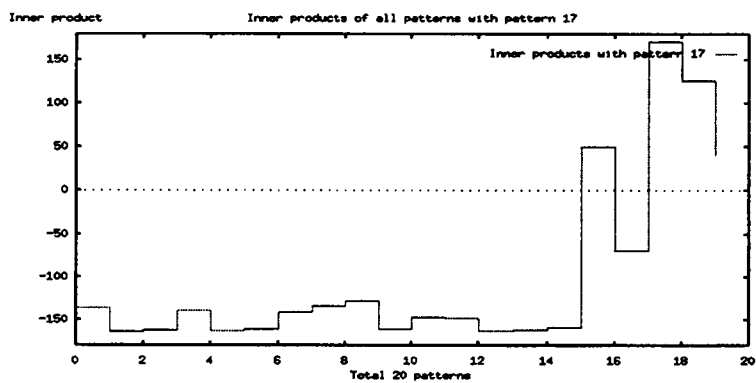
146

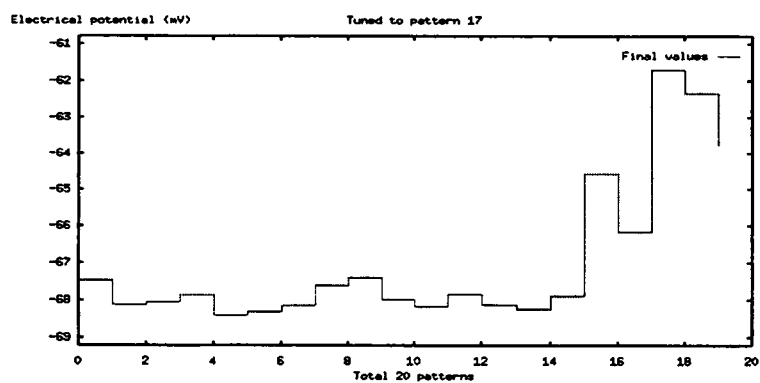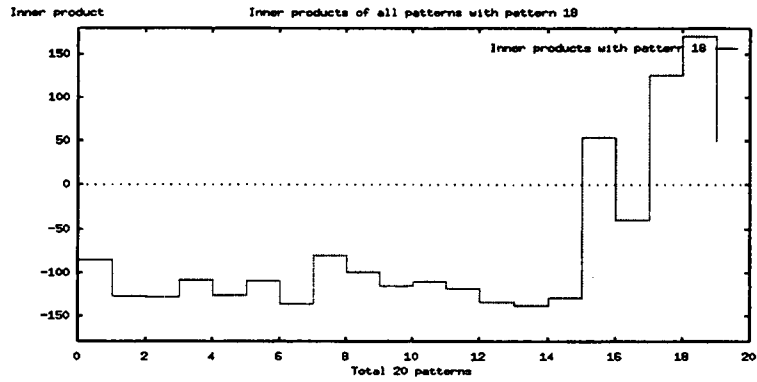Figure D.37: Cross inner products of all patterns with pattern 18
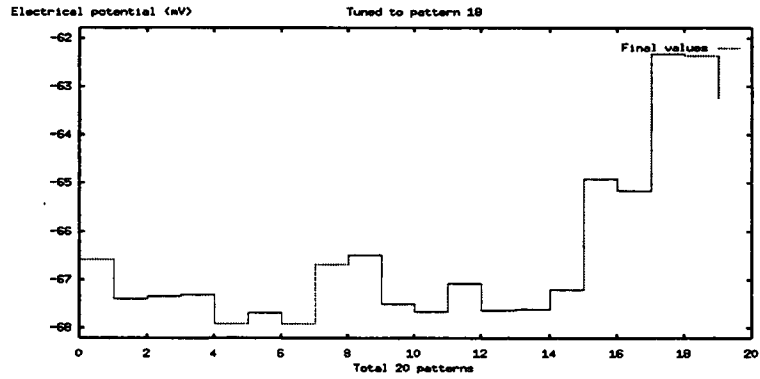
Figure D.38: Final values of the dendritic net tuned to pattern 18 for all input patterns
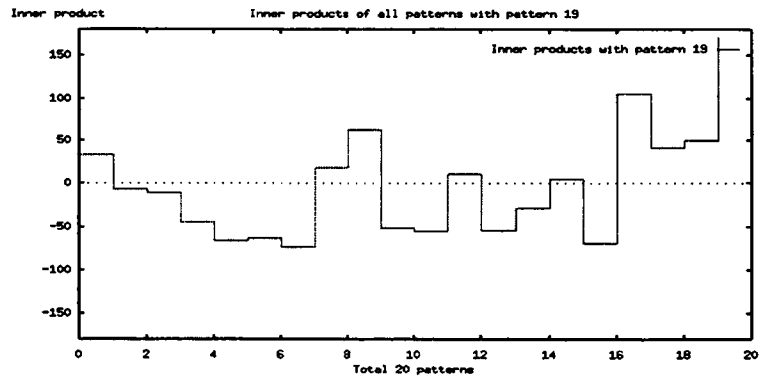

