COMPUTATIONAL STUDY ON SENSORY MOTOR INTEGRATION AND SENSORY
SEQUENCE LEARNING IN SONGBIRD HVC

A Dissertation in

Physics

by

Linli Wang

© 2009 Linli Wang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2009
The dissertation of Linli Wang was reviewed and approved* by the following:

Dezhe Jin  
Assistant Professor of Physics  
Dissertation Advisor  
Chair of Committee

John C. Collins  
Professor of Physics

Gong Chen  
Associate Professor of Biology

Jayanth R. Banavar  
Professor of Physics  
Distinguished Professor and Head of the Physics Department

*Signatures are on file in the Graduate School
ABSTRACT

Intracellular recording shows HVC_{RA} neurons in songbirds have strong selective response to the bird’s own song. This selectivity is especially sensitive to the sequential order of the syllables spanning over several hundred milliseconds. The firing pattern of HVC_{RA} is sparse in the sense that each HVC_{RA} neuron only fires at most once during one motif and has such precise timing that it always gives a burst of spikes lasting around 10 ms. Developed from a song production model using synfire chains, we incorporate the sensory input into the spike propagation process by coincident detection of sensory input from upstream and lateral input inside the synfire chain. The specific wiring structure from sensory input to the synfire chain embeds the sensory sequence which can be recognized by our model. This specific structure can be learned from initial all-to-all weak connections with a postsynaptic voltage dependent learning mechanism. The stability of this model is shown by the robust performances of both recognition and learning under different synaptic noises and various defects in training sequence. In order to handle repetitive inputs, Potassium A-current is induced into our model to check the current surge due to repetitive inputs. The working region and its effect on the sequence recognition of A-current are investigated and its limited role on the single compartment model is concluded. In order to construct a more biological model and leave more room for A-current, the two-compartment model of a single HVC_{RA} neuron replaces the single compartment model of the integrate-fire neuron. A more popular learning mechanism (spike-timing dependent plasticity) is applied to induce the specific wiring from sensory neurons to synfire chain neurons. The sequence recognition and learning of this revised network are studied. The enhanced robustness is partly due to the enlargement of the coincident region by A-current.
# TABLE OF CONTENTS

## LIST OF FIGURES

Chapter 1 Background review ................................................................. 1

1.1 Songbird vocal learning................................................................. 1
  1.1.1 Songbird vocal learning behavior ........................................... 1
  1.1.2 Neuron and cortex ............................................................... 3
  1.1.3 Song related nuclei and pathway........................................... 5
  1.1.4 Song related nuclei intracellular and extracellular recording .. 7

1.2 Synfire chain theory .................................................................. 14
  1.2.1 Premotor model based on synfire chair ............................... 19
  1.2.2 Formation of synfire chain structure ................................... 20

1.3 Synaptic plasticity: STDP and voltage dependent learning rules ...
  1.3.1 Spike timing dependent plasticity ......................................... 21
  1.3.2 Plasticity depending on postsynaptic voltage ....................... 26

1.4 Spatiotemporal sequence coding ................................................. 29
  1.4.1 Perceptron ........................................................................ 30
  1.4.2 Tempotron ......................................................................... 31
  1.4.3 Variance model .................................................................. 34
  1.4.4 Network model .................................................................. 36
  1.4.5 Liquid state machine .......................................................... 38
  1.4.6 Van Rossum distance model ................................................. 39

Chapter 2 Sequence recognition and sequence learning based on plateau potential .... 41

2.1 Introduction .............................................................................. 41
2.2 Methods .................................................................................... 42
  2.2.1 Network Model .................................................................. 42
  2.2.2 Neuron Model .................................................................... 45
  2.2.3 Synapse Model ................................................................... 50
  2.2.4 Recognition model ............................................................. 51
  2.2.5 Plasticity Model .................................................................. 54
  2.2.6 Noise Model ....................................................................... 56

2.3 Robust recognition of sensory sequence with a synfire chain ......... 58
  2.3.2 Noise tolerance for recognition ............................................ 63

2.4 Encoding from learning sequence to feedforward wiring ............ 65

2.5 Learning with synfire chain ....................................................... 67

2.6 Various compromised learning .................................................. 71
  2.6.1 Learning with random deletion ............................................ 71
  2.6.2 Learning with random replacement .................................... 73
  2.6.3 Learning with mixed sequences ......................................... 76
  2.6.4 Learning with homogeneous synaptic decay ....................... 79

2.7 Learning a very long sequence .................................................. 81
  2.7.1 Learning rate..................................................................... 85

2.8 Recognition with repeated sensory inputs and A-current ............ 89
  2.8.1 The violation of coincident detection by sequence with repetitive inputs ... 90
2.8.2 A-current ................................................................. 93
2.8.3 The effect and side effect applying A-current on integrate fire model ............................... 94
2.8.4 Reshape A-current to check current surge .................................................................... 95
2.8.5 Other possibilities about parameter choice ................................................................. 98
2.8.6 Match a realistic dynamic constant ............................................................................. 100
2.8.7 Repetitive lateral inputs and repetitive feedforward inputs ......................................... 104
2.8.8 Sequence of inputs with variable interval .................................................................... 107
2.8.9 Recognition and rejection ......................................................................................... 109
2.8.10 Parameters affect plateau potential ........................................................................... 110
2.8.11 Learning with repetitive sensory inputs ...................................................................... 116
2.9 Conclusion ....................................................................................................................... 117

Chapter 3 Sequence recognition and sequence learning based on two-compartment model .. 119

3.1 Introduction ...................................................................................................................... 119
3.2 Methods .......................................................................................................................... 120
  3.2.1 Front end model ........................................................................................................ 120
  3.2.2 Nif model .................................................................................................................. 130
  3.2.3 Network .................................................................................................................... 130
  3.2.4 Neuron Model .......................................................................................................... 132
  3.2.5 Synapse model ........................................................................................................ 136
  3.2.6 Recognition model ................................................................................................ 137
  3.2.7 Plasticity model ....................................................................................................... 137
  3.2.8 Noise model ............................................................................................................ 140
3.3 Single neuron firing probability ....................................................................................... 140
3.4 Coincidence detection phase diagram ............................................................................ 141
3.5 Sequence recognition ..................................................................................................... 144
  3.5.1 Synfire chain with neuron groups ........................................................................... 144
3.6 Sequence learning .......................................................................................................... 149
3.7 Learning with STDP algorithm and optimal learning .................................................... 152
  3.7.1 Two-Neuron Model ................................................................................................ 152
  3.7.2 Encoding from training sequence to sensory wiring ................................................. 154
  3.7.3 Learning Rate ......................................................................................................... 159
3.8 Conclusion ....................................................................................................................... 160

BIBLIOGRAPHY .................................................................................................................. 161
LIST OF FIGURES

Figure 1-1: The development of the zebra finch song system; adapted from \(^2\) ...................... 2

Figure 1-2: The basic structure of a single neuron; adapted from \(^5\) .............................. 5

Figure 1-3: The pathways of the song system of zebra finch; adapted from \(^6\) ..................... 6

Figure 1-4: The subtypes of HVC neurons and their characteristic response to injected current; The scale bar is 20\(\mu\)m. adapted from \(^11\). ................................................. 8

Figure 1-5: Comparison of selectivity regarding to BOS/REV/CON at HVC, NIf and field L when the bird is anesthetized; adapted from \(^18\) ................................................ 11

Figure 1-6: Top panel: auditory response of NIf and HVC to different stimuli when the bird is anesthetized. Bottom panel: The auditory response of HVC due to inactivation of NIf (applying GABA on NIf); adapted from \(^19\) ...................................... 12

Figure 1-7: An example of synfire chain network with 3 groups (ovals) each of which consists of 3 neurons (big circles). The connections (gray lines) between adjacent groups are all-to-all. The small circles mean the connections are all excitatory............. 14

Figure 1-8: Top graph: Flow chart of spike activity and dispersion along the synfire chain; x-axis: the temporal variation of the spike volleys, y-axis: the number of spikes in the spike volleys; the arrows in the plot mean the change of these two parameters from one group to the next group; Embedded graph shows how the spike volley is defined: top row is synfire chain group, middle row shows the raster plots of each group, bottom row shows the spike volley is the summation of all spikes in one group. The number of spikes in one volley determines the magnitude of the volley and the temporal variation of the spikes determines the width of the volley. (d)raster plot of all neuron groups when the spikes propagate along the synfire chain stable; (e)the raster plot when the spikes gradually dissolve along the synfire chain, adapted from \(^24\) ................................................................. 16

Figure 1-9: Raster plot of 10 HVC\(_{RA}\) neurons and 2 HVC interneurons during birds’ singing and call. ; adapted from \(^10\) ............................................................... 18

Figure 1-10: Different temporal windows for SDTP rule: (a) the window from excitatory neuron to excitatory neuron; (b) the window from excitatory neuron to inhibitory neuron; (c) the window from inhibitory neuron to inhibitory neuron. All temporal windows are in milliseconds; adapted from \(^41\). ............................................................... 23

Figure 1-11: Timeline of eligibility function, reward signal and synaptic weight. A: the profile of eligibility function; B: timeline of reward STDP ingredients. Top row is the spike trains of pre and postsynaptic neuron. The second row is the eligibility function of two spike pairs: pre is before post for red pair and pre is after post for green pair. The black line is the superposition of eligibility function of these two
pairs. The bottom two rows are reward signal and synaptic weight respectively; adapted from 48 ................................................................. 25

Figure 1-12: Tempotron model. (a) 10 afferents to tempotron, the solid spikes belong to + pattern, which means resulting an action potential and the gray spikes belong to – pattern, which means avoiding action potential; (b) the membrane potential trace for each pattern due to failures and the dashed line is the firing threshold; (c) synaptic change from learning mechanism; adapted from 76 ...................................................... 31

Figure 1-13: Variance model. (A) the weighted sum of N afferents gives the total input current, which depends on pattern; (B) An example of 400 afferents lasting 250ms with 5 spikes from each afferent; (C) Normalized synaptic current from 5 afferents; adapted from 88 ................................................................. 34

Figure 2-1: the network connection for recognition. (A) Synfire chain structure: gray ovals represent synfire chain neuron groups consisting of several neurons (big circles). The grey lines are the lateral connections whose magnitude is scaled by the number of neuron in one group. The small circles attached at the end of gray lines means excitatory synapses. (B) Figure 2-1B is taken as one slice (the box in Figure 2-1A) from the synfire chain structure. The white circles and solid dots represent excitatory synapse and inhibitory synapse. 1, 2, 3 are synfire chain neurons belonging to group 1, 2, 3 respectively; s1, s2, s3 are different types of sensory neurons and I1 and I2 are inhibitory neurons. Figure 2-1B is adapted from 96 .................. 45

Figure 2-2: the scheme of coincidence detection .......................................................... 52

Figure 2-3: the formation of plateau potential of neuron 2. Red (blue) dashed lines represent the timing of excitatory (inhibitory) spikes. The embedded network shows the network used in this graph. The parameters are specified in Table 2-1 .................. 53

Figure 2-4: membrane potential dependent synaptic plasticity. When the neuron is hyperpolarized, there is no change; when it is mildly depolarized, it induces LTD, and when it is strongly depolarized, LTP is elicited. LTP will work until the firing threshold (-45 mV) is met .......................................................... 55

Figure 2-5: the illustration of axon withdrawing on post-synaptic neuron. Gray lines are lateral connection, and dashed lines are plastic synapses ......................... 56

Figure 2-6: the mean and standard deviation of membrane potential cause by synaptic noise ................................................................................................................. 57

Figure 2-7: the recognition is demonstrated in 200 trials. (A): averaged subthreshold membrane potential in 200 trials; the blue, green and red lines are from one neuron in 1st, 2nd, 3rd group respectively. (B): averaged excitatory conductance in 200 trials for each group. (C): the raster plot for each group in 200 trials. (D): an example of actual membrane potential from one neuron in each group. (E): the fluctuation of membrane potential of a single neuron caused by synaptic noise. The firing threshold of that neuron is set to plus infinity. The solid line is the averaged
membrane potential in 200 trials, and the dotted line and the dashed line are the upper and lower bound determined by standard deviation in 200 trials.

Figure 2-8: this raster plot in 100 trials shows the recognition performance of a network with 9 synchronized groups, whose size is 5. The connection from sensory neuron is designed to recognize the sequence 1 2 3 4 5 6 7 8 9. The right arrows denote the timing of the sensory inputs projecting to the corresponding neuron groups.

Figure 2-9: The recognition performance of a synfire chain with 15 neuron groups whose population size is 5 and its favorite sequence is 1, 2, 3, 4, 5…14. Solid line with square shows the up state probability of the check group (group 15) fed by its favorite sequence. Other lines are all zeros with different noises as specified. The up state threshold is set as -60.2mV. Other sequences include: (1) 1, 2, 3, 5…13; (2) 1, 2, 3, vacancy, 5…13; (3) 1, 2, 3, 3…3; (4) 1, 14, 4, 9, 7, 13, 11, 7, 2, 12, 7, 9, 12, 13; (5) 1, 10, 14, 9, 11, 12, 2, 5, 6, 3, 4, 13, 8, 7; (6) 9, 3, 6, 11, 13, 4, 2, 10, 5, 12, 7, 8, 14.

Figure 2-10: the selective development of super-synapse is determined by the training sequence. Gray lines are lateral connections, dashed (dotted) lines are plastic synapses and solid lines are matured synapses. (A): initial connection before training. (B): the super-synapse connection trained by the sequence 1 2 3. (C): the super-synapse connection trained by the sequence 1 3 2.

Figure 2-11: the initial connection diagram for sensory learning except those inhibitory neurons and inhibitory connections. The network is a synfire chain with three synchronized groups, whose population is three. The dashed lines and dotted lines are the initial weak connections from s2 and s3 respectively.

Figure 2-12: the training process of the sequence 1 2 3 4 in 272 trials. The synfire chain is composed of 4 synchronized groups whose population size is 5. Initial feedforward connection strength is 0.1 time of the recognition value. (A): the raster plot of 4 neuron groups. (B): the synaptic strength development over 272 trials. Different colors represent corresponding targeting groups. Green lines are for all synapses projecting to group 2, red is for group 2, and blue is for group 4.

Figure 2-13: the training with random deletion of the sensory inputs. The synfire chain network is composed of 4 groups with the population size as 5. (A), (B) show the raster plots of learning process trained the sequence 1 2 3 4 with the missing rate of 0.1, 0.4 respectively. (D): total trials need for learning depend on the missing rate. The dot is the mean value over 50 runs, and the upper and lower bounds specify the standard deviation over 50 runs. Notice the scales of (A) and (B) differs a lot. The total trials needed for high missing rate is 1763 (A), and 388 (B) for low missing rate.

Figure 2-14: the training with random mutation of the sensory inputs. The synfire chain network is composed of 4 groups with the population size as 5. (A), (B) show the raster plots of learning process trained the sequence 1 2 3 4 with the mutation rate of 0.1, 0.4 respectively. (D): total trials need for learning depend on the mutation rate.
The dot is the mean value over 50 runs, and the upper and lower bounds specify the standard deviation over 50 runs. Notice the scales of (A) and (B) differs a lot. The total trials needed for high mutation rate is 906 (A), and 354 (B) for low mutation rate. ................................................................................................................................. 76

Figure 2-15: the mixed training with two sequences (1 2 3 4 and 1 4 2 3). The synfire chain network is composed of 4 groups with the population size as 5. (A) and (B) are the raster plots when two sequences have equal opportunities of presence. (A): final super-synapses are favorable to one of the training sequences; and (B) there is no super-synapses developed after the second synchronized group at the end. (C): the dominance probability with different mix ratios, which is the probability of sequence 1. The blue (red) line is the probability of super-synapses favorable to sequence 1 (2), and the peak is the probability of no super-synapses developed as shown in (B) . . . 79

Figure 2-16: the training with homogenous synaptic decaying. The synfire chain network is composed of 4 groups with the population size as 5. (A), (B) show the raster plots of learning process trained the sequence 1 2 3 4 with the decaying rate of 0.99, 0.93 per second respectively. (C): total trials need for learning depend on the decaying rate. The dot is the mean value over 50 runs, and the upper and lower bounds specify the standard deviation over 50 runs. Notice the scales of (A) and (B) differs a lot. The total trials needed for high decay rate is 1837 (A), and 302 (B) for low decay rate. ................................................................................................................................. 81

Figure 2-17 (A): the raster plot of learning with a long training sequence (1 2 3⋯ 28 29 30). The synfire chain network is composed of 30 groups with the population size as 5. (B): the trial at which each super-synapse has formed vs. group ID. .......................... 81

Figure 2-18: another representation of Figure 2-17A. ................................................................. 83

Figure 2-19: raster plots with different noises. All other parameters are the same with Figure 2-17. (A): there is no noise. (B) There is very little noise whose maximum is 0.0001. ........................................................................................................................................ 84

Figure 2-20: the learning rate for different parameters. (A): learning rate for different initial value of the weak feedforward projection (from 0.1 to 0.6) (B): learning rate for different population size (1, 2, 5, 10, 20, and 30). (C): learning rate for small noise (0, 0.001, 0.001, and 0.01); (D): learning rate for various noise levels (0, 0.2, 0.4, and 0.8); ........................................................................................................................................ 86

Figure 2-21: the network consists of 30 groups with the population size as 5, trained by sensory sequence 1 2 3 ⋯ 30. (A): the withdraw time vs. group ID at different noises. (B): the standard deviation of withdraw time in same group vs. group ID at different noises; (C): the plateau potential vs. group ID at different noises. (D): the stable plateau potential vs. synaptic noise. The network is the recognition network, which is the training result of the sequence: 1 2 3⋯30. .............................................................................................................. 88

Figure 2-22: the wiring from sensory neurons to synfire chain neurons in order to recognize sensory sequence 1, 2, 2, 4, 5. .................................................................................................................. 89
Figure 2-23: comparison with recognition of sequence 12245 and sequence 12345. Top row are the raster plots of 5 neuron groups in 100 trials and bottom row are plots of the mean membrane potential in 100 trials. Left side is for 12245 and right side is for 12345. The noise is 0.2.

Figure 2-24: membrane potentials of recognition of 12245 without noise. The inset shows the full scale membrane trace of all groups. The main plot focuses on the activities of g2 (green) ad g3 (red). The dashed lines represent the timings of excitatory inputs with the same color as its target neuron group.

Figure 2-25: the components of A-current, left side lists the expression of A-current we used, in the middle are the parameters, and right panels are the overall static A-current, activation part, and inactivation part from top to bottom given the parameters listed in the middle. The parameters are adapted from Jin’s model.

Figure 2-26: comparison of membrane potential of the neuron of g3 group with and without A-current. The top two rows are the membrane potential and A-current without A-current and the bottom two rows are the membrane potential and A-current with $g_A = 3$.

Figure 2-27: comparison of A-current model from different authors. (A) Peter Dayan & LF Abbott’s book; (B) Simpolinsky’s model; (C) Jin’s model; (D) my model.

Figure 2-28: recognition of sequence 12245 with reshaped A-current. Top row are the raster plots of 5 neuron groups in 100 trials and bottom row are plots of the mean membrane potential in 100 trials. Left side is for noise as 0.2, and ride side is for no noise.

Figure 2-29: parameter set B

Figure 2-30: parameter set C

Figure 2-31: A-current model about activation and inactivation of A-current in Dayan & Abbott’s book P.224.

Figure 2-32: the working version of static A-current.

Figure 2-33: the performance of recognizing sequence with repetitive inputs. Top row are the raster plots of 5 neuron groups in 100 trials and bottom row are plots of the mean membrane potential in 100 trials. Left side is for parameter set C, and ride side is for parameter set D.

Figure 2-34: recognition performance for parameter set C

Figure 2-35: recognition performance for parameter set D
Figure 2-36: recognition performance for parameter set D with the input interval varying in 12-20ms. .............................................................. 108

Figure 2-37: the left column is the recognition of the same sequence 12245, which is not the preferred sequence (12345) of the network beside A. The right column is the recognition of the preferred sequence 12245 of the network beside A. (A) and (C): the raster plot of five groups with 100 trials and each group contains 5 neurons. (B) and (D): the mean subthreshold membrane potentials of one neuron from each group. .................................................................................................................. 109

Figure 2-38: surface plot of membrane potential with different lateral excitation strengths. The model use single neuron with parameter set C except lateral excitation. The inset is the top view of the surface plot and the red line specify the lateral excitation in parameter set C. The synaptic input is as table: .......................................................... 111

Figure 2-39: surface plot of membrane potential with different feedback inhibition strengths. The model use single neuron with parameter set C except feedback inhibition. The inset is the top view of the surface plot and the red line specify the feedback inhibition in parameter set C. The synaptic input is as table: ............... 113

Figure 2-40: surface plot of membrane potential with different feedback inhibition strengths. The model use single neuron with parameter set C except feedback inhibition. The inset is the top view of the surface plot and the red line specify the feedback inhibition timing in parameter set C. The synaptic input is as table: ............... 114

Figure 2-41: surface plot of membrane potential with different feedforward excitation strengths. The model use single neuron with parameter set C except feedback inhibition. The inset is the top view of the surface plot and the red line specify the feedforward excitation in parameter set C. The synaptic input is as table: ............... 115

Figure 2-42: the learning process of repeated input sequence 12245 of synfire chain composed of 5 groups with the population size as 5. (A): the raster plot of learning process of 5 groups in 282 trials. (B): the synaptic growth of each neuron groups and different target groups are specified with corresponding colors in panel A. ............... 116

Figure 3-1: raw sound pressure wave and song spectrogram. ....................................................... 122

Figure 3-2: smooth spectrum to get acoustic signal. The bottom spectrum is one slice sample of the whole spectrogram. The window is 20ms, and the frequency spans from 0 to 8000 Hz. Above is the plot of the log power vs. frequency. The red line is the original signal and the blue line is the smoothed signal by the Savitzky-Golay method with a span of 16 points. ................................................................. 123

Figure 3-3: 100 FFT coefficients as feature vectors. The top panel is the FFT of the raw spectrum overlapped in many windows, the bottom one is the FFT of the smoothed spectrum overlapped in many windows. The smoothed FFT decays quickly to zero around 100, so we adopt the first 100 FFT coefficients of the smoothed signal as feature vectors. ........................................................................................................... 124
Figure 3-4: vector quantization algorithm flowchart. Adapted from page 54.

Figure 3-5: The effectiveness of the clustering algorithm on 2D space. In A and B: there are 5 2D Gaussian clusters (red dots). The blue dots in A (B) give the prototype vectors of 4 (8) categories. In C and D, the first two components of the feature vectors of a bird song are represented by black dots. The blue (green) dots in D give the prototype vectors of 8 (16) categories. The red dots in D give the prototype vectors of 32 categories.

Figure 3-6: Acoustic feature clustering result. This is an example to show the validity of the feature vectors clustering algorithm. Each column shows the sample of the original spectrum in 20ms belonging to one category.

Figure 3-7: The symbolic symbol representation of three renditions of the same motif. Green dots are the symbols at different moments. Bottom panel is the corresponding spectrogram.

Figure 3-8: scheme of synfire chain network with feedforward sensory inputs. All lateral synaptic connections are denoted by light gray arrows. Each chain receives the same feed-forward inputs from Nif, which are denoted by dark gray arrow.

Figure 3-9: the two-compartment model of HVC_{RA} neuron. The colorful lines are synaptic connections.

Figure 3-10: spontaneous firing probability at different membrane potential. Panel A is the firing probability measure for 200ms and panel B is for 30ms.

Figure 3-11: Coincident detection phase diagram. The scale bar on the right represents the spike time of the HVC neuron. (A) is the model with A-current, \( g_A = 1 \); (B) is the model without A-current.

Figure 3-12: Network to recognize serial sequence input 1 2 3 4 5 6 7 8 9. In order to recognize the symbol sequence 123456789, each spaces 20ms.

Figure 3-13: spike propagation of synfire chain. (A) Spike propagation along a single excitatory line without any noise. (B) Spike propagation along a single excitatory line with noise in 50 trials. (C) Spike propagation along a synfire chain with 30 excitatory lines in single trial. (D) Spike propagation along a synfire chain with 30 excitatory lines in 50 trials. (E) No spike propagation along synfire chain without lateral connections. (F) Spike propagation with repetitive inputs.

Figure 3-14: This network is based on a symbol sequence desired to recognize. The sequence is 6,15,14,14,16,16,16,16,9. We see there a lot of repetitions on the same symbol, so one Nif neuron may target on several HVC neurons. The synaptic connections are set as before.

Figure 3-15: The simplest model for learning. There are two Nif neurons: 1 and 2, which corresponds to symbol input 1 and 2 respectively. There are two HVC neurons: s and 2. The connection Nif1 to HVCs has already been made as well as the lateral
connection from HVCs to HVC2. Initially, there is only a weak synaptic projection (dashed line) from Nif2 to HVC2, which will not elicit a spike on HVC2 solely........ 152

Figure 3-16: Spontaneous activity and the learning. Panel A is the raster plot of the network above when only trained by sequence 1 at 25ms, Panel B is the raster plot when the training sequence is 1(25ms), 2(40ms). The histogram (blue) in panel B is the distribution of the formation of super-synapse projecting to HVC2 over 100 trials.......................................................... 153

Figure 3-17: The minimal model involving axon withdrawing. Except the starter HVC neuron (HVCs), there are all-to-all weak connections from Nif neurons to HVC neurons, which is denoted by dashed line. Which connection becomes the super-synapse depends on the symbol sequence training the network, and the other one withdraws.......................................................... 155

Figure 3-18: The top row is the raster plot of HVC neuron 3, the middle row is the raster plot of HVC neuron 2, and the bottom row is the raster plot of HVC neuron s. Each raster plot is based on 100 trials. Column A is the raster plot of HVC neurons trained by sequence 1(25ms). Column B is the raster plot of HVC neurons trained by sequence 1(25ms) 2(40ms). Column C is the raster plot of HVC neurons trained by sequence 1(25ms) 2(40ms) 3(60ms). The histograms (blue) in panel B and C are the distributions of the formation of super-synapses projecting to HVC2 and HVC3 over 100 trials.......................................................... 156

Figure 3-19: Selectivity of synaptic growth. A: trained by sequence 1 (25ms), 3(40ms), 2(60ms). B: trained by sequence 1 (25ms), 2(40ms), 3(60ms)...................................................... 158
ACKNOWLEDGEMENTS

I am greatly indebted to my advisor Dr. Dezhe Jin. He helped me all the way through my research and thesis writing. Your encouragement helped me progress toward graduation and your generous ideas always shed new light on problems.

Professor John Collins has constantly joined our journal club and shared his insight with us and gave a lot of useful suggestions. Professor Banavar always cared about the progress of my thesis. Special thank goes to Dr. Alexay Kozhevnkov for his valuable contribution of bird song recording data.

I am always grateful to my parents, Delong Wang and Yafang Yun, who have always supported my efforts towards this degree and given me freedom to do what I like. Thanks also go to Dr. Joseph Jun for your warm-hearted help.

I want to thank the Department of Physics, which gave me a chance to start my research and gave me financial support. Special thanks to department secretaries Randi and Megan, thank you guys for your numerous reminders. And special thanks to Professor Rick Robinett, you are always approachable and helpful.

I would like to express my gratitude to all the people who assisted the completion of this thesis.
Chapter 1  Background review

1.1  Songbird vocal learning

1.1.1  Songbird vocal learning behavior

Communicating by vocal language has endowed human beings with superior power in the development of intelligence over other animal species. Although the vocalization produced by many other animals differs from human language in complexity and functionality, it is vocal learning that gives the human brain programmable flexibility in communication by language, in contrast with only the innate call for most animals. Nonetheless there exist a small number of animals which share vocal learning analogous to that of humans. Among them, the most commonly studied group is the songbird (oscine). Songbirds are widely raised as pets and are relatively small, thus making them ideal lab subjects to investigate vocal learning and auditory perception. Analogous to a human learner’s waning ability to acquire a second language after puberty, the early phase of vocal learning for songbirds is sensory learning followed by sensorimotor learning, which constitutes the sensitive period that marks a temporal window for the acquisition of a bird’s song for a whole lifetime (zebra finch) or seasonally (parrots). Figure 1-1 shows the vocal development of a zebra finch as an example. The sensory phase occurs from about 15 days after birth to about 60 days after birth and during this period, the juvenile bird listens to the tutor song from its father and memorizes it. The sensory phase is followed by the sensorimotor phase which occurs from about 25 days after birth to about 90 days after birth, when the bird starts to produce plastic song and keep refining its song by comparing to the tutor song it
memorized using auditory feedback. After 90 days, a zebra finch can produce matured song, while at the same time, it loses the ability to learn new song, and it maintains that original song throughout its life span.

![Figure 1-1: The development of the zebra finch song system; adapted from 2.](image-url)

Another common feature shared by birdsong and human language is that auditory feedback is an essential factor for both sensory learning and sensorimotor learning, not only in learning during the sensitive period but also in maintaining matured song in adulthood. Studies of songbirds show the tutor song as well as other songs are memorized as templates in some form during the sensory phase \(^3\)\(^4\). In the sensorimotor phase, the songbird creates plastic song with a relatively simple form, then incorporates its own template and its own improvisation, then refines it with auditory feedback, and finally crystallizes it into its own song. During the whole process, auditory feedback not only provides raw material for the song template but also guides the plastic song to crystallization \(^1\).
1.1.2 Neuron and cortex

Animals and human beings have robust behavior modes controlled by nervous systems. The nervous system is an interconnected structure composed of numerous neurons. Neurons are highly diversified in their morphology and functions. Overall, there are three components for one neuron:

1. there is one cell body called the soma;
2. there are many dendrites which have various morphologies;
3. there is one axon.

Both the cell body and the dendrites can receive afferents, and the cell body generates action potential. The axon is responsible for transmitting action potential to other neurons. Between the axon of the presynaptic neuron and the dendrites of the postsynaptic neuron is the synapse. There are two kinds of synapses: the inhibitory synapse and the excitatory synapse. When a postsynaptic neuron receives a spike input from a presynaptic neuron, its membrane potential will either be hyperpolarized by an inhibitory synapse or be depolarized by an excitatory synapse, which may lead to firing if depolarization is strong enough to reach the firing threshold. After firing, the neuron undergoes a refractory period, during which it will not respond to any external stimuli.

1.1.2.1 Ion channels

Neurons are soaked in body fluids full of ions, and the surface of each neuron is a membrane that separates it from outside world. Because there are many pump-like ion channels composed of complex proteins on the membrane, the concentration of ions are different inside and outside the membrane. These ions include $\text{Cl}^-$, $\text{Na}^+$, $\text{K}^+$, $\text{Ca}^{2+}$. The flow of these ions is
largely dependent on the membrane potential, which is the potential difference inside the membrane and outside membrane, conventionally set at zero.

1.1.2.2 Rate model and spike model

There are two models explaining how neurons communicate with each other. The first is the firing rate model. In this model, the information is encoded in the mean firing rate during a short period; however, this model cannot accommodate rapid changes. The second is the spiking neuron model. In this model, all the information is encoded in an interspike interval. In reality, neurons are normally very noisy; sometimes it is difficult to understand what is encoded. Both models are supported by experimental evidence. Since the phenomena discussed here are all related to spikes, we will concentrate on the approach of spiking neuron model.

1.1.2.3 Single neuron dynamic equation

When the neuron is at resting potential, the membrane potential maintains a constant. When the neuron receives synaptic input from the axon of a presynaptic neuron, it works as an integrator until its membrane potential reaches a firing threshold. The Hodgkin-Huxley model describes the dynamic properties of normal neurons in which the spike is generated by sharply tuning on and off ion channels. Various neurons can be described by this model with additional ion channels. For most neurons, the shape of the spike is quite stereotypical. Thus integrate-and-fire neuron model is widely used for computational convenience. In this model, the neuron works as an integrator below the firing threshold and gives an instantaneous spike output once it is above threshold level. After firing, the neuron undergoes a certain refractory period and then returns to normal integration.
1.1.3 Song related nuclei and pathway

The neural substrate for a song system is a hierarchical structure innervated with three pathways: a song production pathway, a song perception pathway and an anterior forebrain pathway (AFP), pivoting at HVC (used as a proper name).
The motor pathway is indicated by the green lines. It initiates from HVC\textsubscript{RA} neurons projecting to nucleus RA. Then RA projects to the nucleus controlling respiratory muscles and the syrinx. The AFP pathway is specified by the yellow lines. HVC\textsubscript{X} neurons project to area X in the basal ganglia and area X, DLM, and LMAN form a triangle, projecting to each other. Finally, LMAN acts on RA to converge the learning pathway into the motor pathway. The auditory pathway (red lines) starts with the auditory peripherals and goes to the auditory cortex (field L), and field L projects to HVC directly and indirectly via NIf. Experimental evidence supports that NIf is a main auditory signal of HVC neurons.
1.1.4 Song related nuclei intracellular and extracellular recording

1.1.4.1 Morphology of different types of HVC neurons

There are three types of neurons in HVC nuclei. The HVC$_{RA}$ neuron is projecting to nucleus RA, the HVC$_{X}$ neuron is projecting to X area in the anterior forebrain, and the HVC interneuron projects to HVC$_{RA}$, HVC$_{X}$ and other HVC interneurons. Figure 1-4 lists the morphology of different types of HVC neurons in slices and their typical response to a depolarizing current which is used as a verification of neuron type during intracellular recording. HVC$_{RA}$ is a relatively quiet neuron, thus it only gives one spike burst when it receives persistent depolarizing current injection. HVC$_{X}$ gives more spikes than HVC$_{RA}$ and its spike interval is irregular. HVC$_{INT}$ fires in a tonic fashion.

There are many intracellular and extracellular recordings of HVC and other nuclei (LMAN, NIf) of the zebra finch. HVC shows a selective response to each bird’s own song during playback when the bird is anesthetized $^{7-9}$. Intracellular recording of the premotor sequence on antidromically identified HVC neurons obtained when the bird is singing indicates HVC$_{RA}$ neurons shows a very sparse, precise timing spike burst that occurs only once every motif $^{10}$.
Auditory feedback is important to guide plastic song to crystallization. However, how the auditory feedback refines plastic song and where it happens in a bird’s brain are still open questions. HVC neurons show selective response during song production; it may serve as the site of potential sensorimotor integration. It is important to study the auditory response of HVC RA neurons, HVC X neurons, and HVC INT neurons. Intracellular recordings on anesthetized zebra finches indicate all three neurons show strong selective response to BOS vs. REV (BOS: birds’ own song and REV: time-reversed birds’ own song), which negates the previous result showing that HVC RA has no auditory response. Moreover, HVC RA neurons and HVC X neurons have distinct sub-threshold responses: HVC RA neurons have purely excitatory responses selective to BOS, while HVC X neurons have highly selective hyperpolarized responses truncated by a

Figure 1-4: The subtypes of HVC neurons and their characteristic response to injected current; The scale bar is 20µm. adapted from 11.
depolarized component. Since the HVC\textsubscript{RA} neuron is the origin of the motor pathway, the highly song-selective auditory response makes it a potential site of sensorimotor integration.

1.1.4.2 Hierarchical structure of BOS selectivity

HVC in zebra finches is the pivot of sensorimotor integration, the motor pathway and the anterior forebrain pathway. Highly selective neural responses to birds’ own songs are found during both singing and song playback. Because hierarchical selectivity also appears in other sensory systems, it is natural to ask whether the selectivity originates in the HVC or from its upstream nucleus. The afferents of HVC include nucleus interfacialis (NIf) and thalamic nucleus uvaeformis (Uva), both of which demonstrate singing-related activity. However Uva does not have auditory response; thus selectivity could only from NIf or from the thalamo-recipient forebrain auditory structure (field L), which also projects to HVC directly and indirectly (through NIf). As the NIf and field L may both be the auditory inputs to HVC, the single-unit recording of these two areas is obtained in response to different auditory stimuli (BOS: birds’ own song, REV: time-reversed birds’ own song, CON: song of the same species). NIf neurons respond less to REV than BOS; however, both of them are more expanded than the response in HVC, and there is no significant difference in the response in field L to different stimuli. Correspondingly, the $d'$ measure of NIf shows a wide peak on the positive side of the x-axis between 0 and 3.

* Single-unit recording: The electrode records the electrical activities outside the neuron. If the recording electrode is small enough (tip size $\sim$1 µm), the activity of at most one neuron nearby the electrode will be recorded. Recording in this way is generally called "single-unit" recording. The action potentials recorded are very like the action potentials that are recorded intracellularly with reduced magnitude ($\sim$ 1 mV).

† D’ measure $d' = \frac{2(c-\bar{b})}{\sqrt{\sigma_c^2 + \sigma_b^2}}$ is used to measure the selectivity of responses regarding to two stimuli. $\bar{A}$ ($\bar{B}$) is the mean instantaneous firing rate over the course of the stimulus A (B) and $\sigma_A$ ($\sigma_B$) is standard variation of the instantaneous firing rate over the course of stimulus A (B). The higher magnitude d’ is, the more selective the response of one stimulus is versus another stimulus.
whereas the d’ measure of field L is a more evenly expanded bump along -3 to 3. The song selectivity from field L to HVC seems increasing gradual as the auditory information is further processed. A correlation is observed between the responses in NIf or field L single unit recording with the multi-unit\(^\text{4}\) recording in HVC simultaneously. Although correlated activity is found both in neurons of field L and in the neurons of NIf, this phenomenon is most commonly seen only in NIf; furthermore, the activity of NIf often precedes the activity of HVC with different lags, which is not a prevalent occurrence in field L. Considering several subdivisions of field L is also one source of afferents to NIf; part of the selectivity and correlated activity may originate from field L, as a gradually generating phenomena\(^\text{18}\). Based on the data, two possibilities arise to explain the causal relation of HVC and its afferents on the processing auditory information. One is that both HVC and NIf receive a common auditory input from or regulated by Uva; however, Uva seems to be indifferent to auditory response. An alternate view speculates that NIf is the main auditory input of HVC and pre-processes the auditory information from its own afferents which include field L, Uva, and cHV, as the selectivity gradually increases along the auditory pathway from field L, via NIf to HVC. The latter view is confirmed by the intracellular recordings of HVC on BOS playback with the blockade of NIf experiment\(^\text{19}\).

\(^\text{4}\)Multi-unit recording: If the electrode tip is slightly larger (lower resistance), the superposed activities from multiple neurons may be recorded by the electrode. This type of recording is often called "multi-unit recording".
1.1.4.3 Auditory response of HVC and its main auditory afferent

In Figure 1-6, top panel, the multiunit auditory recording of NIf is compared to the intracellular recording of HVC at the same time. HVC_{RA} neurons show highly selective and sparse responses to BOS (birds’ own song), no response to REV (time-reversed BOS), CON (con-specific song), or burst of noise (each burst consists of white noise), and limited response to BOS-RO (BOS with reverse order of syllables); NIf is relatively more spontaneously firing than HVC_{RA} neurons, and has more response in general for REV-RO and CON than for REV and noise. The bottom panel of Figure 1-6 shows the change in auditory response of HVC projecting neurons before and after the inactivation of NIf. It is obvious that when NIf is pharmacologically inactivated, HVC neurons keep silent and mostly maintain resting potential; once NIf is back to function, HVC neurons show similar but attenuated responses with the response before
inactivation. This study confirms that the major auditory input of HVC is from NIf, although other nuclei may modulate the auditory response \(^{19}\).

Figure 1-6: Top panel: auditory response of NIf and HVC to different stimuli when the bird is anesthetized. Bottom panel: The auditory response of HVC due to inactivation of NIf (applying GABA on NIf); adapted from \(^{19}\)
1.1.4.4 *Synaptic connections in HVC neurons*

Selectivity regarding to BOS vs. REV is enhanced from NIf to HVC, and NIf is demonstrated as the dominant auditory input to HVC. The function of brain not only depends on the external input but also on the internal connection of the network. Thus the way in which the three types of HVC neurons (HVC\textsubscript{RA}, HVC\textsubscript{X}, HVC\textsubscript{INT}) are internally innervated with each other is crucial to process the auditory information sent from NIf. Experiments have been conducted to investigate the neural connections inside HVC nuclei. Intracellular recording from pairs of HVC neurons in vitro is analyzed by spike-triggered average to infer the connectivity among HVC neurons. The dominant connection between HVC\textsubscript{RA} neurons and HVC\textsubscript{X} neurons is inhibitory, which could be mediated by HVC interneuron. This connection is also shown to be one way from HVC\textsubscript{RA} to HVC\textsubscript{X}. HVC interneurons make divergent connections to both types of projection neuron, and reciprocal connection between both projection neurons and interneurons is possible, which could serves as a potential mechanism of synchronization within the neuron population \textsuperscript{11}. Studies have been conducted on brain slices to evaluate the connectivity among the three types of HVC neurons in vitro. However, the functional response of neurons to stimuli can only be revealed \textit{in vivo}. Intracellular recording on pairs of different types of HVC neurons \textit{in vivo} was investigated to evaluate the spike-triggered average response with the presence of BOS and REV \textsuperscript{20}. It was found that all three HVC neurons receive a common excitatory input. This point is supported by the following blockade experiment. Either the HVC local circuit is blocked by applying GABA or the GIRK channel which is essential for hyperpolarization for HVC\textsubscript{X} is blocked; the seeming discrepancy of response between HVC\textsubscript{RA} and HVC\textsubscript{X} disappears. This means the song-selective hyperpolarization of HVC\textsubscript{X} neurons originates from local inhibition inside HVC rather than from outside. The HVC local circuit dispatches the selective auditory
information from NIf neurons into two representations, which the HVC interneuron-mediated inhibition plays a major role to differentiate.

In order to explain the selective response of the zebra finch’s HVC neurons and explore the mechanism of vocal learning mediated by auditory feedback, a synfire chain model is suggested given the experimental evidence:

1. the major connection between HVC\textsubscript{RA} neurons is excitatory \textsuperscript{11};
2. the auditory response shows highly selective response to bird’s own song, whose subthreshold is depolarized \textsuperscript{9};
3. the firing of HVC\textsubscript{RA} neuron is quite sparse, at most once every motif \textsuperscript{10};
4. HVC\textsubscript{RA} gives repeated precise timing firing when songbird is singing \textsuperscript{10};
5. each firing of HVC\textsubscript{RA} neuron consists of 4 to 6 individual spikes lasting 5 to 10 ms.

1.2 Synfire chain theory

Figure 1-7: An example of synfire chain network with 3 groups (ovals) each of which consists of 3 neurons (big circles). The connections (gray lines) between adjacent groups are all-to-all. The small circles mean the connections are all excitatory.
A synfire chain (shown in Figure 1-7) is a feedforward network composed of neuron groups. This idea was proposed by Abeles to accommodate the conflict of a limited number of neurons and the prominent robustness of animals’ behavior-controlled neurons \(^{21}\). There is no synaptic connection among the neurons in one group, but the neurons from two adjacent groups have an all-to-all synaptic connections, which means the neuron in the downstream group receives convergent synaptic input from the neurons in the upstream group, and the neurons in the upstream group fan out divergent synaptic connections to the neurons in the downstream group. There are variations within the synfire chain structure: basically, the feedforward connection is all-to-all style, although in some cases one neuron does not receive all outputs from its upstream group; also, the feedforward connection is normally assumed to be excitatory; however, it could be inhibitory, which would allow the avoidance of avoiding global synchronization in a firing rate model. There are two modes of spike propagation along the synfire chain: asynchronous propagation and synchronous propagation. Asynchronous propagation works within the firing rate neuron model. In this mode, the firing of one neuron is not correlated with other neurons in the same group. Synchronous mode works with the spiking neuron model. The firing in one group is temporally correlated and excites the correlated firing of the next group. Thus the spikes propagate along the synfire chain like a volley. It is theoretically \(^{21,22}\) proposed synchronous propagation is the only stable propagation and it is supported by both simulation\(^{23-25}\) results and the study in brain slices \(^{26}\).
Figure 1-8: Top graph: Flow chart of spike activity and dispersion along the synfire chain; x-axis: the temporal variation of the spike volleys, y-axis: the number of spikes in the spike volleys; the arrows in the plot mean the change of these two parameters from one group to the next group; Embedded graph shows how the spike volley is defined: top row is synfire chain group, middle row shows the raster plots of each group, bottom row shows the spike volley is the summation of all spikes in one group. The number of spikes in one volley determines the magnitude of the volley and the temporal variation of the spikes determines the width of the volley. (d)raster plot of all neuron groups when the spikes propagate along the synfire chain stable; (e)the raster plot
In synchronous propagation mode, spikes in the same neuron group are elicited simultaneously without noise, or distribute in a narrow window temporally and then form a pulse packet with a small temporal window, usually in ms, when the environment is noisy. It is argued that only synchronous spikes can propagate along the synfire chain stably and two important parameters are used to characterize this propagating ability: the number of the spikes in a volley and the temporal variation of the spikes. The flow chart (Figure 1-8) about these parameters along the synfire chain shows there are two outcomes for the spike volley: one is the fixed point of stable propagation with a finite number of spikes and finite standard deviation which are stable along the chain, even with the perturbation of noise; the other is trivial dissolution, which means the pulse packet gradually vanishes and dissolves into background noise along the synfire chain. When the spikes propagate synchronously along the synfire chain, the spike time in the same group will be bounded in a small window, which means the neurons in the same group fire synchronously, and thus their spikes share the same precise timing. Different groups of neurons have different timing. Experimentally, the repeated precise firing of multiple neurons from several groups could be an indication of a synfire chain network. Nonetheless, neurons in the same group do not necessarily group together spatially, and they may spread all over a relatively large region comparing with a possible technically exploration region. It is possible to record dozens of neurons simultaneously in experiments; these neurons may belong to different groups and may have repeating patterns of precise spike timing and form sequential firing patterns. There are a few such studies in vivo and in vitro, and in behaving monkeys. In vivo, normal neurons behave noisily with a lot of spontaneous spikes. Is this kind of evidence is just a coincidence, or an indication of a synfire chain network? It is argued that from the

---

8 Spike volley: a pulse packet consisting of all the spikes in one group. The number of spikes and the temporal variation in one spike volley characterize the synchrony of spikes in the group.
In a statistical view, the probability of a synfire chain is not prominent; rather it is more favorable to the correlated fluctuation by noise \(^{31,32}\). In biological neural networks, even if synfire chain structure exists, the neurons in the same group may not be found locally since the soma of HVC neurons is less than 10\(\mu m\), the dendrites expand around 100 \(\mu m\), and the tip size for single-unit recording is around 1 \(\mu m\). The neurons in different regions could be synchronized when the synfire chain is activated. This increases the difficulty to find synfire chain in HVC.

Only recently was a single unit recording of repeating sequential firing patterns in HVC\(_{RA}\) projecting neurons of zebra finch reported \(^{10}\), which is more conclusive evidence for the existence of a synfire chain. First, the HVC\(_{RA}\) projecting neurons are basically very quiet neurons, their spontaneous firing rate is low, and they will only burst once at most during one song motif. Second, firing is strongly correlated with sensory input and motor behavior. HVC\(_{RA}\) projecting only selectively fires when given a bird’s own song as sensory input or during a bird’s singing.

Figure 1-9: Raster plot of 10 HVC\(_{RA}\) neurons and 2 HVC interneurons during birds’ singing and call. ; adapted from \(^{10}\).
Extracellular recording of multiple cortical neurons has been reported on layer V of both anesthetized and awake rats with the use of a probe array. Results indicate the alteration of brainwave-like UP and DOWN states. UP states exhibit persistent and prominent neuronal activities, and DOWN states are relatively quiet for neuronal activities. The onset of the UP state is the activation of neuronal activities, in which different neurons may have different latencies and exhibit different temporal profiles triggered by the onset. The temporal precision immediately after UP onset is accurate to the milliseconds level, and the precision degrades gradually over time. The richness of neuronal response within the UP state comes from both the cellular properties and network dynamics. Current findings are not fully compatible with either the idea the temporal dynamics has no relation with the function of brain or the synfire chain in which the temporal precision is repetitively precise for all the neuron groups. The individual latencies of each neuron are proposed to represent the different levels of information processing. Smaller latency measures may correlate with simple features and longer latency measures may correlate with complex features after processing. This finding indicates the temporal precision degrades with the process of information received, which could be associated with the jitter propagation along a synfire chain. With the random connections between group even inside group and noisy activities, it is possible there could be synfire chain like structure in biological system.