Figure D.39: Cross inner products of all patterns with pattern 19
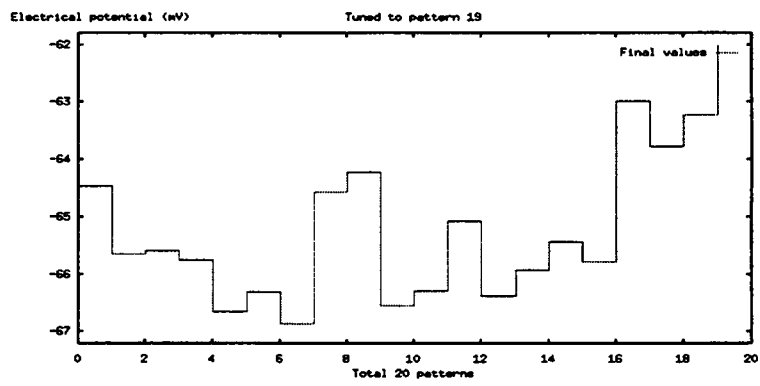
147

Figure D.40: Final values of the dendritic net tuned to pattern 19 for all input patterns

Appendix E

Results of Testing Fault-tolerance

of Random Dendritic Nets

The results of Section 4.5.5 are summarized in the following tables. For convenience, Table C.4 is duplicated here as Table E.3.

Table E.1: Correlation for the random dendritic net with capacitance $= 0.8\ \mu F$, conductance $= 0.001\ S/\text{cm}^2$, and 5 subnets

Run	Correlation
1	0.968713
2	0.945224
3	0.977928
4	0.957361
5	0.927237
6	0.981954
7	0.960796
8	0.955227
Mean	0.9593
Standard deviation	0.0177

Table E.2: Correlation for the random dendritic net with capacitance $= 0.8\ \mu F$, conductance $= 0.001\ S/\text{cm}^2$, and 10 subnets

Run	Correlation
1	0.9777
2	0.9247
3	0.9831
4	0.9672
5	0.9694
6	0.9807
7	0.9468
8	0.9777
Mean	0.9659
Standard deviation	0.0203

Table E.3: Correlation for the random dendritic net with capacitance $= 0.8\ \mu F$, conductance $= 0.001\ S/\mathrm{cm}^2$, and 15 subnets

Run	Correlation
1	0.9768
2	0.9621
3	0.9774
4	0.9681
5	0.9866
6	0.9846
7	0.9648
8	0.9684
Mean	0.9736
Standard deviation	0.0091

Table E.4: Summary of the performances of the random dendritic nets

Membrane conductance	Membrane capacitance	Number of subnets	Mean of correlations	Standard deviation	Confidence interval of 99% for the mean of any sample of 8 correlations from 8 random dendritic nets
$0.001 S/\mathrm{cm}^2$	$0.8\mu F$	5	0.9593	0.0177	$0.9374 \leq \mu \leq 0.9812$
$0.001 S/\mathrm{cm}^2$	$0.8\mu F$	10	0.9659	0.0203	$0.9409 \leq \mu \leq 0.9910$
$0.001 S/\mathrm{cm}^2$	$0.8\mu F$	15	0.9736	0.0091	$0.9623 \leq \mu \leq 0.9849$

151

Table E.5: Correlation for the random dendritic net with capacitance = 0.8 μF, conductance = 0.001 S/cm^2, and 1 subnet

Run	Correlation
1	0.3580
2	0.7613
3	0.2943
4	-0.0278
5	0.7789
6	0.9593
7	0.4393
8	0.8368

Table E.6: Ranks of maximum values and final values for a dendritic net with 10 random subnets and membrane conductance 0.005 S/cm^2

Pattern number	Rank for maximum value	Rank for final value
0	1	2
1	3	4
2	3	4
3	16	1
4	12	1
5	13	2
6	13	1
7	4	1
8	5	2
9	11	2
10	15	2
11	7	1
12	7	5
13	8	2
14	1	8
15	18	1
16	1	1
17	1	1
18	2	2
19	1	1

Vita

Tzusheng Pei was born and raised in Taiwan. He received a BS in Physics and an MS in Nuclear Engineering from National Tsing Hua University, Hsinchu, Taiwan in 1979 and 1982 respectively. In 1986, he came to the University of Lowell (now University of Massachusetts at Lowell) for studies in computer science. He received an MS in Computer Science in 1988. In early 1989, he started to work on machine translation and natural language processing at Industrial Technology Research Institute, Hsinchu, Taiwan first as a software engineer and then as an Associate Researcher. In the Fall of 1991, he came to the University of Tennessee, Knoxville, to obtain a Ph.D. in Computer Science. From the summer of 1992 to the end of 1993, he worked as a research assistant on transportation projects for UT Transportation Research Center. He was a computer system analyst at Cancer Fund of America from early 1994 to the end of 1998.