1.2.1 Premotor model based on synfire chair

Computational models based on a synfire chain are explored to analyze the spike burst propagation existing in songbirds’ HVC. In order for the spike burst to propagate stably without runaway or avalanche, the networks parameters must be fine tuned if the neuron model is only a single compartment model, regardless of whether the integrate-and-fire model or the Hodgkin-Huxley model is used. The intrinsic bursting mechanism must be introduced
achieve the robustness of motor behavior modeled. This mechanism is based on a two-compartment neuron model with a soma and a dendrite. When there is enough stimulation, a wide calcium spike will form on the dendrite and last tens of ms; this is very common on the postsynaptic dendritic area, which serves as a plateau for the spike burst on soma. The robustness is greatly enhanced by incorporating intrinsic firing, which is a plausible assumption; the length of the spike burst is relatively short (around 6 to 10 ms), and given such a short scale, the individual spike in each burst is more likely to be generated based by a cellular mechanism rather than synaptic transmission and integration.

1.2.2 Formation of synfire chain structure

One of key issues when using a synfire chain model concerns its formation. Instead of having everything dictated by genetic blueprint, the neural network substrate is gradually self-organized into a specific structure suitable to handle external stimuli repeatedly appearing during the response of those stimuli, and the whole process is governed by establishing rules with a few realistic restraints. Early studies on the formation of synfire chains from random assembly by self-organization yielded results with restricted value: the parameters need to be fine tuned; based on the firing rate model \(^{36}\), the formed chain is short and forms a close loop. Even the spiking neuron model with the training of STDP yields similar results \(^{37,38}\). In a recent study of the formation of synfire chain networks from a random ensemble \(^{39}\), axon remodeling is applied as a realistic constraint in addition to the STDP rule, which avoids the fine tuning problem and achieves the stability of synapse development, thus resulting in a chain structure along which the number of divergent and convergent connections on each neuron group is approximately conserved. Based on axon remodeling and STDP, a very long chain can be formed easily.
1.3 Synaptic plasticity: STDP and voltage dependent learning rules

The biological neural system exhibits great complexity not only in different types of neurons but also in the connectivity among these neuron ensembles, which is dynamic and may change depending on outside input. Of course the input itself can be sensory input or manual stimulation, or even spontaneous activities. The plasticity problem is about the modification of synapses and how the modification depends on various factors.

1.3.1 Spike timing dependent plasticity

Long before any experimental evidence, the influential argument from Hebb was:

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of cells firing B, is increased.”

Since Hebb first postulated this learning rule in 1949, it has been supported by significant experiment evidence related to different brain areas. It has also been extended by the anti-Hebbian learning rule. However, the correlational form is based more on the firing rate neuron model than is required by Hebbian’s learning rule. With the popularity of the spiking neuron model, STDP has been discovered in many brain areas wherein the synaptic change is determined by the difference of the timings of pre- and postsynaptic spikes, typically lasting only a few milliseconds. STDP is thought to capture the essence of Hebbian’s learning rule: the causality between the firing of presynaptic neurons and the firing of postsynaptic neurons; it logically and naturally extends the learning rule to the anti-Hebbian theory. The exact form of STDP varies with the types of neuron involved and the different brain areas. What is more complicated and unclear is the cellular mechanism underlying the process,
acting like STDP with traditional LTP and LTD induced by tetanic stimulation or pairing protocol.

Traditionally, in various brain areas Hebbian’s postulate (and the anti-Hebbian rule) is strongly supported by two types of plasticity: long term potentiation (LTP) and long term depression (LTD), both of which can be activity dependent and involve substantial change on the neural substrate, possibly serving as the base for learning and memory. In general, LTP can be induced by high frequency stimulation (HFS) on a presynaptic neuron or by low frequency stimulation (LFS) on a presynaptic neuron, coupled with a large postsynaptic depolarization. LTD can normally be induced by LFS with or without a small postsynaptic depolarization. For both kinds of inducement, NMDA receptors play a crucial role in bridging the presynaptic activation and postsynaptic depolarization and finally lead to a rise of the postsynaptic Ca\(^{2+}\) level, the magnitude and time course of which determine the direction of synaptic modification by activating different signal pathways. However, this mechanism is basically applicable on the tLTP “”, and is partially right for tLTD, and may involve more metabotropic channels which may account for the long time course. Different pre-post pairs (combination of excitatory neurons and inhibitory neurons) exhibit different temporal windows on STDP plasticity. Three combinations (excitatory to excitatory, excitatory to inhibitory, inhibitory to excitatory) are thoroughly investigated in terms of their polarity and temporal windows 41.

** tLTP (tLTD) : LTP(LTD) is induced by pairing presynaptic firing with spiking-timing protocol.
Figure 1-10: Different temporal windows for SDTP rule: (a) the window from excitatory neuron to excitatory neuron; (b) the window from excitatory neuron to inhibitory neuron; (c) the window from inhibitory neuron to inhibitory neuron. All temporal windows are in milliseconds; adapted from 41.
Compared to the fairly simple experimental stimulation protocol in vitro, the complex spike stimulation is more realistic in vivo. It shows the integration of STDP on complex patterns is highly nonlinear and depends on not only on spike timing but also on various factors such as stimulation frequency and stimulation duration \(^{41}\). Considering the complex morphology of the dendrite tree, the spatial dependence of STDP response is important to determine the polarity and the temporal window: a quantitative change over small distance finally leads to qualitative change along a long dendrite \(^{41}\). Additionally, STDP can be greatly affected by various neuromodulators such as dopamine.

In order to understand the functional role of the STDP learning rule embedded at a network level, spontaneous activities (especially on postsynaptic site) must be considered because STDP depends heavily on the timing of pre- and post-synaptic spikes, which could be greatly disturbed by spontaneous activities \(^{41}\). It has been reported that spontaneous activities lower the persistence of synaptic potentiation and depression \(^{42,43}\). On the other hand, the sensory-evoked spike pattern is found to echo through spontaneous activities in the peripheral sensory system and to replay during sleep in the hippocampus.

In the spiking neuron model, the information carried by a neuron is represented by the timing of its firing, which is thought to be sensitive to noise or ongoing network activities leading to a temporal variation of spikes \(^{44}\). In some biological systems, such as a barn owl auditory system, a precision of 10 ms must be achieved in order to locate the source of the sound \(^{45}\). Theoretical studies show the tuning of such a system can be obtained by training with STDP \(^{46,47}\).
Although STDP is widely accepted and supported by experiments using different species, the underlying microscopic mechanism is still a mystery. Reward-modulated spike-timing dependent plasticity is proposed as a biological mechanism to explain common accepted STDP phenomena. Like STDP, spike-timing still plays the determinant role on the bidirectional synaptic modification dependent on stimuli. Unlike STDP, the synaptic modification is not realized immediately after a postsynaptic spike. The pre and post spike-timing difference only activates a so-called eligibility function: it rises up immediately after a postsynaptic spike and decays to zero slowly. If the reward signal (neuromodulator) arrives at the synapse before the eligibility function decays to zero, the real modification of synapse happens; otherwise the

Figure 1-11: Timeline of eligibility function, reward signal and synaptic weight. A: the profile of eligibility function; B: timeline of reward STDP ingredients. Top row is the spike trains of pre and postsynaptic neuron. The second row is the eligibility function of two spike pairs: pre is before post for red pair and pre is after post for green pair. The black line is the superposition of eligibility function of these two pairs. The bottom two rows are reward signal and synaptic weight respectively; adapted from 48.
synapse does not change. Eligibility function serves as a permission window to synaptic modification, and a reward signal is the trigger for synaptic modification. This work incorporates reinforcement learning into STDP and becomes a more biological model. It naturally deals with the modulation on plasticity by various neuromodulators such as dopamine and acetylcholine.

1.3.2 Plasticity depending on postsynaptic voltage

In reality, synaptic plasticity exhibits various learning rules. For STDP, the change of synapse is determined by the temporal difference between pre and post spikes. The pairing protocol (low frequency stimulation pairing with postsynaptic depolarization) indicates there is another learning rule, one in which the postsynaptic voltage plays an important role.

Three series of experiments have been done to show the postsynaptic voltage dependence of homosynaptic LTP (LTD) induction in slices of the visual cortex of adult rats. The change of synaptic strength is induced by applying a tetanic stimulation (50Hz) on white matter and the response of postsynaptic potential is probed by applying low frequency stimulation (0.03Hz) on white matter (axon). In the first series of experiments, bicuculline is used to block the GABA_A receptor, thus raising the excitability of the postsynaptic neuron. Three neuron groups are treated with three different concentration of bicuculline (0, 0.1uM, and 0.3uM). The postsynaptic response is recorded before (as the baseline) and 20 minutes after the tetanic stimulation. The postsynaptic response shows no change without bicuculline, LTD with the concentration of bicuculline as 0.1uM, and LTP with the concentration as 0.3uM. Since bicuculline is the antagonist of inhibition, higher concentration results in more excitability, which implies LTD and LTP may have different thresholds of voltage dependence. Prolonged current (one minute) is injected to hyperpolarize the membrane potential by -40mV, or to depolarize the membrane.
potential by +20mV, during the tetanic stimulation. The postsynaptic response before and 20 minutes after tetanic stimulation shows no change for membrane potential hyperpolarized, LTD for no current injection, and LTP for membrane potential depolarized. Further investigation shows the LTP with 0.3uM bicuculline could fail to be induced and even is replaced by LTD instead when APV is applied to block the NMDA receptor. 54.

In general, there exist three regions in the voltage–plasticity function for the profile of LTD and LTP in naive synapses: at a postsynaptic voltage lower than the first threshold, synapses undergo no modification; at a postsynaptic voltage between two thresholds, LTD is elicited; and finally, at postsynaptic voltage higher than the second threshold, LTP is induced. The two thresholds could be shifted according to the past experience of the neuron: if the synapse is depressed, the movement is favorable to the induction of LTP; if the synapse is potentiated, the movement is favorable to the induction of LTD 55. The increase of postsynaptic Ca\(^{2+}\) is essential for the induction of both LTP and LTD. It is proposed that it is the magnitude of calcium that determines the direction of modification of the synaptic strength: LTD is elicited by a modest rise in intracellular Ca\(^{2+}\) influx and LTP is elicited by a large rise in intracellular Ca\(^{2+}\) influx 56. The key aspect of intracellular Ca\(^{2+}\) is confirmed by significant experimental evidence in the hippocampus and the neocortex 57-60. It is further found that the duration of postsynaptic Ca\(^{2+}\) rise as well as the magnitude of Ca\(^{2+}\) is an important factor in determining the direction of modification of synaptic strength 61. If the postsynaptic spike occurs in a short window after excitatory postsynaptic spike potential (EPSP), LTP is induced; if the postsynaptic spike occurs long after EPSP, LTD is induced 62. NMDA receptor-gated conductance is found to be related to the postsynaptic voltage threshold 54. Experiments show LTP cannot be induced by HFS if NMDA receptors are blocked and will recover to normal after the blockade drug is washed out, whereas LTD is not affected by the blockade of NMDA receptors in the neocortex of a freely moving rat 63. The activation of the NMDA channel and the increase of postsynaptic Ca\(^{2+}\)
concentrations are both required by the inducement of LTP and LTD, and the NMDA receptor works as a coincident detector of the glutamate by presynaptic activation and the removal of the Mg$^{2+}$ block on Ca$^{2+}$ influx by postsynaptic depolarization. The mechanism of postsynaptic depolarization is still controversial. It is assumed to be postsynaptic back propagation to dendrite area. However, it was recently found that back propagating action potential is neither sufficient nor necessary for postsynaptic depolarization. There is also experimental evidence which shows LTD induced only by subthreshold postsynaptic depolarization (dLTD) shares similar properties with LTD induced by the postsynaptic spiking (tLTD), which means that tLTD and dLTD are indistinguishable. The subthreshold depolarization appears to be a more fundamental mechanism than postsynaptic spiking. It shows that the mechanism based on the voltage dependence of postsynaptic depolarization is a more basic mechanism which could lead to an STDP-like mechanism, but STDP cannot reproduce the dependence on postsynaptic membrane potential.

A mathematical model is proposed to explain the transition from early plasticity to late plasticity and the maintenance of late plasticity. There are three successive phases in this model. Tagging occurs during E-LTP (E-LTD). There are three states of the synapse: a non-tagged state, a high state (E-LTP), and a low state (E-LTD). The transition from a non-tagged state to a high state or low state is controlled by the transition rates PH and PL, respectively, and once the synapse is on high (low) state, it is labeled with a high (low) tag; the tagged states return to non-tagged state with rates of KH and KL for the high state and the low state, respectively. All the transition rates depend on the stimulation protocol. As the tagging accumulates, when the total number of up state tags (E-LTP) and down state tags (E-LTD) exceeds a certain threshold number, the synthesis of protein necessary for plasticity is triggered, and its concentration will decay back to zero once the threshold is not met. As the protein is produced, the landscape of the

\[\text{dLTD: LTD is induced by pairing presynaptic firing with postsynaptic subthreshold depolarization.}\]
phenomenological potential surface is altered such that the potential of one of the bi-stable fixed points (non-tagged state and high or low state) is elevated, and the consolidation variable will move to the remaining fixed point and stay there for a long time (LTP or LTD), even when the landscape recovers its original shape, until the fixed point where it is resting is elevated again. Given the linear-like dependence of the tagging rate (PH and PL) on the postsynaptic membrane potential, this model gives ABS-like voltage dependence plasticity when 100 synapses are stimulated with 2Hz input lasting 50 s. This model also gives the dependence of synaptic change on stimulation frequency. Since the trigger for synthesis of plasticity protein is determined by the total number of E-LTP tags and E-LTD tags, the cross-tagging is accommodated by this model. Therefore the trigger and consolidation are not homosynaptic but a population behavior, although its acting region can be greatly affected by neuromodulators such as dopamine. This model is not able to explain the homosynaptic plasticity either with postsynaptic depolarization or STDP unless the neuromodulator is assumed to play a key role in lowering the triggering threshold for protein synthesis.

1.4 Spatiotemporal sequence coding

Neural coding is the internal representation of external stimuli and other information. The coding process extracts the useful information from sensory input or upstream input and makes it readable to other neurons. A single neuron basically works as an integrator with a firing threshold. When the threshold is reached, an action potential is sent out to all targets. Neuron ensembles work independently unless they communicate with action potential. Experimental data show the amplitude and shape of action potential is relatively stereotypical, so the action potential itself may not be a good candidate to
convey information from sensory stimuli. Based on analysis of the spike train of neurons, spike timing exhibits great variability for different types of neurons and for different trials of the same neuron. Two different views exist depending on the scale of time considered. In the firing rate view, the mean firing rate represents sensory information; the firing rate model is robust and simple, although its response is slow and individual spike timing is discarded. Another view assumes every spike timing is useful to convey sensory information. This temporal code is fast and rich in capacity. Sometimes it is not as reliable as the firing rate model because of a noisy environment in vivo, and the exact meaning of each spike is difficult to analyze. Since there are different types of stimuli and different features in natural stimuli, the sensory inputs exhibit very rich variability. In reality, a neuron may receive a spike train from different sensory neurons at different timings; thus this spike train forms a spatiotemporal sequence. For example, different odors are encoded as different spatiotemporal patterns in locust antennal lobes. How the brain uses spatiotemporal sequence is still an open question.

1.4.1 Perceptron

The perceptron model \(^{71,72}\) and multilayer perceptron \(^{73}\) are proposed to recognize spatial patterns. The perceptron model \(^{74}\) is a linear classification method involving a weighted sum of different inputs. The realization of perceptron is the artificial neuron network in which weights can be determined through supervised algorithms. The perceptron model is not applicable to pattern varying with time because it contains no internal representation of time. In order to deal with spatiotemporal patterns time is
treated as an additional spatial dimension. In the first method, spatiotemporal pattern is transformed into spatial pattern by binning the time. However, the temporal correlation is destroyed by using time as an extra spatial dimension. Other fixes include the introduction of a former state by recurrent connections and using the delay response function to represent time. Neither of them raises a natural solution for realistic spiking neuron.

1.4.2 Tempotron

Figure 1-12: Tempotron model. (a) 10 afferents to tempotron, the solid spikes belong to + pattern,
Temporal information in a spike train is shown to be an essential part of neural coding. Different applications (spike latency code and pairwise pattern) of temporal coding may be studied using the tempotron model 76, a pattern recognition and decision-making unit based on a single neuron model. The tempotron is simply a leaky integrate-and-fire neuron 77,78 with a firing threshold. The membrane potential of the tempotron is the weighted sum of all synaptic currents with all synaptic inputs (different timing). The exact detail of a single synaptic current is determined by a synaptic kernel function governed by a membrane decaying constant and a synaptic integration constant. Tempotron is used to classify two groups (+ and -) of multichannel spike trains. Each channel is an input to one synapse of the tempotron. The learning mechanism is based on error behavior in response to two kinds of patterns; for example, the + group requires the tempotron to fire once during the spike train, and the – group requires no spike during the spike train. If the goal fails, synapse strength should be modified proportionately to the kernel function, which is determined by the temporal difference between each input timing and the timing of the maximum voltage. Kernel function is a measure of the factor of the contribution by each spike input to the firing of the tempotron. The learning results indicate a specific pattern: those synapses whose spikes are close prior to the maximum voltage are modified drastically, and those synapses whose associated spikes are far away from maximum voltage have almost no change. Thus the kernel function is derived from integrate-and-fire neuron model; it also works as a significance window on the maximum voltage of tempotron. Tempotron learning can be derived from the gradient of the cost
function, which is the difference between the firing threshold and the maximum voltage from erroneous behavior. Two types of spatiotemporal patterns can be classified by the tempotron model. The first is embedded with the spike latency code which is suggested in the olfactory system ⁷⁹ and the retina ganglion ⁸⁰, in which there is only one spike in each channel. The tempotron model has more capacity than the perceptron model in terms of the ratio between pattern number and channel number. For the perceptron model, the limit of this ratio is bounded by 2 ⁷¹,⁸¹. For tempotron, in optimal kernel function, this limit approaches 3 for different channel numbers. For pairwise synchronous patterns ⁸²-⁸⁵, tempotron is able to discriminate two categories associated with category specific pairwise pattern. This is not accidental, because the pairwise pattern actually intensifies the short-term temporal correlation between channels. This means the tempotron is good at picking out the temporal correlation in input patterns. The online version of tempotron learning is based on the convolution of postsynaptic voltage using the kernel function. The synaptic change depends on whether the convolution exceeds a threshold value or not. The online version is biologically plausible because experimental evidence supports the plasticity pairing presynaptic spike and postsynaptic voltage ⁵⁴,⁵⁵,⁶⁷,⁸⁶. The tempotron model has a few drawbacks. First, it is based on single neuron and thus is too risky to assign important tasks biologically. How to incorporate the output of a population of tempotrons is another question. Second, tempotron is a supervised learning model, and there is no experimental support on the biological substrate of the supervision. Even if one exists, it seems an inefficient way to train single neuron. Third, due to the limit of the kernel function, the tempotron is very good at recognizing temporal features within the scope of the kernel function ⁸⁷. For temporal complexity which goes
beyond the scope of the kernel function, the combination of features separating far away from each other may not be recognized.

1.4.3 Variance model

Figure 1-13: Variance model. (A) the weighted sum of N afferents gives the total input current, which depends on pattern; (B) An example of 400 afferents lasting 250ms with 5 spikes from each afferent; (C) Normalized synaptic current from 5 afferents; adapted from 88.

Another single neuron model is proposed to classify spatiotemporal patterns perceived by a single readout neuron as weighted synaptic current 88. The training pattern is randomly
chosen from an ensemble of random multichannel spike input patterns. The aim is to
distinguish the training patterns from all other patterns not included as training sets.
Theoretical and experimental studies show the firing rate of an integrate-and-fire
neuron monotonically depends on both the mean and the variance of the input currents.
Thus the learning rule is applied to maximize the variance of the sum of the weighted
normalized synaptic currents, since only variance is pattern specific. Because the total
synaptic current is a weighted sum of synaptic currents from individual channels, this is
analogous to Hebbian learning in the sense that normalized individual current is the input
and total current is the output. The variance of total current depends on $W^T C W$, where $C$
is the covariance matrix of individual currents and $W$ is weight. When the learning
converges, the variance approaches to the largest eigenvalue of covariance matrix and the
optimal weight is the leading eigenvector of the covariance matrix of the normalized
synaptic currents, which shares the same mathematical structure predicted by Oja’s rule.

The same as how Oja’s rule formalizes Hebbian rule, positing a simplified version
only with only presynaptic and postsynaptic correlation. As the Hebbian rule leads to
weights growing problem, the simplified correlation learning cannot make the output
variance approach to the eigenvalue of input covariance (as Oja’s rule can) but to a
suboptimal value between the background variance and the leading eigenvalue. This
learning mechanism performs robustly with time warping. The covariance matrix
depends on the time difference of spikes from different channels, serving as a
coincidence window detector. As long as the coincident spikes are still within the
window after a time warp, the readout neuron can recognize the time-warped pattern.
Since recognition depends on the statistical properties (mean and variance) of input
currents and the covariance matrix detects coincident spikes from different channels by counting on the absolute value of spike timing difference, the model cannot discriminate the time-reversed input from original input. In addition, final output is the firing rate of the readout neuron; this model may not work well for brain areas where sparse and precise timed spikes occur. Their simulation results show the firing of readout is sparse, and all the discussions are based on the variance of input current of the readout neuron instead of the output of the readout neuron, which makes the readout unit unnecessary. After all, this supervised learning mechanism depends totally on the presynaptic input, which is not very biological from a casual point of view. Although there have been related plasticity experiments conducted without eliciting a postsynaptic spike, postsynaptic membrane potential is shown to be a more fundamental component of LTP and LTD, which is plausible both in experiments and in theory.

1.4.4 Network model

Although these models have the capability to discriminate patterns randomly generated, the learning process is supervised. The readout process depends on a single neuron which could present a problem with robustness. Network models also have been proposed to deal with the spatiotemporal problem.

Delayed activity is used as coding time in the model proposed by Hopfield and Brody. Delayed activity means the firing rate of the neuron decays over time when receiving stimuli. Different stimuli can be associated with different decaying rates. By choosing appropriate decaying rates for a specific spatiotemporal pattern, the firing rates
caused by different stimuli could converge to the same point at some time. When this firing is used to drive readout neurons, the convergence of the firing rate will make readout neurons act synchronously at that moment. This kind of transient synchrony marks the recognition of pattern. Other spatiotemporal patterns will not make the firing rate converge, thus there is no transient synchrony for readout neurons. This model can be applied to speech recognition lasting 0.5 s or longer when the input neuron has both frequency tuning and decay activity and the readout is composed of a neuron population with balanced excitation and inhibition. This model is quite robust in two aspects: first the recognition is invariant of time warping, because the convergence of firing rates is not affected by it; second, the readout uses the population of neurons instead of using them as logic items, which is characterized by the transient synchrony. However, the key to convergence is the decaying rate for each stimulus, which is chosen manually. There is no corresponding biological evidence to support this manual convergence.

Jin proposed a network model to recognize spatiotemporal sequence. In this model, the network is composed of simple integrate-and-fire neurons connected in a feedforward manner. Each neuron receives excitatory input from both lateral input and sensory input, and works as an AND gate waiting for lateral input and sensory input. The UP state is a platform to integrate coincident spikes in the scale of membrane constant, thus it holds the memory of correctness of past input. This model is invariant of the interval between sensory inputs because the duration of UP state can be elastic. The capacity of this recognition model is not bounded since it only depends on the interaction of two synapses. When there are more synapses, there are more patterns to recognize. By replacing the single neuron with a neuron population, robustness can be greatly enhanced.
However, this recognition does need specific wiring from sensory neurons to the neuron chain. We solve this problem by introducing a biologically feasible learning mechanism in the following chapters.

1.4.5 Liquid state machine

There is yet another approach for tackling spatiotemporal problems. Spatiotemporal patterns can be encoded through the interaction between external stimuli and the temporal varying network state instead of using specific temporal mechanisms such as delayed lines and oscillators. The liquid state machine (LSM) is proposed as a generic computing device similar to the Turing machine. The computation is divided into two separate parts: the first stage involves a randomly connected recurrent neural network receiving external stimuli, which maps the temporal stimuli to a higher dimensional state space non-linearly. The state has two components: an active state and a hidden state. The active state is the ongoing activities of each neuron and the hidden state includes other factors that may affect the ongoing activities, such as short-term plasticity and neuromodulators. Ongoing activities can be depicted by the firing rate of each neuron. The firing rates of all neurons in the network at one time are represented by a point in N-dimension state space. The evolution of the firing rates depicts a trajectory in that N-dimension state space. Based on theoretical estimate, this N-dimension representation of two stimuli is more likely to be linearly separable than just a temporal representation of them. Thus in the next stage, normal perceptron can be used as the readout unit to recognize different stimuli. The randomly connected recurrent
network responds to external stimuli as a fading memory, like the ripple caused by a pebble thrown into a pond. Therefore time is naturally represented in the aftermath of the network response and computation in multiple aspects and multiple scales can be handled at one time. However, the model is only a vague idea about the whole process: there is no systematic way to extract the information useful to a specific goal from the soup of a randomly connecting recurrent network. There is no specific guideline for constructing a randomly connected network to make the trajectories of stimuli separable enough. On the one hand, the network can process many computations at one time; on the other hand, it is not very efficient for a specific function because it lacks of a measure for key information and other information.

1.4.6 Van Rossum distance model

Another model concerns about auditory object recognition \(^{105}\). This model is inspired by the van Rossum metric \(^{106}\) which is a tool used to measure the distance between two spike trains. The van Rossum metric is the Euclidean distance of the difference of convolution of two spike trains with an exponential kernel. The decaying factor of the exponential kernel can be tuned to accommodate the rate model and the spiking neuron model. When the decaying factor is a small number (~1 ms), the van Rossum metric is almost the distance of two spike trains and depends on the number of non-coincident spikes. When the decaying factor is a large number (~1 s), the van Rossum metric depends on the difference between the firing rates of two spike trains. Based on this metric, a spike train can be compared to template spike trains and then be classified into different categories. A neural network composed of integrate-and-fire neurons is constructed to implement van Rossum metric’s function. Using the spike train from a spike recording of field L
(songbird during playback), this model proves to accomplish the classification task with the analytical van Rossum metric and outperforms in robustness two other models based solely on coincidence detection and rate detection. Although this paper concerns auditory object recognition, it does not actually deal with coding and decoding problems of spatiotemporal sequence. The model needs a template as an input spike train, which is an oversimplification for an auditory object, because an auditory object contains both spatial and temporal structure. Compared using the van Rossum model, the circuit can finish the same task but efficiency could be a problem.
Chapter 2  Sequence recognition and sequence learning based on plateau potential

2.1 Introduction

Synfire chain structure is proposed to explain the repetitive precise timing of spikes \(^{10,34,35}\) shown in different brain areas. However there are three questions waiting for answer:

First since synfire chain can result in precise timing; however in reality noise and ongoing activities of the network always make the firing of a specific neuron not so precise. How do we assess the affect of noise and the impact of the functionality of the synfire chain network? How does the synfire chain network resist the effect of random activities to mess the precise timing embedded in the network by noise?

Secondly, inspired by the evidence that HVC in songbird is thought to be a site of sensorimotor integration with highly selective, sparse, and precise spike timing during singing and auditory process, numerical simulation shows that synfire chain could provide spike burst propagation for premotor activity with precise timing, nonetheless how does the model integrate sensory information with the properties of synfire chain thus serves as a potential basis for sensorimotor integration is still an open question.

Finally, an essential factor of songbird vocal learning is the auditory feedback, which plays important role in refining the plastic song during sensorimotor phase, even maintaining the matured song after the song is crystallized, however the exact mechanism is still unknown since it involves the interaction among several nuclei. More importantly, the search for the template of tutor song yields quite mixed and dynamic changing results \(^{107}\). Although less is known for what happens during sensory learning phase, here we propose a possible mechanism to encode the song structure in large scale by synaptic plasticity among synfire chain neurons and its sensory input neurons, which could constitute the formation of the song repertoire. It is the first attempt
to build a perceptual learning on a premotor structure. Also we are interested at how the perceptual learning begins in a noisy environment.

2.2 Methods

There are four types of neurons in our model, namely (1) binary neuron; (2) integrate-fire neuron; (3) quadratic integrate-fire neuron; (4) multicurrent integrate-fire neuron. We also consider two types of synapses. They are (1) excitatory AMPA synapses (2) inhibitory GABA<sub>A</sub> synapses. The synfire chain is comprised of synchronized neuron groups, projecting to one another in a cascade manner, the sensory neurons send feedforward excitation to synfire chain neurons, and the interneurons provide global feedforward and feedback inhibition to synfire chain neurons. The wiring from sensory neurons to synfire chain neurons can be plastic with a voltage dependent learning algorithm.

2.2.1 Network Model

There are three parts in our network model: (1) the synfire chain network (Figure 2-1A) consists of groups of excitatory neurons projecting in a feedforward manner between two adjacent groups. This part processes the information from sensory neurons; (2) different kinds of sensory neurons provide a spatio-temporal spike sequence as the input to the synfire chain; (3) two kinds of interneurons provide feedback and feedforward inhibitions respectively to synfire chain neurons. The last two parts are illustrated in Figure 2-1B.

There are three kinds of network connections in our model: (1) lateral connections in synfire chain; (2) feedforward connections from sensory neurons; (3) connections with interneuron.
The excitatory lateral connection is from one group to its adjacent group in all-to-all style (Figure 2-1A). The feedforward connection is from sensory neurons to synfire chain neurons as shown in Figure 2-1B. Thus the firing of different sensory neurons provides a spatiotemporal spike sequence as the input to synfire chain network. The exact wiring from sensory neurons to synfire chain neurons depends on the spatiotemporal sequence intended to recognize or learn from.

In order to regulate the firing of synfire chain neurons shown, i1 and i2 in Figure 2-1B provide fast feedback and delayed feedforward inhibitions (small solid circles) respectively to synfire chain neurons. For every excitatory line in a synfire chain, there are two and only two interneurons (i1 and i2) to regulate the firing along that line. Thus for a synfire chain with the group size as n, there will be n excitatory lines, each of which has 2 interneurons associated with. Figure 2-1B shows the connections with interneurons and sensory neurons for one of excitatory lines in the synfire chain in Figure 2-1A. Other excitatory lines in the synfire chain will copy the configuration shown in Figure 2-1B. Global inhibition is not adopted because the lateral excitation works together with the immediately followed feedback inhibition to produce the plateau potential, which is important for coincident detection of lateral and feedforward inputs. With noise, the firing timings of the neurons in same group follow some distribution, thus the firing of the feedback interneuron (i1) may precede the lateral excitation to the next group.

i1 and i2 work in a complementary way: i1 fires if the synfire chain neuron fires when receiving sensory input; and i2 fires only when the synfire chain neuron is silent when receiving sensory input.

G is the increment of synaptic conductance when a neuron receives a synaptic input. When the input is excitation (inhibition), the excitatory (inhibitory) conductance changes certain amount (G). These parameters are listed in the Table 2-1.

---

‡‡ Excitatory line is a degenerated synfire chain structure with the group size as one.
Table 2-1: Network connection parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_t = 13.0$</td>
<td>lateral excitation between chain neurons</td>
</tr>
<tr>
<td>$G_{ff} = 5.5$</td>
<td>feedforward excitation from the sensory neurons</td>
</tr>
<tr>
<td>$G_{fbi} = 30$</td>
<td>feedback inhibition from I1</td>
</tr>
<tr>
<td>$G_{fbi} = 20.2$</td>
<td>feedback inhibition from I2</td>
</tr>
<tr>
<td>$G_{exif} = 0.3$</td>
<td>excitation from sensory neuron to I2</td>
</tr>
<tr>
<td>$G_{ext1} = 5.0$</td>
<td>excitation from chain neuron to I1</td>
</tr>
<tr>
<td>$G_{ext2} = 1.0$</td>
<td>excitation from chain neuron to I1</td>
</tr>
<tr>
<td>$G_{in12} = 15.0$</td>
<td>inhibition from I1 to I2</td>
</tr>
<tr>
<td>$G_{f} = 10.0$</td>
<td>feedforward excitation to n1</td>
</tr>
</tbody>
</table>

These parameters works together in order to achieve the following effects:

1. The feedforward connection to n1 ($G_f = 10.0$) is set to give the n1 an initial start firing from down state because n1 has no upstream neuron.
2. Neither feedforward input $G_{ff}$ alone nor lateral input $G_t$ alone will make the synfire chain neuron fire from resting potential with the presence of inhibition;
3. Lateral input $G_t$ alone will excite synfire chain neuron to a plateau potential with the presence of inhibition;
4. Feedforward input $G_{ff}$ will make synfire chain neuron fire if it is on plateau potential.
Figure 2-1: the network connection for recognition. (A) Synfire chain structure: gray ovals represent synfire chain neuron groups consisting of several neurons (big circles). The grey lines are the lateral connections whose magnitude is scaled by the number of neuron in one group. The small circles attached at the end of gray lines means excitatory synapses. (B) Figure 2-1B is taken as one slice (the box in Figure 2-1A) from the synfire chain structure. The white circles and solid dots represent excitatory synapse and inhibitory synapse. 1, 2, 3 are synfire chain neurons belonging to group 1, 2, 3 respectively; s1, s2, s3 are different types of sensory neurons and I1 and I2 are inhibitory neurons. Figure 2-1B is adapted from ⁹⁶.

### 2.2.2 Neuron Model

We model different kinds of neuron with different model as the below table.

**Table 2-2: Neuron models.**

<table>
<thead>
<tr>
<th>Modeled neurons</th>
<th>Neuron models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensory neuron</td>
<td>Binary neuron</td>
</tr>
<tr>
<td>Synfire chain neuron</td>
<td>Integrate-fire neuron</td>
</tr>
<tr>
<td>Feedback interneuron</td>
<td>Multicurrent integrate-fire neuron</td>
</tr>
<tr>
<td>Feedforward interneuron</td>
<td>Quadratic integrate-fire neuron</td>
</tr>
</tbody>
</table>

The spiking properties of these neurons are investigated as Jin showed in a previous paper ⁹⁶.
2.2.2.1 Binary neuron

We use binary neuron to model sensory neurons, which send spatiotemporal spike sequence to the synfire chain neurons. Here the term “spatial” means different kinds of sensory output and the term “temporal” means the temporal order of different kinds of sensory spike output. Basically, sensory neuron just gives a spike at a specific time and sends it out to its target synfire chain neurons and remains silent at other time.

2.2.2.2 Integrate-fire neuron

We use integrate-fire neuron to model the neurons along the synfire chain. In addition to normal integrate-fire neuron model, we have a special A-current channel \( I_A \) in our equation.

\[
\tau \frac{dU}{dt} = g_L (L - U) - g_E U - g_I (U - E_I) + I_A.
\]

\( g_E, g_I \) are excitatory conductance and inhibitory conductance respectively. They are dimensionless parameters.\(^{85}\) When the neuron receives an excitatory (inhibitory) input, \( g_E (g_I) \) changes certain amount specified by Table 2-1.

Other parameters are set as the below table, which are set as usual:

Table 2-3 parameters for integrate-fire neuron model

<table>
<thead>
<tr>
<th>( \tau = 20ms )</th>
<th>time constant for membrane potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L = -70mV )</td>
<td>Resting potential</td>
</tr>
</tbody>
</table>

---

\( ^{85} \) Single neuron can be regarded as a simple resistor-capacitor circuit. Different ion channels correspond to different resistors in parallel. A differential equation similar to RC circuit equation can be used to describe the time variation of the member potential. If the unit on both sides of the equation is in voltage, the conductance in the equation is a dimensionless number; if the unit is in current, the conductance is in mS/cm\(^2\).
Once the synfire chain neuron receives the lateral spike input, the membrane potential is quickly pushed to around -60 mV (up state as we will see in section) with slowly decaying, once at the up state, the neuron will fire if there is an incoming sensory input.

### 2.2.2.3 Quadratic integrate-fire neuron

Quadratic integrate-fire neuron provides feedforward inhibition (i2) to the neurons along one excitatory line whenever there are any sensory inputs. This inhibition is also controlled by multicurrent integrate-fire neuron (i1). This model is adapted from Hansel and Mato’s model.
to match the property of hippocampal spiking interneuron. In our frame, this model can produce inhibitory spike with relatively large delay.

\[
C \frac{dU}{dt} = A(U - U^*)^2 - g_E U - g_I (U - E_I) - I_c
\]

Table 2-4 parameters for quadratic integrate-fire neuron

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C )</td>
<td>0.9467 ( \mu F/cm^2 )</td>
</tr>
<tr>
<td>( U^* )</td>
<td>-59.5462 mV</td>
</tr>
<tr>
<td>( A )</td>
<td>0.012875 mS/cm(^2)/mV</td>
</tr>
<tr>
<td>( E_I )</td>
<td>-75 mV</td>
</tr>
<tr>
<td>( I_c )</td>
<td>0.1601 mA/cm(^2)</td>
</tr>
<tr>
<td>( \Theta )</td>
<td>-26.3462 mV</td>
</tr>
<tr>
<td>( R )</td>
<td>-64.1462 mV</td>
</tr>
</tbody>
</table>

**2.2.2.4 Multicurrent integrate-fire neuron**

Multicurrent integrate-fire neuron provides feedback inhibition (i1) to synfire chain neurons. The inhibition is triggered by the firing of the synfire chain neurons. This model follows the modification proposed by Lewis and Gersnter\textsuperscript{109} from conductance based interneuron model by Erisir\textsuperscript{110}. This model has fast deactivating potassium channel and simple threshold and reset mechanism and can model fast spiking interneuron.

\[
C \frac{dU}{dt} = -I_L - I_{Na} - I_{K} - I_{K3} - g_E U - g_I (U - E_I)
\]
Table 2-5 parameters for multicurrent integrate-fire neuron

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C = 1 \mu F / cm^2$</td>
<td>the membrane capacitance</td>
</tr>
<tr>
<td>$I_L = g_L (U - L)$</td>
<td>the leak current</td>
</tr>
<tr>
<td>$g_L = 1.25 \mu S / cm^2$</td>
<td>the leak conductance</td>
</tr>
<tr>
<td>$I_{Na} = g_{Na} m^3 h (U - E_{Na})$</td>
<td>the $Na^+$ current</td>
</tr>
<tr>
<td>$g_{Na} = 112.5 \mu S / cm^2$</td>
<td>the $Na^+$ conductance</td>
</tr>
<tr>
<td>$E_{Na} = 74 mV$</td>
<td>$Na^+$ reversal potential</td>
</tr>
<tr>
<td>$I_{K1} = g_{K1} n^4 (U - E_K)$</td>
<td>the Kv1.3 current</td>
</tr>
<tr>
<td>$g_{K1} = 0.225 \mu S / cm^2$</td>
<td>the Kv1.3 conductance</td>
</tr>
<tr>
<td>$I_{K3} = g_{K3} p^2 (U - E_K)$</td>
<td>the K3.1-Kv3.2 current</td>
</tr>
<tr>
<td>$g_{K3} = 225 \mu S / cm^2$</td>
<td>the K3.1-Kv3.2 conductance</td>
</tr>
<tr>
<td>$E_K = -90 mV$</td>
<td>the $K^+$ reversal potential</td>
</tr>
<tr>
<td>$E_t = -75 mV$</td>
<td>the reversal potential</td>
</tr>
</tbody>
</table>

The dynamical equation of the gating variables

$$\frac{dy}{dt} = \alpha_y (U) \cdot (1 - y) - \beta_y (U) \cdot y, y = m, h, n, p$$

With the parameters are:

Table 2-6 dynamical parameters for gating variables

<table>
<thead>
<tr>
<th>$\alpha_n$</th>
<th>$\beta_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_n = (3020 - 40U) / \exp\left[ -(U - 75.5) / 13.5 \right] - 1$</td>
<td>$\beta_n = 1.2262 / \exp(U / 42.248)$</td>
</tr>
</tbody>
</table>
Other parameters in this model are:

Table 2-7 other parameters in multicurrent integrate-fire neuron model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta$</td>
<td>$-30mV$</td>
</tr>
<tr>
<td>$R$</td>
<td>$-85mV$</td>
</tr>
<tr>
<td>$\tau_r$</td>
<td>$1.5ms$</td>
</tr>
</tbody>
</table>

And the gating variables $m$, $h$, $n$, $p$ are reset to $0$, $0.16$, $0.88$, $0.2$ respectively after each spike.

2.2.3 Synapse Model

2.2.3.1 Excitatory Synapses

The excitatory synapse is modeled by a kick-and-decay algorithm. When there is pre-synaptic spike, the excitatory synaptic conductance ($g_E$) of the post-neuron jumps by certain amount $G$ specified by in Table 2-1, then it decays exponentially ($\frac{dg_E}{dt} = -g_E / \tau_E$), the decaying time constant $\tau_E$ is $3ms$. 

\[
\begin{align*}
\alpha_g &= 0.0035 / \exp(U / 24.186) \\
\beta_g &= -(0.8712 + 0.017U) / \{ \exp[-(U + 51.25)/5.2] - 1 \} \\
\alpha_n &= -(0.616 + 0.014U) / \{ \exp[-(U + 44)/2.3] - 1 \} \\
\beta_n &= 0.0043 / \exp[(U + 44)/34] \\
\alpha_p &= (95 - U) / \{ \exp[-(U - 95)/11.8] - 1 \} \\
\beta_p &= 0.025 / \exp(U / 22.222)
\end{align*}
\]
2.2.3.2 Inhibitory Synapses

The inhibitory synapse is also modeled by a kick-and-decay algorithm. When there is pre-synaptic spike, the inhibitory synaptic conductance \( g_i \) of the post-neuron jumps by certain amount \( G \) specified by in Table 2-1, then it decays exponentially \( \frac{dg_i}{dt} = -g_i / \tau_i \), the decaying time constant \( \tau_i \) is 3ms.

2.2.4 Recognition model

In order to understand the function of a particular brain circuit, we have to take account into both the properties of neurons and synapses and the architecture of the network.

As we know, synfire chain is proposed as a spike propagation mechanism in brain to achieve both the robustness and feasibility with limit number of neurons. Without considering the properties of individual neuron, the synaptic parameters of synfire chain has to be fine tuned in order to achieve stable synchronous transmission \(^{34}\), otherwise the synchronous spikes will either be runaway or dissolving. The underlying reason is the propagation is only affected by synaptic transmission and neuron integration, both of which are dynamically changing. Ikeda pointed out random delay of synaptic transmission has the same effect on synchronous spike propagation as synaptic noise \(^{111}\). So it is necessary to use some single neuron property for example: plateau potential to regulate the propagation of synchronous spike propagation. Because single neuron property is embedded in single neuron and less affected by network states such as the accumulation of jitter, we expect to achieve more flexibility on spike propagation and more robustness by plateau potential or up and down states.
2.2.4.1 Up and down states

We want every neuron of synfire chain neuron work on two states: down state and up state as shown in Figure 2-2. Down state is defined as the neuron is at resting potential, being hyperpolarized even slightly depolarized. Up state is defined as the neuron is depolarized, and maintains that plateau potential for a while. Of course normal excitation will not make neuron stays on a plateau potential, so we use a slightly shifted excitation (red dashed lines) and inhibition (blue dashed lines) to produce a plateau potential as shown in Figure 2-3. In this figure, when the neuron is at resting potential, it receives an unchecked excitation from lateral connection along the synfire chain right after 109ms and its membrane potential rises up just like what a typical integrate-fire will do until the feedback inhibition occurs. Since the excitatory neuron in synfire chain projects to an interneuron which provides fast feedback inhibition along the excitatory line. So whenever a neuron receives a spike input from its upstream group, it is immediately followed by an inhibition from the feedback interneuron projecting that excitatory
line. When the neuron receives the feedback inhibition, the membrane potential is already depolarized but still below the proximity of firing threshold, and after that the membrane potential will keep this depolarized potential for a while because the excitation and inhibition are balanced, despite that there is still some drop due to leaky current. If there is no other input, the membrane potential will slowly return to resting potential. In this figure since there is second sensory input (just before 117ms in Figure 2-3), the downstream neuron will fire quickly and elicit another feedback inhibition, which will send the upstream neuron back to down state.

Figure 2-3: the formation of plateau potential of neuron 2. Red (blue) dashed lines represent the timing of excitatory (inhibitory) spikes. The embedded network shows the network used in this graph. The parameters are specified in Table 2-1.

Up and down states differ with each other in terms of excitability when receiving a sensory input. The down state has a lower membrane potential thus the excitatory neuron will not fire even when there is an input spike from a sensory neuron, accompanying with delayed feedforward inhibition. Once an excitatory neuron receives lateral input spikes from its upstream neuron along the synfire chain, the membrane potential will be depolarized and maintain this plateau potential for a while depending on the membrane time constant, which means the down
state transforms to the up state. Once on up state, any input from sensory neurons makes the excitatory neuron fire accompanying with fast feedback inhibition projecting on all the neurons and the feedforward interneuron along the line; then a lateral excitation spike is sent to its downstream neurons, whose state will be turned from down state to up state. Based on this coincidence detection mechanism, a synfire chain can only recognize the sensory sequence whose order corresponds how the sensory neurons connect to the synfire chain neurons and reject any other sequences.

### 2.2.5 Plasticity Model

#### 2.2.5.1 Learning algorithm

The recognition is determined by how the sensory neurons connect to the excitatory neurons in synfire chain. Starting from all-to-all connections, the exact wiring needed for recognition can be developed with an ABS-like voltage dependent learning mechanism \(^{54,55,67,68,70,76,86}\) (Figure 2-4), which depends on the membrane potential of the target neuron at the arrival of the sensory spike input. There are three regions in the learning curve: hyperpolarization (< -70mV), no change; slight depolarization (-70 ~ -63mV), LTD; more depolarization (> -63mV), LTP. In the first two regions (LTD and no change) the membrane potential covers the down state of the synfire chain neuron and in LTP region membrane potential overlaps with the up state of the synfire chain neuron. Since the up state differs from the down state with a higher membrane potential, therefore LTD is elicited only when there is a small depolarization, more depolarization exceeding the plateau potential will elicit LTP and there is no synaptic changes due to hyperpolarization. This mechanism gives a bidirectional synaptic change dependent on
sensory input also screen the synaptic fluctuation due to random noise which lead to small depolarization.

2.2.5.2 Axon remodeling

The process of axon withdrawing is illustrated in Figure 2-5. Initially, all sensory neurons connect to all excitatory neurons along the synfire chain with weak strength and the same training sequence is repeated in every trial. Based on the learning rule, all synapses will differentiate to different strengths as time elapses. For each postsynaptic synfire chain neuron, the first synapse to

![Figure 2-4: membrane potential dependent synaptic plasticity. When the neuron is hyperpolarized, there is no change; when it is mildly depolarized, it induces LTD, and when it is strongly depolarized, LTP is elicited. LTP will work until the firing threshold (-45 mV) is met.](image)
reach a threshold is the super-synapse and remains, while other synapses connecting the same target neuron will withdraw from it. However this process can be reversed if the super-synapse drops below the threshold, then all the withdrawn synapses can wake up from withdrawn state. Finally, a stable connection trained by the sequence is just the wiring needed for recognition of the training sequence. It is worth to mention that the withdrawing happens on the postsynaptic neuron, which is analogous to the synapse elimination process in neuron muscular junction\textsuperscript{112,113}.

2.2.6 Noise Model

Synaptic noise is introduced in the excitatory neurons in the synfire chain. The noise is both excitatory and inhibitory with the same maximum strength, both of which are Poisson spike train with a frequency of 100Hz. Each time the noise strength is randomly chosen between 0 and its maximum strength, which is also tunable.
Figure 2-6: the mean and standard deviation of membrane potential cause by synaptic noise.

Figure 2-6 show the relation of membrane potential and synaptic noise. The neuron is an isolated neuron with synaptic noise whose maximum strength is specified along x axis. The right panel is the mean membrane potential, and we can see raising the noise strength will lift the resting potential above -70mV. The left panel is the standard deviation of membrane potential verses noise. The relation is not linear since there is A-current activation and deactivation above the resting potential. The working noise of the simulation is set at 0.2, which produces fluctuation of 1.5mV above resting potential.
2.3 Robust recognition of sensory sequence with a synfire chain

2.3.1.1 Recognition example

The recognition mechanism of single excitatory chain is proposed by Jin 96. Whether single excitatory neuron is on down state or up state is determined by the history of sensory input sequence and regulated by the interneurons, because the up state is triggered by the firing from the upstream neuron along the synfire chain and serves as the basis to relay the spike to its downstream neuron. However this is the ideal situation where there is no noise; in reality there are not only various kinds of noises but also a number of small synaptic inputs randomly coming from other neurons. It is important to investigate how robust this synfire chain model is with different noises, which is more close to the actual environment in vivo. In order to study the robustness with noise, we add synaptic noise on the excitatory neurons in synfire chains, which are both excitatory and inhibitory with a frequency of 100 Hz. Under noisy circumstances the average subthreshold membrane potential and synaptic conductance over 200 trials are shown in Figure 2-7. The network is composed of 3 synchronized neuron groups and each group contains 5 neurons. The feedforward connection (\( G_f = 5.5 \)) is set up in order to perform recognition of the sequence 1 2 3 along with the lateral connection (\( G_l = 13.0/5 \)). The synaptic connection and strength are set as shown in Figure 2-1 in order to perform recognition. The desired sequence intended to recognize is 1 2 3, which are the efferent from sensory neurons s1, s2, s3 respectively. The connections from sensory neuron to synfire chain neuron are s1 -> n1, s2 -> n2, and s3 -> n3; here n1, n2, and n3 represent the excitatory neurons along the synfire chain. From the raster plot (Figure 2-7C), we can see the spikes propagate stably along the chain neurons given the desired sequence 1 2 3 with small fluctuation on membrane potential caused by noise. The left panel on the top row (Figure 2-7A) shows the averaged membrane potential, which shows how plateau
potential is formed one after another along synfire chain. Except the starter neuron, every excitatory chain neuron has up state and down state; and the transition from up state to down state is caused by the lateral excitation from the upstream neuron(s) along the synfire chain, in this example the up state of n2 is excited by the spike from n1 and the up state of n3 is excited by the spike from n2. The firing of the excitatory neuron along the chain happens only during the up state if there is a sensory input, but there is only transitional depolarization during the down state under the same situation, which provides the mechanism to reject other sequences which are not the desired one. Thus every spike emitted from excitatory chain neurons is the consequence with the coincident impact from two sides: one is the lateral excitation from upstream neuron(s) and the other is the feedforward from sensory neurons. The middle panel on the top row (Figure 2-7B) shows each neuron receives two spike inputs: one is lateral input and one is feedforward input, which will finally result to the firing of that excitatory neuron. Once the excitatory chain neuron gives a spike, feedback inhibition will turn the upstate of the neuron back to down state, on which the neuron refuse to response to any sensory input. This shows the up state and down state is not destroyed by the noise, because the synchronized spikes are more likely to propagate along the chains than a single spike as shown in Figure 2-7C. Figure 2-7D gives an example of the actual membrane potentials of three groups in one of 200 trials.
2.3.1.2 Spontaneous activities caused by synaptic noise

The membrane potential of the synfire chain neuron suggests noise does not ruin the base of recognition. But we may wonder why noise has little impact on recognition and there is any other side effect caused by synaptic noise? Figure 2-7E directly shows the effect of noisy synaptic inputs on the membrane potential over 200 trials. The neuron is from the third group along the synfire chain and its spike threshold is set to plus infinity in order to study the subthreshold influence of noisy inputs. The network set up is the same as Figure 2-7A. Sensory input sequence
(123) starts from 100ms every 15ms. To start from noise equilibrium state, all the sensory inputs do not show up until 100ms. Before 100ms we can see the membrane potential is lifted from -70mV (resting potential) to around -68mV, and the standard deviation is about 1.5mV. This indicates the excitatory noise has greater impact on membrane potential at resting potential than inhibitory noise does because of the inhibitory reversal potential which is -75mV, however it is not advisable to keep absolute equivalence between excitatory and inhibitory noise at resting potential, because this will also results in the dominance of inhibitory noise when the membrane potential is depolarized, which is not favorable to maintain the up state. With noise, the up state (about -60mV) and down state (-68mV) are still distinguishable. The noise only produces limited fluctuation when the membrane potential is stable or slow varying; however when the neuron receives either a lateral input or feedforward input so that the membrane potential experiences a fast change, those three lines almost overlap on each other, which means the membrane potential is much more confined. This indicates the noise such as 100Hz will not affect the transition between the states of the membrane potential very much, only plays a role when the neuron stays on one state for a while. Thus the mechanism of recognition without noise should still work under noisy circumstance since noise can not disturb the transition between states.

2.3.1.3 Recognition and rejection test

Since the mechanism of recognition is not affected by the synaptic noise, with the magnitude as we mentioned above. It is confident to see how this network recognizes the desired sensory sequence and rejects other sequences.
Figure 2-8 shows the raster plot for synfire chain network composed of 9 synchronized groups (g1, g2, g3...), each of which contains 5 neurons. So each row is actually the collapse of 5 individual neurons in the same synchronized group. The sensory neurons project to each of the group in a serial manner so that spikes can propagate along the synchronized groups only if the sensory input sequence is 1 2 3 4 5 6 7 8 9, which means the sequence is recognized. The right arrows represent the timing of the sensory input projecting to neuron group labeled in that row. Before 350ms, there is no stable spike propagation along the chain except that the starter (g1) will give spike whenever it receives a sensory input and sometimes the second neuron group (g2) will respond to sensory input as long as the sensory input is 1 2 as a fragment from the complete sequence 1 2 3 4 5 6 7 8 9. Only if the sensory input is from 1 to 9, the spikes can sweep along all
the neuron groups. The actual sequence is from 1 to 8 after 350ms, which leads the spike propagates from g1 to g8. After that we shows two kinds of special rejection: 1) the deficiency to fire for g9 shows the spike cannot be stably elicited only by lateral excitation; 2) following that g8 group receives several repeated sensory inputs, the deficiency to fire shows the stable spike raster cannot appear without lateral excitation because at this time g8 group is at down state. This plot demonstrates the actual performance of a recognition mechanism based on coincident detection: only the desired sequence or part of the desired sequence starting from the beginning of the sequence can elicit spike propagation along the neuron groups. If we use last neuron group as readout for recognition, in this sense, only the desired sequence can be recognized by the network, which means there is stable spike raster for the last neuron group. However each neuron group in synfire chain can be used as readout, because they not only pass information to downstream but also give a spike as indicator to recognize a fragment of desired sequence. Although there is synaptic noise in all the excitatory neurons with the same magnitude as in Figure 2-7E, as we expected, the noise does not diminish the working of the recognition, just spread over the spike raster, as long as it is inside the window in which the coincident detection works, this network is robust for recognition.

### 2.3.2 Noise tolerance for recognition

For recognition, noise is a challenge for the spike propagation of coincident detection, because spontaneous spikes can be elicited without coincident spikes. So it is important to study the propagation of spontaneous spikes with different noise amplitude so there is no room for false recognition caused by spontaneous spikes. For single excitatory chain, we can set the readout of recognition as the spike output of the last neuron along the chain. With noise situation and with neuron population model, we have to consider the questions: how to discriminate the false signal
of the spontaneous spikes caused synaptic noise. When noise is small, the spontaneous spikes cannot propagate along the synfire chain, the spontaneous spikes on each neuron in the same group can be regarded as independent to each other, this also means they tend to be unsynchronized temporally. In order to pick out the synchronized spikes of one group which are the indicator of the sequence recognition, an additional group is attached at the end of the last group receiving sensory input. The up or down state of membrane potential of the neurons in the additional group indicates the success of recognition or not.

Figure 2-9: The recognition performance of a synfire chain with 15 neuron groups whose population size is 5 and its favorite sequence is 1, 2, 3, 4, 5…14. Solid line with square shows the up state probability of the check group (group 15) fed by its favorite sequence. Other lines are all zeros with different noises as specified. The up state threshold is set as -60.2mV. Other sequences include: (1) 1, 2, 3, 3, 5…13; (2) 1, 2, 3, vacancy, 5…13; (3) 1, 2, 3, 3, 3…3; (4) 1, 14, 4, 9, 7, 13, 11, 7, 2, 12, 7, 9, 12, 13; (5) 1, 10, 14, 9, 11, 12, 2, 5, 6, 3, 4, 13, 8, 7; (6) 9, 3, 6, 11, 13, 4, 2, 10, 5, 12, 7, 8, 14.
Figure 2-9 shows the recognition probability (up state probability) of a synfire chain. The recognition probability of its favorite sequence (1 2 3 4 5…14) decreases from 1 to below 0.1 as the noise increases and other probability curves rise from 0 to 1. So for noise is 0.2, which is used in most of simulation, the recognition performance has an excellent contrast with favorite sequence to other sequences. For all other sequences, the recognition probabilities are zero for all noises.

2.4 Encoding from learning sequence to feedforward wiring

The selective recognition ability of the above model depends mostly on how the sensory neurons project to the excitatory neuron in synfire chain. How this specific wiring forms from a random initial environment is an open question. We assume all the sensory neurons connect to all the excitatory neurons in the synfire chains with a weak synaptic strength at the beginning of training. The learning algorithm is stated in above section. The change of synapses depends on the membrane potential of the postsynaptic neuron. With small depolarization, the change is LTD, and with more depolarization, the change is LTP. The membrane potential of the excitatory neuron is determined by whether it is on up state or down state. This means LTP will work only when the target neuron is on up state, otherwise LTD will play the dominant role if there is only small fluctuation caused by noise or get hyperpolarized. The neuron is pushed from down state to up state by the lateral excitation, which is sent by its immediate upstream neuron (group) along the synfire chain with a developed synapse. So the excitatory neuron (group) along the chain will be wired with the sensory neuron one after another.
Figure 2-10: The selective development of super-synapse is determined by the training sequence. Gray lines are lateral connections, dashed (dotted) lines are plastic synapses and solid lines are matured synapses. (A): initial connection before training. (B): the super-synapse connection trained by the sequence 1 2 3. (C): the super-synapse connection trained by the sequence 1 3 2.

Figure 2-10 shows a simplified model of the learning process without showing inhibitory neurons: There are three neuron groups each of which contains only one neuron. The sensory neurons are specified by square box with label on it. The lateral connections are denoted by gray lines, which are not plastic. Feedforward connections are denoted by dotted (projecting from n2) and dashed lines (projecting from n3). Solid lines are the matured synapses, which maintain a stable strength with plasticity. All dotted and dashed lines are the weak synapses, which are set as 0.1 of full blown strength (the strength needed for recognition) at beginning, then undergo synaptic plasticity. If the training sequence is 1 2 3, n1 will send a spike to n2 because it is the starter neuron and could fire without lateral excitation., then n2 is at up state when s2 send an
sensory spike output to it, and this will lead LTP acting on the synapse s2→n2. At this time n3 is at down state or hyperpolarized, so s2→n3 will not grow. Later on in the same trial when s3 comes, n2 has already returned to down state because of either inhibition, so s3→n2 will not grow. If the super-synapse of n2 has not formed yet, n3 will be at down state when s3 comes, since it does not receive lateral excitation from n2. As trials go on, s2→n2 reach the super-synapse threshold and s3→n3 withdraws, then n2 is able to send lateral excitation to n3, thus n3 will be on up state when s3 comes and s3→n3 will grow to be the next super-synapse along the synfire chain. Finally, only s2→n2 and s3→n3 are the super-synapses, all other synapses withdraw from their target neurons. By the same mechanism, the super-synapses will be s3→n2 and s2→n3 if the training sequence is 1 3 2. We can see that this mechanism guarantees the initial all-to-all feedforward wiring will finally converge to the wiring structure needed for recognition of the very training sequence.

2.5 Learning with synfire chain

The learning algorithm seems reasonable for a single excitatory chain. How good it is for a synfire chain composed of neuron population? First, we use a small synfire chain with 3 groups to exemplify the initial network connections.
Figure 2-11: the initial wiring of the network. There are three synchronized groups g1, g2, and g3, each of which contains 3 excitatory neurons. The synchronized neuron group projects to its downstream group in an all-to-all manner: all the neurons in pre-synaptic group project to all the neurons in post-synaptic group, which are denoted by the grey lines. The lateral connection ($G_l = 13.0/3$) is scaled by the population of the group. The feedforward is also the all-to-all style: all sensory neurons project to all the neurons in all groups with weak strength, denoted by dashed lines (projecting from s3) and dotted lines (projecting from s2), which ($G_{ji} = 0.55$) is set to 0.1 of matured feedforward synaptic strength. G1 group is the starter group, whose synapses are
already set up so that whenever there is a sensory input from s1, so neurons in group 1 will fire. The above diagram neglects the sets of interneurons, which are responsible to provide inhibition for each single excitatory chain. The connection for interneurons is specified as Figure 2-1A.

Figure 2-12: the training process of the sequence 1 2 3 4 in 272 trials. The synfire chain is composed of 4 synchronized groups whose population size is 5. Initial feedforward connection strength is 0.1 time of the recognition value. (A): the raster plot of 4 neuron groups. (B): the synaptic strength development over 272 trials. Different colors represent corresponding targeting groups. Green lines are for all synapses projecting to group 2, red is for group 2, and blue is for group 4.

Figure 2-12A is the learning raster plot for a synfire chain of 4 groups with 5 neurons inside each group over 200 trials, and the training sequence is 1 2 3 4. So the final connection should be s1->g1 (already there at the beginning), s2->g2, s3->g3, and s4->g4. The learning is finished around 150 trials and the super-synapses are quite stable in the following trials, and it is
indicated by the successive formation of the raster bar for each neuron groups. For g3 and g4, before the raster bar is formed, there is some spike assemble right after the spike of its upstream neuron. For example, before the red raster bar of g3 forms, there are some spikes appearing right after the green raster bar of g2. This happens where the withdrawing happens: when the super-synapse s2->g2 has formed, all the other synapses projecting to g2 are cut off, however the synapse s2->g3 still exists. So the sensory input from s2 not only excites g2 to fire through the super-synapse, but also gives g3 a small depolarization. With the lateral input from g2 and a weak feedforward input from s2, g3 will have more probability to fire because of the coincidence detection mechanism. And these spikes do not hurt the selective growth of the super-synapse, because the feedforward input from s2 precedes the lateral input from g2 and at that time g3 is at down state. After the correct super-synapse (s3->g3) has been picked out, s2->g3 will be cut off and these unnecessary spikes will disappear without the coincident inputs. Figure 2-12B shows the corresponding synapses’ growth of the above raster plot. The green lines are all the synapses projecting to g2 group and red lines are for g3 group and blue lines are for g4 group. It is worth pointing out that the synapses of one group will not begin to grow unless all the super-synapses in its upstream group have formed. Because the lateral projection strength is scaled by the number of neurons in a synchronized group, each time one upstream neuron’s firing contributes only small portion of total lateral excitation needed for the up state of the target neuron also the noise will spread over the small spikes from the neurons in upstream group. Unless all the super-synapses of the upstream group have formed, there can be synchronized lateral excitation, which is powerful enough to send downstream neurons to up state. Therefore the formation of solid raster bar could represent the formation of all the super-synapses in that group, since this means all the neurons in that group can give stable spike outputs. These plots are the typical result for the training of a complete single sequence, which could be compared with the various situations when the training sequence is not so perfect, which represents when happens in reality.
2.6 Various compromised learning

After we consider the synaptic noise and study the robustness of synfire chain network in both recognition and learning. Following we are going to focus on another category of noise: noise from sensory inputs. This means even the training sequence itself is not perfect, which is a more realistic assumption. We investigate three situations when the training sequence is not perfect: sequence with random deletion, sequence with random replacement, and mixed training.

2.6.1 Learning with random deletion

In this case, each element of the training sequence has a finite probability of not showing up, so in an actual trial, the actual training sequence could be 1 3 4, 1 2 4, 2 3 4, 2 4… instead of 1 2 3 4 every time.
Figure 2-13: the training with random deletion of the sensory inputs. The synfire chain network is
composed of 4 groups with the population size as 5. (A), (B) show the raster plots of learning process trained the sequence 1 2 3 4 with the missing rate of 0.1, 0.4 respectively. (D): total trials need for learning depend on the missing rate. The dot is the mean value over 50 runs, and the upper and lower bounds specify the standard deviation over 50 runs. Notice the scales of (A) and (B) differs a lot. The total trials needed for high missing rate is 1763 (A), and 388 (B) for low missing rate.

In Figure 2-13A, each element in the full sequence 1 2 3 4 has 10% chance of absence. Comparing with the ideal case, the formation of super-synapses is not affected too much except that the raster bar is more spreading. But when the missing rate is large enough and there are so many broken sequences, the training will slow down severely. In Figure 2-13B, each sensory input element has 40% chance missing, and it takes more trials (1783 trials) to develop all the super-synapses because the training sequence is broken which cannot elicit synaptic growth so that some trials are wasted in this sense. However it is just a matter of time to develop a series of matured super-synapse based on the complete training sequence. In a nut shell, this kind of random deletion of the element of the training sequence does not hurt the learning process but only delay the formation of super-synapses if the missing rate is large. Figure 2-13C studies the missing rate dependence of the total trial needed. When the missing rate is small (smaller than 0.25), the total trial needed increase linearly; and when it exceeds 0.4, the total trial needed will exponentially explode, so does the variance of the trials needed, which means the training is very unstable.

2.6.2 Learning with random replacement

Figure 2-14 shows the raster plot of another kind of imperfect training over 200 trials. The complete training sequence is 1 2 3 4, and each element of the complete sequence has some mutation probability to become any of the other sensory inputs. For example, input 2 appears at 115ms in the complete sequence, however at this moment 1, or 3, or 4 could appear instead of 2
according to some mutation probability and this mutation is independent for each element of the sequence. So the actual training sequence could be 1 2 4 1 or 4 2 4 4 or 1 2 1 2 or 2 4 3 4 or 1 1 3 4 or 1 2 1 4 or 2 2 1 4... In Figure 2-14A, the mutation probability is 0.1, which means the probability of the correct sequence (1 2 3 4) is only 73%, the formation of super-synapses, which is indicated by the raster bar, is not affected too much. Four super-synapses have formed in 354 trials, a little more sluggish than the ideal case. And there are some noisy spikes caused by random mutated sensory inputs, which happen anywhere the sensory input can appear because the mutation can happen anywhere. These noisy spikes are more when the mutation probability is 0.4 (Figure 2-14B), which means the probability for the correct sequence is only 24%; and the raster bar is still robustly developed in 906 trials, which also means the correct super-synapses have been developed, however the development of super-synapse is severely slowed down. This learning mechanism shows very good robustness against the random noise from the sensory inputs. Figure 2-14C studies the mutation rate dependence of the total trial needed. When the mutation rate is small (smaller than 0.3), the total trial needed increase linearly; and when the mutation rate is large, the total trial needed will exponentially explode, so does the variance of the trials needed, which means the training is very unstable.
2.6.3 Learning with mixed sequences

Figure 2-15 show the raster plot of the same network trained by two definite sequences with some mix ratio. The two sequences used in our example are sequence 1: 1 2 3 4 and sequence 2: 1 4 2 3, and for each trial only one of them can be selected according to the mix ratio. In Figure 2-15A, the mix ratio is 0.1, the sequence 1 has 90% chances to be selected and sequence 2 has 10% chances to be selected in each single trial. When one of the sequence is dominant (Figure 2-15A mix ratio=0.1), the wiring from sensory neurons to excitatory neurons evolves toward the structure determined by the dominant sequence and the minor sequence has little influence on the formation of the super-synapses. When the probability of one sequence is comparable to another sequence, there are three outcomes: the final connection is favorable to one sequence, to the other sequence, and there is no super-synapses formed favorable to either of the sequences. Most of the time, the evolution trained by one of sequences will beat the evolution trained by another sequence and take over all the super-synapses, which means if sequence 1 wins, s2->g2, s3->g3, s4->g4 are the super-synapses and if sequence 2 wins, s4->g2, s2->g3, s3->g4 are the super-synapses, also remembering that g2, g3, g4 are synchronized neuron groups which contain 5 neurons, so s2->g2 means s2 projects to all the neurons in g2 group. However sometimes two sequences will tie with each other when the sensory input from one sequence cannot take over all the super-synapses from all neurons in one synchronized group. At this time,
for example, s2 from sequence 1 projects super-synapses to 3 neurons of g2 group and s4 from sequence 2 projects super-synapses to 2 neurons of g2 group, then g3 group cannot get enough lateral excitation (lateral excitation is scaled according to the number in the same synchronized group) from g2 group trained by either of the sequences, thus g3 will never on up state, which means no LTP, no super-synapse projecting to g3 group no matter how many trials used. This happens most often when two sequences have equal probability to appear (Figure 2-15B), but even at this situation, the super-synapse evolution could still be dominated by the training from either of the sequences. In Figure 2-15C, the dependence of the probability of one sequence dominating all super-synapses on mix ratio is investigated. The sinking and rising dashed lines are the probability of sequence 1 and 2 respectively based on 50 runs each of which takes 600 trials. The probability of either sequence resembles the sigmoid function. And the peak line is the probability of tie, which is small peak centered at mix ratio = 0.5. When mix ratio = 0.5, there is only 10% chances of tie, another 90% chances will be that either of the sequence will dominate the learning process. This plot shows if the mix ratio of a minor sequence is 40%, the learning process is still very robust, and the major sequence will dominate the growth of all super-synapses definitely.
2.6.4 Learning with homogeneous synaptic decay

We still want to consider a special situation, which is not caused by noise. Since the learning algorithm only works when there is an incoming sensory input, the synapses are supposed to have a tendency to degenerate all the way in the learning and this works for all the plastic synapses.

Figure 2-16 shows the raster plot when there is homogeneous decaying for all the synapses. Figure 2-16A shows the small decaying rate (0.99 per second) does not affect the growth of super-synapses. For large decaying rate (0.93 per second), the growth of super-synapses is slowed down greatly as shown in Figure 2-16B. Figure 2-16C studies the decaying rate dependence of the total trial needed. When the decaying rate is large (larger than 0.96 per second), the total trial needed decrease linearly; and when it is lower than 0.94 per second, the total trial needed will exponentially explode, so does the variance of the trials needed, which means the training is very unstable.
Figure 2-16: the training with homogenous synaptic decaying. The synfire chain network is composed of 4 groups with the population size as 5. (A), (B) show the raster plots of learning process trained the sequence 1 2 3 4 with the decaying rate of 0.99, 0.93 per second respectively. (C): total trials need for learning depend on the decaying rate. The dot is the mean value over 50 runs, and the upper and lower bounds specify the standard deviation over 50 runs. Notice the scales of (A) and (B) differs a lot. The total trials needed for high decay rate is 1837 (A), and 302 (B) for low decay rate.

2.7 Learning a very long sequence

Figure 2-17 (A): the raster plot of learning with a long training sequence (1 2 3… 28 29 30). The synfire chain network is composed of 30 groups with the population size as 5. (B): the trial at which each super-synapse has formed vs. group ID.
Figure 2-17 is to test the stability of the learning mechanism we used. The synfire chain is composed of 30 neuron groups, each of which contains 5 neurons. There are 30 different types of sensory neurons, which have the all-to-all weak projection to all the excitatory neurons in the synfire chain. The network is trained by the sequence 1 2 3 … 30. It takes about 2745 trials for 30 super-synapses developing successively along the synchronized groups. The stable raster plot indicates the super-synapse has formed and the connection on one excitatory neuron has crystallized. From this plot, the raster bar of each neuron is always a little bit shorter than the raster bar of its upstream neuron and a little bit longer than the raster bar of its downstream neuron. This means only if the upstream neuron formed a super-synapse so that it can provide stable lateral excitation to its successor, the neuron receiving lateral excitation can stay on the upstate, which is the LTP working region, and develop the super-synapse of its own, then relay the lateral excitation to the next excitatory neuron. It seems this mechanism works quite stably and efficiently. Considering the sensory period of zebra finch is about 50 days, and assume the bird could accept the training about 50 trials a day, at the end of the sensory period, the bird could develop a sensory selective network composed of ~30 neuron groups, which could handle the auditory recognition lasting about 500-600ms, the stereotypical length of the syllables of bird’s song. Figure 2-18 is a magnification of Figure 2-17A. The noise in Figure 2-17 is 0.2, which will produce 1.5mV around resting potential.
Figure 2-18: another representation of Figure 2-17A.
We note that there are unwanted spikes when the plateau potential is formed. Surely it is caused by synaptic noise, because there are no such spikes when the noise is zero shown in Figure 2-19: raster plots with different noises. All other parameters are the same with Figure 2-17. (A): there is no noise. (B) There is very little noise whose maximum is 0.0001.
2-19A. In Figure 2-19B, where the noise is only 0.0001, as the spikes propagate along the chain, the unwanted spikes gradually appear. This means these spikes are caused by the asynchrony of spikes in the same group. If there is any noise, no matter how small it is, as it propagates along the synfire chain, the spikes in the same group will spread over a small window, sometimes this window is pretty large, which means it lies out of the coincident detection range, especially when the synapse is forming, at this time, one of neurons may not fire or fires not accompanying with the firings of other neurons in that group, which will result the absence of feedback inhibition to synfire chain neuron of its downstream group. The plateau potential is formed by a lateral excitation and an immediately following feedback inhibition, and if that feedback inhibition is absent, even if there is less lateral excitation than normal, the downstream neuron without feedback inhibition will fire. After the super-synapse has formed so that the lateral excitation and feedback inhibition is stably elicited, the extra spikes just disappear.

2.7.1 Learning rate

From the raster plot trained by a long sequence (Figure 2-17), the super-synapses seem to grow very regularly, which is favorable to the formation of super-synapse of a long synfire chain. Following we want to know what is the learning rate of the super-synapse and what is the important factor that could affect the learning rate.
Figure 2-17 suggests the learning rate is a constant along synfire chain, which is found to be affected by the initial value of the weak synapses. Before learning takes place, the initial state of the feedforward connection is the all-to-all weak connection. Figure 2-20A shows the learning rate varies with the value of the initial value of the weak synapse. It is reasonable because the initial value serves as the baseline of learning and the growth of synapses can be regarded as a logarithmic random walk with drifting. It is really a tradeoff to determine how weak it is for these weak synapses. Too large value is not a biological assumption, and too small value will take too much time to develop a super-synapse.

Another factor seems important to the learning rate is the size of neuron group. It is interesting to see if the growth of the super-synapses in the same neuron group interferes with each other or not? Figure 2-20B shows the learning rate of different population. Except the population size equals 1, the learning rates of other population sizes (2, 5, 10, 20, 30) are close to
each other. When the population size is more than 1, the learning rate is lower than the single excitatory chain (population size=1). Because the population size is a finite number other than 1, there must be a distribution of the trials needed for the super-synapse in the same group, also remember that the neurons in immediate downstream group will not be pushed to up state unless all the super-synapses in its upstream group have been developed, which is determined by the longest trials needed in upstream group, therefore the learning rate is lower for multiple population by the distribution of trials needed inside the population, and this distribution does not varies too much with different population size, which means the synaptic growth inside the same group is independent with each other. This is a good result in the sense that the learning rate is not affected with the population model, which is a more realistic situation.

Although in our learning model, noise is not an essential part in the sense that learning works no matter noise present or not, however whether the learning works faster or slower with noise is still an open question. It shows the learning rate can be affected by synaptic noise. As the maximum of noise amplitude increases, the learning is first expedited then greatly slowed down comparing with the baseline when there is no noise as shown in Figure 2-20D. However when the noise is really small (less than 0.01), it takes longer than the case without noise as shown in Figure 2-20C.
Figure 2-21A shows the withdraw time dependence on group ID at different noise. It is obvious that the super-synapses form linearly without no noise and little noise (0.001 and 0.2). As the noise increases (0.8), the withdraw time does not follow the linear relation with group number. At the same time, we studied the standard deviation of the withdraw time in a same group, and found the standard deviation will approach to a constant value along the synfire chain when the noise is small and will explode when the noise (0.8) is high. Since the learning algorithm is dependent on the plateau membrane potential of the target neuron, we study the plateau membrane potential of one neuron from every group in a recognition network for 1, 2, 3…30. Figure 2-21C shows the plateau potential of each neuron when the feedforward input arrives. We see that when the noise is small, the plateau potential will be stable along the synfire chain, and when the noise is high, the downstream neuron cannot maintain the same plateau potential as the upstream neuron as the error propagates along the synfire chain. Figure 2-21D shows how the plateau potential varies...
with noises, and the plateau potential is from 30th group, so we can regard the plateau potential as stable along the synfire chain. When the noise is higher than 0.4, the plateau potential is destroyed, and when the noise is small, the plateau potential is slightly lifted. This means when the noise is higher than 0.4, the learning is greatly slowed down because the plateau potential lies on downstate.

### 2.8 Recognition with repeated sensory inputs and A-current

In the above we assume all the sensory inputs in one sequence are different; however in reality there is no reason to prevent one (type of) sensory neuron has more than one target neuron. First let us see the problem with the recognition of sequence 1, 2, 2, 4, 5. This sequence is different with the sequences talked before in the sense that sensory input “2” appears twice consecutively. According to the encoding rule between the recognition sequence and the feedforward wiring, the corresponding network structure is shown Figure 2-22. Here neuron group is represented by single neuron because we focus on the scheme of the wiring from sensory neurons to synfire chain neurons.

![Figure 2-22: the wiring from sensory neurons to synfire chain neurons in order to recognize sensory sequence 1, 2, 2, 4, 5.](image)
2.8.1 The violation of coincident detection by sequence with repetitive inputs

The following simulation is based on a model without A-current. Integrate fire neurons also have different firing threshold and network has different connection strength with the rest of the paper. The different parameters are listed as below table: These parameters are from Jin’s model \(^9\). These parameters can also give recognition of non-repetitive sequence based on plateau potential.

<table>
<thead>
<tr>
<th>( \Theta = -54mV )</th>
<th>Firing threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G_{esf} = 2.4 )</td>
<td>Feedforward excitatory conductance</td>
</tr>
<tr>
<td>( G_{esl} = 4.8 )</td>
<td>Lateral excitatory conductance</td>
</tr>
<tr>
<td>( G_0 = 4.6 )</td>
<td>Initial excitatory conductance</td>
</tr>
<tr>
<td>( g_L = 1 )</td>
<td>Leaky conductance</td>
</tr>
<tr>
<td>( g_A = 0 )</td>
<td>A-current conductance</td>
</tr>
</tbody>
</table>

However the raster plot (Figure 2-23) shows that g2 and g3 fire almost together, and the firing of g3 splits into two branches. This is not what we designed the network for because not only the branches will relay false timing information to downstream neuron groups but also the branches will weaken the activity of the synchronized spike volley from one group gradually and eventually the spike raster will dissolve (Diesmann 1999). And we know the key point of recognition is to maintain the plateau potential. If the plateau potential is not stable, the recognition is not well grounded. Compared with the membrane potential for the normal recognition of 12345, the plateau potential becomes lower and lower along the synfire chain. The membrane potential of g3 (red line) has two weak peaks for 12245 instead of one prominent peak.
for 12345. For the first peak (around 120ms) the membrane potential surges which is far beyond the plateau potential produced by lateral excitation and fast feedback inhibition.

In order to study the cause of this surge of g3 neuron, we studied repetitive case without noise, and found coincidence detection mechanism is violated here as shown in Figure 2-24. Without noise, the network can still relay the spike correctly as shown in the inset; however there is still a protrusion for g3 between 115ms and 120ms. During this period, the excitatory input to g3 includes not only the lateral input from g2 but also feedforward input from s2, thus g3 receives two excitatory inputs in a short time, which elicits the protrusion instead of the plateau potential. Without noise, g3 does not fire but forms a protrusion which is not a serious problem, but it
becomes serious when there is noise which greatly enhances the probability of spiking on that protrusion. In order to accommodate this special situation in our model, the current surge due to two close excitations has to be controlled and the protrusion has to be flattened.

Although \( g_3 \) receives two excitations in a short time, they are quite different from the two excitations bridged by plateau potential because membrane potential rises up so quickly. We want to seek a way to enhance the capacitance of those excitatory neurons to the current surge. Including \( A \)-current channel in our integrate fire neuron model is one of the choices. Up to now, the functionality of our network has nothing to do with \( A \)-current. Actually \( A \)-current is introduced into the model to deal with recognition and learning with the sequence containing repetitive inputs.

Figure 2-24: membrane potentials of recognition of 12245 without noise. The inset shows the full scale membrane trace of all groups. The main plot focuses on the activities of \( g_2 \) (green) and \( g_3 \) (red). The dashed lines represent the timings of excitatory inputs with the same color as its target neuron group.
2.8.2 A-current

A current is a member of vast voltage dependent Potassium channel family which includes Kv3.3, Kv3.4, Kv4.1, Kv4.2, Kv4.4 channels\textsuperscript{114}. It usually activates around resting potential and quickly inactivates when depolarized, so it is also called transient Potassium current\textsuperscript{115}. Traditionally, it provides a mechanism to give low firing rate which is not possible only with Hodgkin-Huxley model\textsuperscript{74}.

Like most of Potassium currents, it is inhibitory. It is composed of an activation part and inactivation part which both depends on the voltage of the membrane. Originally, A-current model is adopted from Jin’s previous work\textsuperscript{116} which is shown in Figure 2-25.

\[
I_A = g_A \cdot \alpha^3 \cdot b \cdot (E_K - U)
\]

\[
\alpha = \frac{1}{1 + c_a \cdot e^{-(U+U_a)/t_a}}
\]

\[
b_\infty = \frac{1}{1 + c_b \cdot e^{(U+U_b)/t_b}}
\]

\[
\tau_A \cdot \frac{db}{dt} = -b + b_\infty
\]

\[
g_A = 1;
U_a = 70mV;
U_b = 80mV;
t_a = 5ms;
t_b = 6ms;
\tau_A = 3ms;
\]

Figure 2-25: the components of A-current, left side lists the expression of A-current we used, in the middle are the parameters, and right panels are the overall static A-current, activation part, and inactivation part from top to bottom given the parameters listed in the middle. The parameters

*** Page 197
are adapted from Jin’s model 116.

2.8.3 The effect and side effect applying A-current on integrate fire model

Figure 2-26 gives an example of applying A-current on integrate fire model. Obviously, there is an unwanted peak without A-current. Let us see if we can diminish this peak. From bottom two rows, we can see the peak is somewhat suppressed since the local highest point drops from -55mV to nearly -60mV and the shape is more smooth; however the biggest drawback is g3 neuron does not fire even without noise because the A-current lowers the whole plateau potential and even accelerates the decaying process of plateau potential. From bottom row, we see the A-current spans all the way when the membrane potential lies on the plateau potential. Also we found the membrane potential is hyperpolarized before 100ms, at which the sensory input sequence begins. This is caused by the static inhibitory A-current near resting potential. Suggested by these two side effects of A-current, we propose A-current has to be confined between the UP state and resting potential (DOWN state), at least for the static A-current, otherwise the excitability of both UP state and resting potential (DOWN state) will be compromised.
2.8.4 Reshape A-current to check current surge

Figure 2-27 compares the static A-current before and after reshaping with other’s people’s model. (A) and (B) are the static A-current shape used by other authors\textsuperscript{74,117} respectively, and (D) is the original A-current model I used, which is from Jin\textsuperscript{116}. They all widely span from hyperpolarized region (-80mV) to over firing threshold (-40mV). It is obvious that if we use these models, it is inevitable to shift down membrane potential as a whole. So we adjust the expression
of A-current, and confine static A-current between -70mV and -65mV as shown in Figure 2-27C. We also set the dynamic constant for inactivation $\tau_A = 0.7ms$ to assure A-current works as a surge protector to response quickly to rapid current and disappear quickly if there is none. The detailed expression of reshaped A-current is listed as the below table.

<table>
<thead>
<tr>
<th>Parameters in Figure 2-27C</th>
<th>Parameters in Figure 2-27D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_A = g_A a^3 b (E_K - U), g_A = 100$</td>
<td>$I_A = g_A a^3 b (E_K - U), g_A = 70$</td>
</tr>
<tr>
<td>$a = 1/(1 + e^{-(U+70)/5})$</td>
<td>$a = 1/(1 + 0.02e^{-(U+68)/0.5})$</td>
</tr>
<tr>
<td>$b_\infty = 1/(1 + e^{(U+80)/6})$</td>
<td>$b_\infty = 1/(1 + 0.01e^{(U+68)})$</td>
</tr>
<tr>
<td>$\tau_A \frac{db}{dt} = -b + b_\infty, \tau_A = 3ms$</td>
<td>$\tau_A \frac{db}{dt} = -b + b_\infty, \tau_A = 0.7ms$</td>
</tr>
</tbody>
</table>
Figure 2-27: comparison of A-current model from different authors. (A) Peter Dayan & LF Abbott’s book \cite{74}; (B) Simpolinsky’s model \cite{117}; (C) Jin’s model; (D) my model \cite{116}.

Figure 2-28 gives the performance of recognition with repetitive input sequence 12245 when the applied A-current is reshaped. Using parameter specified in Figure 2-27D, we can see the protrusion at the beginning of plateau potential is chopped off without shifting the UP state as a whole, and the current surge is still in control even there is moderate noise.

Figure 2-28: recognition of sequence 12245 with reshaped A-current. Top row are the raster plots of 5 neuron groups in 100 trials and bottom row are plots of the mean membrane potential in 100 trials. Left side is for noise as 0.2, and ride side is for no noise.
2.8.5 Other possibilities about parameter choice

There are also other possibilities about parameters listed in the following table. Parameters set A is used above to explain the functionality of A-current. In set B and set C, the firing threshold is raised to \(-48\text{mV}\) and \(-45\text{mV}\) respectively. To minimize the side effect, A-current should only play a role between resting potential and UP state which lies in below firing threshold. So raising firing threshold should leave more room to the playground of A-current.

Basically lateral excitation \((G_{exl})\), fast feedback inhibition \((G_{inb})\), how fast it is \((G_{exf})\) and inhibitory A-current \((g_A, a, b, \tau_A)\) work together to forge a plateau potential and feedforward excitation \((G_{exf})\) elicits a spike from plateau potential. Since we raised the plateau potential and A-current, so it is better to reduce the leaky conductance in order to maintain the stability of plateau potential.

Table 2-8

<table>
<thead>
<tr>
<th>Parameter set A</th>
<th>Parameter set B</th>
<th>Parameter set C</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Theta = -54\text{mV}), (g_L = 1)</td>
<td>(\Theta = -48\text{mV}), (g_L = 0.5)</td>
<td>(\Theta = -45\text{mV}), (g_L = 0.5)</td>
</tr>
<tr>
<td>(G_0 = 4.6, G_{exf} = 2.4), (G_{exf} = 4.8)</td>
<td>(G_0 = 8, G_{exf} = 3.5), (G_{exf} = 10)</td>
<td>(G_0 = 12, G_{exf} = 4.0), (G_{exf} = 11)</td>
</tr>
<tr>
<td>(G_{ex1} = 4.5, G_{inb} = 22.8)</td>
<td>(G_{ex1} = 4.5, G_{inb} = 22.8)</td>
<td>(G_{ex1} = 5.0, G_{inb} = 25.0)</td>
</tr>
<tr>
<td>(G_{ex2} = 1.0, G_{inf} = 20.2)</td>
<td>(G_{ex2} = 1.0, G_{inf} = 20.2)</td>
<td>(G_{ex2} = 1.0, G_{inf} = 20.2)</td>
</tr>
<tr>
<td>(G_{exif} = 0.3, G_{in12} = 15.0)</td>
<td>(G_{exif} = 0.3, G_{in12} = 15.0)</td>
<td>(G_{exif} = 0.3, G_{in12} = 15.0)</td>
</tr>
<tr>
<td>(I_A = g_A a^3 b (E_K - U)), (g_A = 70)</td>
<td>(I_A = g_A a^3 b (E_K - U)), (g_A = 6.5)</td>
<td>(I_A = g_A a^3 b (E_K - U)), (g_A = 10)</td>
</tr>
</tbody>
</table>
\begin{align*}
a &= \frac{1}{1 + 0.02e^{-(U_{+67}/0.5)}} \\
b &= \frac{1}{1 + 0.01e^{(U_{+68})}} \\
\tau_A &= \frac{db}{dt} = -b + b_\infty,
\tau_A &= 0.7\text{ms}

\tau_A &= \frac{db}{dt} = -b + b_\infty,
\tau_A &= 0.7\text{ms}
\end{align*}

Figure 2-29: parameter set B
2.8.6 Match a realistic dynamic constant

In the above parameters, the dynamic constant of inactivation of A-current is set as 0.7ms in order to achieve the agility of its response to current surge. Considering in animals, the inactivation time of A-current ranges from several milliseconds to hundreds milliseconds. It seems 0.7ms is too small. In Dayan & Abbott’s book, they used a more delicate model of A-current shown in Figure 2-31. In this model, the dynamic constant of activation $\tau_a$ is small, as an approximation, we can treat it as instantaneous. For inactivation part $\tau_b$, the value is around 3ms in the membrane potential region (~60mV) we are interested. However this raises other issue with...
higher dynamic constant: the A-current does not decay as quickly as before therefore there could be excessive accumulated A-current even the neuron is away from the inactivation. So it is advisable static A-current keep a distance away from UP state and resting potential.

Figure 2-32 gives the final version of static A-current and approximate region to UP and DOWN states. We can see the static A-current stays sway from resting potential (-70mV) and UP state (-55mV) with current surge about 5mV cushion region, which is reserved for accumulated A-current. The excitability on resting potential will not be affect because there is no static A-current on resting potential. And firing threshold is set as -45mV to leave more room for A-current. A-current only activates between a little bit higher than resting potential and up state, and checks
any current surge, when there are repeated inputs and feedforward excitation and lateral excitation reach the excitatory neuron in a short time period.

The underlying parameters are shown as following table

Table 2-9 A-current parameters

<table>
<thead>
<tr>
<th>A-current parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta = -45mV, g_L = 0.5$</td>
</tr>
<tr>
<td>$I_A = g_A a^3 b (E_K - U)$</td>
</tr>
<tr>
<td>$a = \frac{1}{1 + e^{-(U+63)}}$</td>
</tr>
</tbody>
</table>
For this A-current, at least two sets of parameters can recognize repetitive inputs sequence correctly with the difference that they produce different plateau potential as shown in Figure 2-33. For parameter D, although there is sharp protrusion instead of smooth edge for parameter C, the raster plot still behaves normally due to higher threshold.

Table 2-10

<table>
<thead>
<tr>
<th>Parameter set C</th>
<th>Parameter set D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_0 = 16$, $G_{exf} = 5.0$, $G_{exl} = 18$</td>
<td>$G_0 = 10$, $G_{exf} = 5.5$, $G_{exl} = 13$</td>
</tr>
<tr>
<td>$G_{exl1} = 5.0$, $G_{inb} = 30$</td>
<td>$G_{exl1} = 5.0$, $G_{inb} = 30$</td>
</tr>
<tr>
<td>$G_{exl2} = 1.0$, $G_{inf} = 20.2$</td>
<td>$G_{exl2} = 1.0$, $G_{inf} = 20.2$</td>
</tr>
<tr>
<td>$G_{exf} = 0.3$, $G_{in12} = 15.0$</td>
<td>$G_{exf} = 0.3$, $G_{in12} = 15.0$</td>
</tr>
<tr>
<td>$g_A = 12$</td>
<td>$g_A = 6.0$</td>
</tr>
</tbody>
</table>

\[ b_\infty = \frac{1}{1 + e^{U - U_d}} \]

\[ \tau_A \frac{db}{dt} = -b + b_\infty, \tau_A = 3 \text{ms} \]
2.8.7 Repetitive lateral inputs and repetitive feedforward inputs

Up to now we discussed the recognition of sequence with repetitive inputs, which is the recognition target of synfire chain network, and following we want to investigate how the network rejects those repetitive sequences which are not the desired one. Since the excitatory inputs include lateral inputs and feedforward inputs, we want to focus on two special situations: one is repetitive lateral inputs, and the other one is repetitive feedforward inputs. In order to test

Figure 2-33: the performance of recognizing sequence with repetitive inputs. Top row are the raster plots of 5 neuron groups in 100 trials and bottom row are plots of the mean membrane potential in 100 trials. Left side is for parameter set C, and ride side is for parameter set D.
this, we design a special sequence 1, 1, 2, 2, 2, 1, 2, 2, 4, 5, 6, 7, 8, 9, 10 as input sequence to synfire network intending to recognize 1, 2, 2, 4, 5, 6, 7, 8, 9, and 10. Since g1 is the starter group, the first three “1, 1, 1” give repetitive lateral inputs to g2 group and if g2 does not fire, the test is passed. The next three “2, 2, 2” is to test the network ability to reject repetitive feedforward inputs: for the first two “2, 2”, g2 and g3 fires respectively because they are following the third “1”, in other word, “1, 2, 2” is a segment of the desired sequence“1, 2, 4, 5…9”. For the third “2”, it should be rejected and no group fires. Following, the “1, 2, 2, 4, 5, 6, 7, 8, 9, 10” is to confirm the recognition for desired sequence, also test the stability of spike volley propagation along the synfire chain with lots of groups. For this part, it is good to see the sequential firing from g1 to g10. Comparing with the performances of parameter set C and D (Figure 2-34 and Figure 2-35), we see the network fails to reject successive lateral excitation for parameter set C, but still works for parameter set D. This is reasonable because there is larger lateral excitation and higher plateau potential with parameter set C. and It seems parameter D is more superior than parameter set C, however successive lateral excitation is just a special situation which could only happen by such sequence as “1, 1, 1” and no other possibilities. Actually because of higher plateau potential, it should be more stable and robust for parameter C, which is shown in the following section. Parameter D fulfills the requirements in all aspects in price of some level of robustness.
Figure 2-34: recognition performance for parameter set C
2.8.8 Sequence of inputs with variable interval

Up to now we assume the interval between inputs is the same. However, biologically, the fixed interval is quite rare because of the noise from sensory system, and information could be encoded in the intervals between spikes. In order to study the ability to handle variable interval sequence, we have to figure out the maximum and minimum interval this network can handle. The following table lists the maximum and minimum interval allowed by parameter set C and parameter set D. We can see parameter set C is more robust against the jitter of sensory inputs since it can elicit higher plateau potential. Figure 2-36 gives an example of the recognition with parameter set C with variable interval.
Table 2-11

<table>
<thead>
<tr>
<th></th>
<th>Parameter set C</th>
<th>Parameter set D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum interval</td>
<td>20ms</td>
<td>17ms</td>
</tr>
<tr>
<td>Minimum interval</td>
<td>12ms</td>
<td>14ms</td>
</tr>
</tbody>
</table>

Figure 2-36: recognition performance for parameter set D with the input interval varying in 12-20ms.
2.8.9 Recognition and rejection

In the following, we are going to study two categories regarding to repeated inputs: in the first case the repeated inputs are not of desired sequence determined by the network structure and in the second case the repeated inputs are the preferred input sequence or part of the sequence determined by the network structure.

Figure 2-37: the left column is the recognition of the same sequence 12245, which is not the preferred sequence (12345) of the network beside A. The right column is the recognition of the preferred sequence 12245 of the network beside A. (A) and (C): the raster plot of five groups with 100 trials and each group contains 5 neurons. (B) and (D): the mean subthreshold membrane potentials of one neuron from each group.
As shown in Figure 2-37A and B, the preferred sequence for this network is 12345, while the actual input sequence is 12245. For the second 2, there is no coincident detection on neuron group 3, and we want to check whether coincident detection mechanism works here or not. From raster plot (A), there is no chance for spikes to go through group 3 and there is no trace of spikes in the following downstream neuron groups. From the subthreshold membrane potential (B), we can see when the second 2 comes (130ms), both group 2 and group 3 turn to down state. For group 2, repeated sensory inputs cannot excite the neurons to up state; and for group 3, even when it is on up state, if there is no sensory input, the neurons will go back to down state with feedforward inhibition.

Figure 2-37 C and D shows the recognition of sequence (12245) with repeated inputs, which is the preferred sequence of the network. On the top of the column is the wiring from sensory neurons to synfire chain neurons, and we see sensory neuron 2 projects to two targets: neuron (group) 2 and neuron (group) 3. From the raster plot (A), the preferred sequence evokes good response with stable spike propagation along the chain from one group to next group. The subthreshold membrane potential shows the up state of neuron from group 3 is higher than the up state of neurons from other group, because the neuron receive both feedforward input (the first 2) and lateral input from g2 in a short time period. Figure 2-37 verify that the coincident detection mechanism works for the coincident inputs reach the neuron bridged by plateau potential and rejects other kinds of input pattern such as two adjacent inputs (too close), two inputs far away from each other.

2.8.10 Parameters affect plateau potential

The core of recognition lies whether the plateau potential can propagate along synfire chain. The plateau potential is formed by balanced but shifted lateral excitation and feedback
inhibition. Basically lateral excitation \( G_{exi} \), feedback inhibition \( G_{inb} \), excitation strength \( G_{exi} \) controlling response speed of feedback inhibition, and inhibitory A-current \( g_A, a, b, \tau_A \) work together to forge a plateau potential.

### 2.8.10.1 Variable lateral excitation

From Figure 2-38, we can see lateral excitation has to be large enough to produce a plateau potential given the feedback inhibition \( G_{inb} = -30 \). The working line (red) is chosen to generate a plateau for membrane potential and avoid excessive hump close to firing threshold.

![Figure 2-38: surface plot of membrane potential with different lateral excitation strengths.](image)

The model uses a single neuron with parameter set C except lateral excitation. The inset is the top view of the surface plot and the red line specifies the lateral excitation in parameter set C. The synaptic input is as table:
### 2.8.10.2 Variable feedback inhibition

Given lateral excitation ($G_{ex} = 13$), feedback inhibition has to be large enough to avoid firing and small enough to avoid pulling down plateau potential. The working line (red) is chosen between high inhibition resulting drop down of plateau to and low inhibition resulting to firing. The plateau is formed with appropriate feedback inhibition as shown in the shaded area of Figure 2-39.

<table>
<thead>
<tr>
<th>Timing</th>
<th>Synaptic input</th>
</tr>
</thead>
<tbody>
<tr>
<td>5ms</td>
<td>variable</td>
</tr>
<tr>
<td>5.3ms</td>
<td>-30</td>
</tr>
</tbody>
</table>
Given lateral excitation \( G_{\text{ex}} = 13 \) and feedback inhibition \( G_{\text{inb}} = -30 \), the delay for feedback inhibition has to be small enough to avoid peak at the beginning of plateau potential. The working line (red) is chosen at the timing (around 0.3ms behind lateral excitation) when the
peak just disappears. This means feedback inhibition has to be fast enough (less than 0.3ms) to generate a plateau potential instead of a peak.

![Surface plot of membrane potential with different feedback inhibition strengths.](image)

Figure 2-40: surface plot of membrane potential with different feedback inhibition strengths. The model use single neuron with parameter set C except feedback inhibition. The inset is the top view of the surface plot and the red line specify the feedback inhibition timing in parameter set C. The synaptic input is as table:

**Table 2-14**

<table>
<thead>
<tr>
<th>timing</th>
<th>Synaptic input</th>
</tr>
</thead>
<tbody>
<tr>
<td>5ms</td>
<td>13</td>
</tr>
<tr>
<td>variable</td>
<td>-30</td>
</tr>
</tbody>
</table>
2.8.10.4 Variable feedforward excitation

When plateau potential is all set up, feedforward excitation has to be strong enough to elicit a spike. As the strength increase, the firing timing approach to feedforward input timing. The working line (red) is chosen to just intersect the asymptotic line, which means the spike is elicited immediately after feedforward input arrives.

Figure 2-41: surface plot of membrane potential with different feedforward excitation strengths. The model use single neuron with parameter set C except feedback inhibition. The inset is the top view of the surface plot and the red line specify the feedforward excitation in parameter set C. The synaptic input is as table:

Table 2-15

<table>
<thead>
<tr>
<th>timing</th>
<th>Synaptic input</th>
</tr>
</thead>
<tbody>
<tr>
<td>5ms</td>
<td>13</td>
</tr>
</tbody>
</table>
### 2.8.11 Learning with repetitive sensory inputs

Axon withdraw can happen on both pre-synaptic neuron and post-synaptic neuron. In order to have more than super-synapses for one sensory neuron, axon withdraw model on target neurons is applied in our learning simulation, which accommodates the learning of a sequence with repeated inputs like Figure 2-37C.

<table>
<thead>
<tr>
<th>Time</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3ms</td>
<td>-30</td>
</tr>
<tr>
<td>20ms</td>
<td>5.5</td>
</tr>
<tr>
<td>Firing time+0.3ms</td>
<td>-30</td>
</tr>
</tbody>
</table>

Figure 2-42: the learning process of repeated input sequence 12245 of synfire chain composed of 5 groups with the population size as 5. (A): the raster plot of learning process of 5 groups in 282
Figure 2-42 shows the learning process of sequence 12245: the raster plot and synaptic growth show similar pattern with the learning without repeated inputs. Actually, the super-synapses (red lines) projecting to group 3 grow faster than other super-synapses, because the up state of group 3 is higher than the up state of other groups.

2.9 Conclusion

1) Synfire chain is quite robust to recognize specific sensory sequence input in presence of synaptic noise. The recognition is based on coincident detection of lateral excitation and feedforward excitation bridged by a carefully constructed plateau potential in synfire chain neuron.

2) The specific sequence can be encoded in the specific wiring from sensory neuron to synfire chain neurons. The encoding process can be achieved by a voltage dependent learning algorithm. This learning is robust to not only synaptic noise but also various defects of input sequence (random deletion, random replacement, decaying). The learning rate is invariant to synfire chain length and the population size of synfire chain group.

3) In order to accommodate sequence with repetitive inputs, A-current has to be introduced in single neuron model to suppress the current surge which is caused by repetitive inputs. This A-current activation region has to be limited between down state (resting potential) and up state (plateau potential) to avoid the weakening of overall excitability.
4) The firing threshold of single compartment model limits the role of a more realistic A-current. A two-compartment model without firing threshold on dendrite can leave more room for A-current to enhance the capacitance to current surge.
Chapter 3  Sequence recognition and sequence learning based on two-compartment model

3.1 Introduction

In last chapter, we discuss how synfire chain network can work with the plateau potential of single neuron to recognize a specific sensory sequence also how to encode the sequence by synaptic plasticity (sensory learning process). There are several inherent problems:

First, the computation starts with manual made sensory sequence, so it is hard to compare this computational scheme with biological system receiving physical stimuli. In this chapter, we try to make the first step to approach this problem by analyzing the song of zebra finch, a highly structured vocalization and generate sensory sequence from them.

Secondly, A-current plays an important role to enhance the robustness of coincident detection seen from the example of chapter 2; however, we have to limit the activation region of A-current with the constraints of firing threshold and resting potential to avoid the overall excitability suppressed. Also the neuron model is an over simplification of biological neuron in the aspect that we omit the spatial structure of biological neuron and pack all the functionalities into one integration unit (soma). A quick fix is to use two-compartment model which describes the spatial structure of biological neuron in the simplest form: a soma and a dendrite. The soma carries information integration and action potential generation and the dendrite collects and preprocesses all the sensory information. Since dendrite is not involved in action potential generation, its membrane potential could have broader variation than the membrane potential of soma, thus leave a greater room for A-current to play its role.

Finally, in chapter 2, the formation of the plateau potential needs balanced and an excitation and a closely following inhibition, mediated by a fast feedback interneuron. This mechanism shows some level of susceptibility to various parameters including synaptic strengths,
noise, etc. Here we propose another way to produce plateau potential, which is based on the intrinsic properties of HVC_{RA} neurons of zebra finch. We use two-compartment model to model this neuron. Once it fires, it always gives a burst of 4-6 spikes which lasts around 10ms. This leads to the plateau potential of its downstream neuron, which is not dependent on interneuron.

3.2 Methods

3.2.1 Front end model

The auditory pathway is a hierarchical structure composed of three parts: the front end, the sensory neurons (Nif) and HVC local circuit. Front end includes all the peripheral auditory sensory process and translates auditory signal to sound features. Nif is the primary auditory input to HVC nucleus and is responsible to translate different sound features arriving at different time into spatiotemporal spike sequences. Finally HVC local network is up to read the sequences sent from Nif and recognize them.

The front end model includes the neural pathway before Nif in the sound processing pathway. It detects the sound pressure wave and analyzes it in frequency-time space then generates the key features finally categories them according to the library.

The sound wave is tuned into spectrogram along the auditory pathway before it reaches Nif. We want to extract the acoustic feature from the spectrogram within a short time window comparing with the song length, but still long enough to generate specific pattern. For some reason, we set this short time window as 20ms. In order to extract the acoustic feature within a short time window, we smooth the spectrum in each short time window, thus we get the slow varying acoustic signal. To reduce the data further, we do the Fourier transformation on smoothed acoustic signal, we find it is a reasonable approximation to represent the smoothed spectrum only
by its first 100 FFT coefficients, thus we use the first 100 coefficients as the acoustic feature vectors to represent the spectrogram in that short time window. We manually cluster the feature vectors in all time windows into 16 categories. Each category shares some common feature shown in the examples of each category. We can represent all the acoustic features in the same category only by their category ID. Now we could represent the song spectrogram by feature sequence. The whole process is borrowed from speech recognition except the smooth part. We think the smooth part is a procedure to reduce noise.

The whole process starts from the raw song representation: the pressure wave.

### 3.2.1.1 Spectrogram

The sound pressure wave is sampled with 40 kHz frequency shown in Figure 3-1 top panel. There is not too much information from pressure wave except the relative sound magnitude. In order to get the structure of sound in frequency space, short-time fast Fourier transformation (SFFT) is applied to sound wave. In SFFT, the wave signal is analyzed in a short time window, which is short enough to assume frequency distribution in that window not varying too much. If we concatenate frequency distribution in each window, specific pattern appear in this two dimensional (time domain and frequency domain) representation of song (spectrogram). Our spectrogram (Figure 3-1 bottom panel) is calculated with a 20 ms time window sliding over the song wave, and the shift of the sliding window is 4ms. In order for better contrast, the epsilon 8e-4W/Hz is added to the raw spectrogram so anything value smaller than 8e-4 W/Hz should be around 8e-4W/Hz. Because the color bar is distributed evenly between maximum value and minimum value, the plot has better contrast with epsilon. The logarithm of the spectrogram is calculated in each window. In order to get the decibel, the baseline is set as 2e-5 W/Hz, which is
the standard value to calculate sound level. Normally, zebra finch’s song spans up to 8 kHz, thus we discards the frequency components higher than 8 kHz.

![Figure 3-1: raw sound pressure wave and song spectrogram.](image)

From spectrogram, the song of zebra finch shows a well structured vocalization. It usually begins with a few introductory notes, then followed by repeating pattern called motif lasting around 500ms. Typically each motif comprises of 3-5 distinct vocal structures, which is called syllable. The whole rendition of the song is called bout.

### 3.2.1.2 Generate feature vector

For the spectrogram in a 20 ms window, the frequency domain has 161 components spanning from 0 to 8 kHz with an increment of 50 Hz (the inverse of the window time). Our goal is to extract the acoustic feature in 20 ms window, which can represent different patterns within 20 ms in the bird’s song. Then syllable and motif can be represented by a sequence of these
features. In order to get feature vectors in this window, these 161 components need not to be present because there is clear structure for frequency distribution such as harmonic stack. For the acoustic feature vectors: it should be relatively slow-varying in frequency domain. In order to get acoustic feature, we do a smoothing algorithm on each slice of data.

![Smooth spectrum](image)

Figure 3-2: smooth spectrum to get acoustic signal. The bottom spectrum is one slice sample of the whole spectrogram. The window is 20ms, and the frequency spans from 0 to 8000 Hz. Above is the plot of the log power vs. frequency. The red line is the original signal and the blue line is the smoothed signal by the Savitzky-Golay method with a span of 16 points.

In Figure 3-2, we plot a slice of the spectrogram in frequency domain, and the smoothed curve represents the slow-varying components. The smoothed signal captures the acoustic features of the original signal and reduces the high frequency noises.
Second we want to pick up those features which are most relevant to slow-varying components. Thus we overlap the FFT of all slices of a song spectrogram in one plot. And we don’t want to lose too many details of the original spectrogram.

![FFT of Raw Spectrum](image)

![FFT of Smoothed Spectrum](image)

Figure 3-3: 100 FFT coefficients as feature vectors. The top panel is the FFT of the raw spectrum overlapped in many windows, the bottom one is the FFT of the smoothed spectrum overlapped in many windows. The smoothed FFT decays quickly to zero around 100, so we adopt the first 100 FFT coefficients of the smoothed signal as feature vectors.

Figure 3-3 is a comparison of the overlapped FFT of original spectrogram and the overlapped FFT of the smoothed spectrogram. We see the smoothed FFT resembles the original in the overall shape, and quickly decays to zero after 100. So we use the first 100 FFT
coefficients of each slice of the spectrogram as the feature vector to represent the acoustic feature in that 20ms window.

### 3.2.1.3 Clustering feature vectors and generates the symbol

Now we get the feature vectors which represent most of the acoustic feature in the spectrogram. We ask the question how bird’s brain comprehends these feature vectors then how to assemble them to various meaningful syllables. We propose that the bridge between acoustic features and syllable is the generation of the conceptual symbol. The symbol is abstracted from a bunch of feature vectors which are clustered in the nearby vector space. So the symbol can be thought of the category identification of the feature vectors. Practically we cluster a pool of feature vectors, which are from bird’s own song and other con-specific songs, then generate the base vectors by taking the mass centroid of all the vectors in the same category. The exact procedure is shown in Figure 3-4, which follows the vector quantization method taken in speech recognition\textsuperscript{118}. 

The reason we want to use not only birds’ own song but also other con-specific songs, we assume this will give us a much broader sample space, because the wild birds hear various acoustic signal not only its tutor song but also other songs. We assume this broader sample contains all the possible feature vectors that could contribute conceptual symbol in the bird’s brain. Once we have the base vectors for each category, the symbol is generated by comparing a Euclidean distance measure the feature vectors with all the base vectors, finding the closest one. This algorithm is limited to the category number of 2, 4, 8, 16…., which are the powers of 2. In the following work, we cluster all the song vocalization into 16 categories. But before we progress to the next stage, we still want to check the reliability of the symbols. If two spectrums belong to the same symbol, they should look similar or share some common feature. First, we did the test on two-dimensional space shown in Figure 3-5. Since the categories that this algorithm can divide

Figure 3-4: vector quantization algorithm flowchart. Adapted from page 54.
the data into have to be the power of 2. We want to see if this affects the clustering. There are 5 Gaussian clusters in Figure 3-5 A and B. In A, the category number is 4 and in B, the category number is 8. We can see that if the category number is less than cluster number, some cluster is not included in the categories such as the central cluster in Figure 3-5A. If the category number is larger than cluster number as shown in Figure 3-5B, all the clusters can be represented. Notice that this result only applies on the data points all of which are from clusters with the same level of distribution. In reality, data points do not always form clusters such as shown in Figure 3-5C and D. The data points are the first two components of the feature vector from a real song. C shows the prototype vectors as blue (green) dots for 8 (16) categories. D shows the prototype vectors as red dots for 32 categories. Since quite a few of data points are not concentrated around a center, but rather form a linear distribution (in higher dimension, maybe form other shapes). There is no big difference with 16 or 32 categories to represent the real data from the view of two-dimension.
Figure 3-5: The effectiveness of the clustering algorithm on 2D space. In A and B: there are 5 2D Gaussian clusters (red dots). The blue dots in A (B) give the prototype vectors of 4 (8) categories. In C and D, the first two components of the feature vectors of a bird song are represented by black dots. The blue (green) dots in D give the prototype vectors of 8 (16) categories. The red dots in D give the prototype vectors of 32 categories.

Figure 3-6 lists the spectrum of a few samples in each category. We found the spectrums in the same category they do share some common features. But the difference among categories is not as prominent as expectation.

Figure 3-6: Acoustic feature clustering result. This is an example to show the validity of the feature vectors clustering algorithm. Each column shows the sample of the original spectrum in 20ms belonging to one category.

Once the symbol is generated, the whole bout (motif, syllable) spectrogram is represented by a sequence of symbols. This is the input to HVC area. Figure 3-7 shows a symbolic sequence representation for three renditions of the same motif. We can see the variability of the same motif is still large.
The variability of symbol sequence can be handled in recognition problem with adding wiring to HVC neurons; but causes severe problem in learning with real song. Because in real song, the same syllable (motif) has variability, and this variability cannot be accommodate naturally in our model; as we get some insight from the mixed training in last chapter, training result could not give wiring structure capable of recognize both of them. Therefore enhancing the robustness of sequence generation process is required to bridge the gap between real song training and sequence training. Alternatively, we try to use gammatone decomposition to generate sound features.

Figure 3-7: The symbolic symbol representation of three renditions of the same motif. Green dots are the symbols at different moments. Bottom panel is the corresponding spectrogram.
3.2.2 Nif model

Considering Nif is the main auditory input of HVC\textsuperscript{19}, we propose that Nif neurons are responsible for detecting these features. There are different types of Nif neurons, which detect the different types of features so that when a specific feature comes in, the corresponding Nif fires and projects the spike to its target HVC neuron;

Without the detailed data on Nif, we model Nif conceptually. Nif receives information about features generated from front end system, and turns them into a symbol sequence; then HVC network receives the feed-forward inputs from Nif and recognizes the symbol sequence.

We use binary neuron to model Nif neurons, which send spatiotemporal spike sequence to the synfire chain neurons. Here the term “spatial” means different kinds of features and the term “temporal” means the temporal order of each feature. Nif neuron just gives a spike at a specific time and sends it out to its target neuron and remains silent at other time.

3.2.3 Network

HVC local network is modeled by groups of neuron projecting to one after another in an all-to-all style, the same as the synfire chain structure in Chapter 2. Two-compartment model is used to model HVC\textsubscript{RA} neuron. All excitations (feedforward and lateral) and inhibition target on dendrites as specified in Figure 3-9; the main function of the soma is to synthesize the information from dendrite and generate action potential and send to the downstream neurons.

As an illustration, Figure 3-8 shows a synfire chain with three groups. Each group is composed of two neurons, which means there are two excitatory lines regulated by two interneurons respectively. Each line receives the same feedforward input from Nif sensory
neurons. All synapses are denoted by small circles: solid ones are inhibitory and circles are excitatory. The synaptic strengths of Figure 3-8 are specified as the following tables.

Table 3-1

<table>
<thead>
<tr>
<th>Color</th>
<th>Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red lines</td>
<td>Feedforward connections</td>
<td>$G_{ef} = 0.3$</td>
</tr>
<tr>
<td>Green lines</td>
<td>Lateral connections</td>
<td>$G_{el} = 0.04$</td>
</tr>
<tr>
<td>Purple lines</td>
<td>Excitation to interneuron</td>
<td>$G_{ei} = 1.0$</td>
</tr>
<tr>
<td>Blue lines</td>
<td>Feedback inhibition</td>
<td>$G_{ib} = 0.04$</td>
</tr>
</tbody>
</table>

Figure 3-8: scheme of synfire chain network with feedforward sensory inputs. All lateral synaptic connections are denoted by light gray arrows. Each chain receives the same feed-forward...
inputs from Nif, which are denoted by dark gray arrow.

Fast feedback interneuron is kept in our model as in the network. It does not involve into the plateau potential directly. It serves as two purposes: one is to mimic more realistic environments and test the robustness of this mechanism; also we hope to regulate parameters of the network.

3.2.4 Neuron Model

3.2.4.1 $HVC_{RA}$ neuron model

![Diagram of HVC neuron model](image)

Figure 3-9: the two-compartment model of $HVC_{RA}$ neuron. The colorful lines are synaptic connections.

The HVC neuron is modeled by a two-compartment model. This model adapted from Jin’s paper \(^{34}\) to model the sequence generation in HVC. This single neuron model is embedded with intrinsic spike burst also has a long refractory period.
\[
\frac{dU_s}{dt} = \left[ (I_{LS} + I_{Na} + I_K + I_{KHT} + I_{KLT} + I_{FF} + I_{sin}) \cdot A_s + I_{ext} + I_{ds} \right] / (C_m \cdot A_s);
\]

\[
\frac{dU_d}{dt} = \left[ (I_{Ld} + I_{Ca} + I_{CaK} + I_A + I_{syn} + I_{din}) \cdot A_d + I_{des} + I_{sd} \right] / (C_m \cdot A_d);
\]

\(U_s\) is the subthreshold membrane potential on the soma in mV; \(U_d\) is the subthreshold membrane potential on the dendrite in mV.

### Table 3-2

| \(I_{ls}\) = \(g_{ls} \cdot (E_r - U_s)\) | the leaky current on soma |
| \(I_{Na}\) = \(g_{Na} \cdot m^3 \cdot h \cdot (E_{Na} - U_s)\) | the Sodium current on soma |
| \(I_K\) = \(g_K \cdot n^4 \cdot (E_K - U_s)\) | the Potassium current on soma |
| \(I_{KHT}\) = \(g_{KHT} \cdot w \cdot (E_K - U_s)\) | the high threshold potassium current on soma; |
| \(I_{KLT}\) = \(g_{KLT} \cdot I \cdot (E_K - U_s)\) | the low threshold potassium current on soma; |
| \(I_{Ld}\) = \(g_{Ld} \cdot (E_r - U_d)\) | the leaky current on dendrite; |
| \(I_{Ca}\) = \(g_{Ca} \cdot m^2_e \cdot (E_{Ca} - U_d)\) | the Calcium current on dendrite |
| \(I_{CaK}\) = \(g_{CaK} \cdot q \cdot (E_K - U_d)\) | the Calcium dependent Potassium current on dendrite; |
| \(I_A\) = \(g_A \cdot a^2 \cdot b \cdot (E_K - U_d)\) | the inhibitory potassium A current; |
| \(I_{syn}\) = \(-g_{des} \cdot U_d\) | the excitatory current on dendrite; |
| \(I_{din}\) = \(g_{din} \cdot (E_I - U_d)\) | the inhibitory current on dendrite; |
3.2.4.2Multicurrent integrate-fire neuron

Multicurrent integrate-fire neuron provides feedback inhibition (i1) to synfire chain neurons. The inhibition is triggered by the firing of the synfire chain neurons. This model follows the modification proposed by Lewis and Gersnter\(^{109}\) from conductance based interneuron model by Erisir\(^{110}\). This model has fast deactivating potassium channel and simple threshold and reset mechanism and can model fast spiking interneuron.

\[ C \frac{dU}{dt} = -I_L - I_{Na} - I_{K_1} - I_{K_3} - g_L U - g_I (U - E_I) \]

Table 3-3 parameters for multicurrent integrate-fire neuron

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C = 1\mu F / cm^2)</td>
<td>the membrane capacitance</td>
</tr>
<tr>
<td>(I_L = g_L (U - L))</td>
<td>the leak current</td>
</tr>
<tr>
<td>(g_L = 1.25 \mu S / cm^2)</td>
<td>the leak conductance</td>
</tr>
<tr>
<td>(I_{Na} = g_{Na} m^3 h (U - E_{Na}))</td>
<td>the (Na^+) current</td>
</tr>
</tbody>
</table>
The dynamical equation of the gating variables

\[ \frac{dy}{dt} = \alpha_y(U) \cdot (1 - y) - \beta_y(U) \cdot y, \quad y = m, h, n, p \]

With the parameters are:

Table 3-4 dynamical parameters for gating variables

| \( g_{Na} = 112.5 \mu S / cm^2 \) | the \( Na^+ \) conductance |
| \( E_{Na} = 74 mV \) | \( Na^+ \) reversal potential |
| \( I_{K1} = g_{K1} n^3 (U - E_K) \) | the Kv1.3 current |
| \( g_{K1} = 0.225 \mu S / cm^2 \) | the Kv1.3 conductance |
| \( I_{K3} = g_{K3} p^2 (U - E_K) \) | the K3.1-Kv3.2 current |
| \( g_{K3} = 225 \mu S / cm^2 \) | the K3.1-Kv3.2 conductance |
| \( E_K = -90 mV \) | the \( K^+ \) reversal potential |
| \( E_I = -75 mV \) | the reversal potential |

The parameters are:

- \( \alpha_n = (3020 - 40U) / \{ \exp[-(U - 75.5) / 13.5] - 1 \} \)
- \( \beta_n = 1.2262 / \exp(U / 42.248) \)
- \( \alpha_m = 0.0035 / \exp(U / 24.186) \)
- \( \beta_m = -(0.8712 + 0.017U) / \{ \exp[-(U + 51.25) / 5.2] - 1 \} \)
- \( \alpha_h = -(0.616 + 0.014U) / \{ \exp[-(U + 44) / 2.3] - 1 \} \)
- \( \beta_h = 0.0043 / \exp[(U + 44) / 34] \)
- \( \alpha_p = (95 - U) / \{ \exp[-(U - 95) / 11.8] - 1 \} \)
- \( \beta_p = 0.025 / \exp(U / 22.222) \)
Other parameters in this model are:

Table 3-5 other parameters in multicurrent integrate-fire neuron model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta$</td>
<td>$-30mV$</td>
<td>spiking threshold</td>
</tr>
<tr>
<td>$R$</td>
<td>$-85mV$</td>
<td>reset potential</td>
</tr>
<tr>
<td>$\tau_g$</td>
<td>$1.5ms$</td>
<td>refractory period</td>
</tr>
</tbody>
</table>

And the gating variables m, h, n, p are reset to 0, 0.16, 0.88, 0.2 respectively after each spike.

3.2.5 Synapse model

3.2.5.1 Excitatory synapse

The excitatory synapse is modeled by a kick-and-decay algorithm. When there is presynaptic spike, the excitatory synaptic conductance of the post-neuron jumps by an amount $\Delta G$, then it decays exponentially ($\frac{dg_{ex}}{dt} = -g_{ex} / \tau_e$), the decaying time constant $\tau_e$ is 5ms.

3.2.5.2 Inhibitory synapse

The inhibitory synapse is modeled by a kick-and-decay algorithm. When there is presynaptic spike, the inhibitory synaptic conductance of the post-neuron jumps by an amount $\Delta G$, then it decays exponentially ($\frac{dg_{in}}{dt} = -g_{in} / \tau_i$), the decaying time constant $\tau_i$ is 5ms.
3.2.6 Recognition model

When the sound signal passes the Nif area, the song is transformed into symbol sequence, each symbol spaces about 20ms. The HVC takes these sequences as input and try to recognize the very sequence implied by its specific wiring from Nif to HVC.

Assume the start HVC neuron is always giving a spike burst whenever the pre-synaptic Nif neuron fires. With the specific wiring from Nif to HVC, when the desired sequence is presented, the following HVC neuron is able to get the lateral input and feed-forward input coincidently, thus it will relay the spike to the next HVC neuron. So the spike propagates along the synfire chain. If the sequence inputted from Nif is not the desired one, some or all HVC neurons will not receive the lateral and feed-forward inputs at the same time window, this only results a depolarized membrane potential. The spike propagation will finally fail somewhere in the chain. In the sense of spike propagation, the desired sequence is recognized.

3.2.7 Plasticity model

3.2.7.1 STDP algorithm

The LTP and LTD for each pair of neurons depends on the time difference of the current spike and all the spikes before it. For example, neuron m fires at \( t_m \). Another spike is fired before \( t_m \) by neuron n. If n is the pre-synaptic neuron of m, then the synapse from n to m is strengthened by long term potentiation (LTP), because the pre-synaptic neuron fires before the post-synaptic neuron. If n is the post-synaptic neuron of m, then the synapse from m to n is weakened by long term depression (LTD), because the pre-synaptic neuron fires after the post-synaptic neuron.
The exact change of LTP is determined by the maximum percentage change $A_{LTP}$, the original synaptic strength $G_{nm}$, and the curve factor $P(\Delta t)$. The standard value of $A_{LTP}$ is 0.5%, however in our simulation, it can be adjustable to expedite learning rate. The factor $P$ depends on time difference ($\Delta t = t_m - t_n'$) of the spike of neuron m and the spike of neuron n. The time difference is definitely positive, because $t_m$ is always the current time. When the time difference $\Delta t$ is smaller than 5ms, $P$ factor increases linearly from 0 to 1. After 5ms, $P$ decays exponentially with a time constant as 20ms. Final LTP change factor is calculated by summing up all the $P$ factors caused by all previous spikes of neuron n, thus the synapse from n to m is updated.

$$G_{nm} \rightarrow G_{nm} + A_{LTP} \cdot G_{nm} \sum_{i} P(t_m - t_n')$$

$$P(\Delta t) = \frac{\Delta t / 5}{e^{-(5\Delta t - 5)/\tau_{LTP}}} \text{ if } \Delta t \geq 5ms$$

The exact change of LTD is determined by the maximum percentage change $A_{LTD}$, the original synaptic strength $G_{mn}$, and the curve factor $D(\Delta t)$. The standard value of $A_{LTD}$ is 0.525%, however in our simulation, it can be adjustable to expedite learning rate. The factor $D$ depends on the time difference ($\Delta t = t_m - t_n'$) of the spike of neuron m and the spike of neuron n. The time difference is definitely positive, because $t_m$ is always the current time. As the time difference $\Delta t$ is smaller than 5.25ms, $D$ factor increases linearly from 0 to 1. After 5.25ms, $D$ decays exponentially with a time constant as 20ms. Final LTD change factor is calculated by summing up all the $D$ factors caused by all previous spikes of neuron n, thus the synapse from n to m is updated.

$$G_{mn} \rightarrow G_{mn} + A_{LTD} \cdot G_{mn} \sum_{i} D(t_m - t_n')$$
\[ D(\Delta t) = \begin{cases} \frac{\Delta t}{5.25} & \text{if } \Delta t \leq 5.25 ms \\ e^{-(\Delta t - 5.25)/\tau_{ms}} & \text{if } \Delta t \geq 5.25 ms \end{cases} \]

3.2.7.2 Axon remodeling

All synapses projecting from sensory neuron to sensory neuron start from all-to-all weak connections. The synaptic strength evolves depending on the training sequence repeating in every trial. When one of Nif neuron’s synapses reaches the super-synapse threshold, this synapse becomes the super-synapse, and the Nif neuron then withdraws all its other synapses. The withdrawn synapses are like cut-off from its targets. They don’t evolve with STDP mechanism, but they do experience homogeneous decaying. This process is totally reversible. Because the super-synapse still undergoes STDP mechanism, if the super-synapse drops under the threshold, it is no longer a super-synapse, and all the other withdrawn synapses become alive again and evolve with STDP mechanism.

3.2.7.3 Homogeneous synaptic decay

There is homogeneous decaying on all the synapses from Nif to HVC neuron. At every time step, all synapse becomes $\beta$ times of itself (\( G_{nm} \rightarrow \beta \cdot G_{nm} \)). The decaying allows the memory of previous synaptic strength gradually diminishes. In practice, it is also useful to check the spontaneous activities; otherwise it is impossible to form a very long chain.
3.2.8 Noise model

The spontaneous activity is driven by the synaptic noise on soma and on dendrite of all HVC neurons. The synaptic noise includes both excitatory noise and inhibitory noise. The noise spikes are a Poisson spike train. The frequency of noise spike train is 100Hz, the amplitude is adjustable in recognition, and is fixed at 0.04mS/cm$^2$ in learning. The resulting standard deviation around resting potential is about 5mV.

3.3 Single neuron firing probability

Noise is an inevitable part during recognition and learning in reality. We want to study the firing probability of a single neuron with noise at different membrane potential, when it is excited by different injected current.

Figure 3-10 gives the single neuron excitability with both excitatory and inhibitory noise present. The x-axis is the membrane potential from -85 to -60 mV, which is pushed by injected current. The y-axis is maximum noise synaptic input amplitude for both excitatory and inhibitory noise. The noise is injected as a Poisson spike train with frequency as 100Hz, whose magnitude is chosen randomly from 0 to the maximum noise synaptic input. Higher membrane potential and intense noise will increase the firing probability of the HVC neuron. In our recognition and learning scheme, the membrane potential is a result of integrating all inputs in a short time window; so higher membrane potential by coherent inputs gives a higher firing probability. A is measured over a long run (200ms after noise is introduced). B is measured over a short run (30ms after noise is introduced). There is steep jump around -65mV in both plots when noise is 0; we can regard it as a threshold to discriminate the UP and DOWN state, noise just smears the shape of the state transition. In order to make the desired synapse to grow, we need the noise to elicit
spike when the synapse is immature; on the other hand, the noise should not be too strong to force the HVC neuron to spike when its membrane potential is far away from the UP state as shown in panel A.

For both panels, a firing threshold is presented. The noise should be limited in such region that it only affects the firing probability when the membrane potential is around UP state in a short period statistics and does not have significant impact on firing probability when the membrane potential is far away from the UP state.

![Figure 3-10: spontaneous firing probability at different membrane potential. Panel A is the firing probability measure for 200ms and panel B is for 30ms](image)

### 3.4 Coincidence detection phase diagram

In order to incorporate the sensory information into neural network as past experience, the smallest unit doing this job is a single neuron. There are two kinds of synaptic inputs projecting into the HVC-RA neuron in the synfire chain: the feedforward inputs from Nif (external) and lateral inputs from HVC network (inside). Coincidence detection of these two kinds of synaptic inputs is essential to integrate the sensory information into HVC network. In order to determine the parameter region where coincident detection works, the phase diagram of
the timing of elicited spike is plotted against the strength of lateral input and feedforward input. In Figure 3-11, lateral input (occurring at 5ms) precedes feedforward input (occurring at 25ms) by 20ms, both of which are tuned by the magnitude. Different colors mean the timing of elicited spike if there is one. On the right, the color bar specifies the timing in millisecond. Mainly, there are two area separated by a jigsaw curve: the yellow area is the parameter region where there is a spike after feedforward input (25ms) and the blue area is the parameter region where there is no spike at all after feedforward input arrives. The vertical and horizontal dashed lines specify the parameters where a spike can be elicited by only lateral input or feedforward input. So the dashed lines and the jigsaw curve enclose the parameter region where coincident detection works toward a spike elicited. Working in this region, every HVC neuron will fire only if both lateral input and feedforward input are present coordinately. The recognition based on coincident detection happens in this region, and the synaptic strength in this region is also the goal for the learning mechanism. The HVC neuron will fire if either of the feedforward input or the lateral input is represented when the parameters is outside the dashed box. In the blue area, HVC neuron will not give a spike.

In Figure 3-11A, the jigsaw curve is like a step down function with the increase of lateral strength, which means the excitability of HVC neuron is prominently lowered by lateral input. Considering the flatness of the curve when lateral strength is higher than 0.04, the jigsaw curve can be regarded as bi-stable state regarding to the excitability of HVC neuron. We call UP state where the lateral strength is larger than 0.04, because the HVC neuron is really sensitive to feedforward input since the lateral input pushes the membrane potential UP. The DOWN state is the region where lateral strength is smaller than 0.04, because the HVC neuron will not give a spike unless feedforward input is much larger than the magnitude needed for UP state. According to the shape of the curve, the optimized relative strength of these two inputs can be chosen to achieve coincident detection but both are small. For learning, the lateral connections are already
made, and feedforward connections are all-to-all weak connections, which cannot elicit a spike by itself. However with the help of noise, the HVC neuron will jump across the curve into the coincidence detection region thus it will elicit the growth of the desired synapse from Nif to HVC.

For comparison, we plot the coincident detection phase diagram without A-current in Figure 3-11B. We can see how coincident detection region is shrunk in the aspect of both the magnitude of the dashed lines and the shape of the jigsaw curve (The stepwise curve is caused by discretization). The first means the threshold of coincident detection is lowered without inhibitory A-current, and the second indicates the distinction of UP and DOWN state is smeared without A-current.

Figure 3-11: Coincident detection phase diagram. The scale bar on the right represents the spike time of the HVC neuron. (A) is the model with A-current, $g_A=1$; (B) is the model without A-current.
3.5 Sequence recognition

3.5.1 Synfire chain with neuron groups

In order to recognize a sequence, there must be specific wiring from Nif to HVC. The length of the sequence determines the length of HVC synfire chain. The interval between symbols is 20ms. There is a corresponding Nif firing for each symbol in the sequence. Since HVC neurons are sequentially connected with each other, there is a HVC neuron associated with that time window in which one Nif neuron fires. If the firing Nif neuron projects to the HVC associated with that time window, the HVC neuron takes the input from Nif firing at that time window. So the wiring is totally determined by the sequence. The ID of symbol determines the pre-synaptic neuron, and the order of the symbol determines post-synaptic neuron. Nif has the same specific connection to all the chains. When there is sound signal coming in, Nif emits feed-forward output to HVC synfire chain network.

Based on synfire chain network model, HVC could recognize a specific sequence, and reject other sequences. The sequence is delivered from Nif by analyzing the spectrogram feature in a short time window. The recognition depends on the synfire chain network in HVC as well as the characteristic wiring from Nif to HVC, so that each neuron in synfire chain could get the lateral input and feed-forward input at the same time window, and the spike burst propagates along the chain.

Figure 3-12: Network to recognize serial sequence input 1 2 3 4 5 6 7 8 9. In order to recognize
the symbol sequence 123456789, each spaces 20ms.

According to the coherence detection, the network is composed of locally excitatory chain of HVC and the specific wiring from the Nif neuron to its target HVC neuron. All the following feed-forward synaptic strength and lateral synaptic strength are tuned so that either one of them cannot make the HVC neuron fires; HVC neuron will not fire until both feed-forward input from Nif and lateral input from its predecessor are present in a short time window. The first synapse from Nif1 to HVC1 is specially set so that HVC1 can fire with only feedforward input.

Figure 3-12 is an example of network for the recognition. The symbol sequence intended to recognize is 123456789. In order to recognize this symbol sequence, the network must have the following structure as shown in Figure 3-12. According to the coherence detection scheme, the HVC neuron will not fire until the feed-forward input from Nif neuron meets the lateral input from its predecessor in a short time window.
When the input sequence is 123456789, X axis is time and each row represents the histogram of the spikes of a HVC neuron. The network is shown in Figure 3-12. Without noise, this network only recognizes the sequence 123456789 and rejects all other combinations, the HVC neuron doesn’t respond at all. This network only recognizes 123456789, which means the HVC neurons will fire sequentially one after another as shown in Figure 3-13A, and rejects all other sequences, which means there is no response for all HVC neurons. Without any noise, this discrimination is definite. In realistic world, noise is everywhere. We have to consider the impact of noisy input to our recognition scheme. We inject Poisson spike trains to the synapses of both

Figure 3-13: spike propagation of synfire chain. (A) Spike propagation along a single excitatory line without any noise. (B) Spike propagation along a single excitatory line with noise in 50 trials. (C) Spike propagation along a synfire chain with 30 excitatory lines in single trial. (D) Spike propagation along a synfire chain with 30 excitatory lines in 50 trials. (E) No spike propagation along synfire chain without lateral connections. (F) Spike propagation with repetitive inputs.
dendrites and somas. Excitatory and inhibitory noises both have a frequency of 100Hz. The final results will give a 2-3mV standard deviation around resting potential.

With noisy synaptic input on both soma and dendrite, both excitatory and inhibitory with a frequency 100Hz, which results 2-3mV standard deviation around resting potential. Figure 3-13B is raster histogram for HVC neurons in 50 trials. Overall, the spike pulse packet gradually loses its shape. Eventually there is a little chance for a spike to propagate to the last HVC neuron group in the chain.

As our expectation, the recognition based on single excitatory chain is not deterministic, as shown in Figure 3-13B, a Gaussian spike pulse packets instead of spike bursts still propagate along the excitatory chain with multiple trials. We can see along the chain, the pulse packet gradually degrades, so in the sense of single trial, the recognition is not robust, since there is a little chance that the sequential spike propagation is not destructed by random noise. In order to be more resistant to random noise, the excitatory chain is duplicated to several copies, each copy receive the same feed-forward input and is laterally interconnected with each other, as we show the scheme in Figure 3-13C.

This raster histogram in Figure 3-13C is based on 30 duplicated excitatory lines interconnecting with each other. As before each line has 9 HVC neurons laterally connected. Noise is as the plot above. The histogram is generated on single trial; we count the number of spikes of HVC neurons at the same position in a chain from multiple chains. Obviously, even in single trial sense, the multiple chains model is more robust than single chain, because there is either spike propagation or fail somewhere in the chain. It is obvious that the spike pulse packet not only propagates along the chain but also keeps the bell shape all the way. The more lines interconnecting, the spike bursts propagate more steadily, because during propagation, the spike pulse packets almost keep the same bell shape without degradation.
Figure 3-13 D is the histogram for each neuron for 50 trials. The spike pulse packet shows a Gaussian shape, and the packet propagates along the chain. According to this picture, this structure can be called as a synfire chain.

For the sequences other than 123456789, the spike or spike pulse packet will not propagate along the chain. For example, if we reverse the song, thus the input from Nif is 987654321 instead of 123456789; there is very little response from HVC neurons as shown in Figure 3-13E, because as each input comes, there is no coherent lateral input thus results no firing. Actually, if we get rid of the lateral connections, the network will lose the ability to discriminate its own song and other songs.

To verify the coherence detection idea, the lateral connections are blocked off. Even with desired input 123456789; the HVC neurons only give very little response except the starter HVC neuron. Thus the recognition depends equal importantly on HVC local lateral excitatory network and the specific wiring from Nif to HVC.

Figure 3-13E shows very little response even we present the sequence 123456789. For the real birds’ song, the symbol sequence input is not necessarily to be like 12356789. Let’s check one syllable in symbol, which is 6 15 14 14 16 16 16 16 7 and the corresponding network is Figure 3-14.

![Figure 3-14](image)

Figure 3-14: This network is based on a symbol sequence desired to recognize. The sequence is 6,15,14,14,16,16,16,16,7. We see there a lot of repetitions on the same symbol, so one Nif neuron may target on several HVC neurons. The synaptic connections are set as before.
We see most of sequence concentrates on several symbols, and there are some repetitions sometimes. Figure 3-13F is the raster plot of spike histogram of each HVC neuron along the chain over 50 trials. Noise is present as before. We can see this algorithm is robust to synaptic noise and to symbol repetitions as well. However, repeated symbols will increase the probability of firing before coherent inputs arrive. This still works in our scheme as shown in Figure 3-13F.

3.6 Sequence learning

The synfire chain can be formed up from neuron pool using STDP mechanism. This specific wiring can also rise up from all-to-all connections using the same algorithm trained by the sequence intended to recognize. We first talk about the spontaneous activity and two-neuron model to introduce the principle of learning, then study a chain composed of many neurons in series, finally expand the model to many layers in order to achieve robustness. We also discuss the factors that affect the learning rate and the role of axon withdrawing, and homogeneous decaying.

First all Nif neurons connect to all HVC neurons with a weak synaptic strength except the start neuron, whose connection is assumed to be already made. Second we assume the matured connection from Nif1 to HVC1 is already made, so HVC1 will fire definitely whenever Nif1 fires. We are going to show that the specific wiring can rise up from the all-to-all connections with the STDP mechanism trained by the song sequence we are intended to recognize. Although the lateral connections between HVC neurons have already formed before training, however the spikes cannot propagate along the synfire chain because the lateral strength is so chosen that the spikes from previous HVC neuron can only depolarize the next HVC neuron but not lead to firing directly. Once the HVC neuron is depolarized, spontaneous activity before training is very
important to the synaptic growth. Actually the rate of spontaneous activity determines the rate of the learning.

Before training process, the spontaneous activity of the neuron next to the start neuron is almost bounded in the time window right after the start neuron as shown in Figure 3-15A. This is beneficial to the growth if there is Nif input just at the beginning of the time window, because Nif (pre-neuron) fires before the HVC (post-neuron), thus the LTP mechanism will strengthen the connection between the Nif and the HVC neuron. We don’t want the spontaneous activity before the time window, because LTD then will work to reduce the magnitude of the desired synaptic connection. We don’t want the spontaneous activity located in the next time window either, because this will lead synaptic growth between the HVC neuron and the Nif neuron firing next, also it will reduce the contribution to the synaptic growth between the HVC neuron and the Nif neuron firing currently. So the well bounded firing (firing appearing only after sensory input and before the next sensory input, which is a 20ms window) of the HVC neuron is not only characteristic for the synfire chain network, but also the key to the formation of synaptic connections between Nif to the synfire chain. Another interesting point is only the spontaneous activities of the neuron on UP state are well bounded, and spontaneous activities of neurons on DOWN state are totally randomly distributed. So at the beginning of the training process, only the HVC neuron right after the start neuron fires well inside the next time window, the spike distribution of the following neurons loses the bounded feature quickly, the distribution will spread over all the time windows, which is totally depends on noise. Only after the super-synapse between the HVC neuron right after the start neuron and the Nif neuron has been made and there is reliable spikes emitted every time in that time window when the desired input present, then it is ready to connect the Nif to the next HVC neuron. The new connection just been made will deliver reliable spikes in that time window, which forms the bases for the following connection. Initially, we assume there are all-to-all connections between each Nif-HVC pair, and eventually the all-to-
all will connections evolve into the desired connection which is able to recognize the training sequence.

Axon withdrawing and homogeneous decaying play an important part to shape the synaptic connections. With the presence of the training sequence, the synaptic connection favorable to the sequence will be reinforced, some other synaptic connection will also be leveled up because of its random walk behavior; of course most of the other synaptic connection will keep silent because of the LTD and homogeneous decaying. Even in the very extreme situation, the desired synaptic connection will grow a lot faster than the other synaptic connections. So we set a threshold, any synapse reaching this threshold first will becomes a super-synapse, and other synapse from the same Nif neuron will be withdrawn from its targets. The withdrawal actually means the Nif neuron has already found the correct HVC neuron to connect with in order to recognize the training sequence. However, the process is reversible, if the strength of the super-synapse drops under the threshold, and all the withdrawn synapse will be recovered. The homogeneous decaying has multiple functions in the synapse growth. The decaying will control the accumulation caused by spontaneous activity, this is especially important to form a very long chain, because the learning time is so long that there is a relatively large probability for spontaneous to become runaway without decaying. Secondly, the decaying could gradually remove those withdrawn synapses, since they don’t receive the stimuli any more. Finally, considering the noisy input from Nif, we don’t want any substantial growth appearing when Nif just gave a series of random inputs, this is guaranteed by the homogeneous decaying, and otherwise substantial growth will finally accumulate to the withdrawn threshold. For the desired connection, it should not be affected too much by homogeneous decaying, because it is repeated strengthened by the firing caused by the training sequence. With STDP axon withdrawing and homogeneous decaying, we can show that the desired connection from Nif to HVC can be formed sequentially along the synfire chain network in order to recognize the training sequence.
3.7 Learning with STDP algorithm and optimal learning

3.7.1 Two-Neuron Model

![Diagram of two-neuron model](image)

Figure 3-15: The simplest model for learning. There are two Nif neurons: 1 and 2, which corresponds to symbol input 1 and 2 respectively. There are two HVC neurons: s and 2. The connection Nif1 to HVCs has already been made as well as the lateral connection from HVCs to HVC2. Initially, there is only a weak synaptic projection (dashed line) from Nif2 to HVC2, which will not elicit a spike on HVC2 solely.

The simplest synfire chain network is shown in Figure 3-15. We only consider a single chain with two neurons laterally connected. Since the number of HVC neuron is in accordance with the elements of the sequence, in this model we only consider the sequence composed of two symbols. For simplicity, we say there are two different symbols: symbol 1 and symbol 2. These symbols are generated in Nif from the spectrogram in a short time window. The first neuron is the starter neuron, whose input from Nif is assumed to be well developed already. If the first symbol in the training sequence is 1, Nif 1 should project to HVC 1 reliably so that HVC 1 will fire whenever Nif 1 fires, because the synaptic strength is set to achieve this goal. Initially there are two HVC neurons: HVC 1 and HVC 2, two Nif neurons: Nif 1 and Nif 2. The elements in the training sequence are 1 and 2. Since HVC 1 is the starter neuron, the synapse from Nif 1 to HVC
1 is assumed to be a super-synapse and it does not involve in evolution, nor does it project to other neurons. Nif 2 projects to HVC 2 only weakly so that there is only small depolarization on HVC 2 caused by the firing of Nif 2. There is synaptic noise on soma and dendrite of HVC neurons. The standard deviation around resting potential is about 5mV without external inputs. The probability of spontaneous firing is rather low. There are two symbols 1 and 2. Symbol 1 is projected to HVC 1 at 25ms and symbol 2 is at 40ms. We will show how to get these values. A trial is within 60ms and whenever there is a spike, the associated synapse will undergo change by STDP mechanism, and all synapses experience homogeneous decaying. First we study the spontaneous activity of HVC 2 without the symbol 2.

Figure 3-16: Spontaneous activity and the learning. Panel A is the raster plot of the network above when only trained by sequence 1 at 25ms, Panel B is the raster plot when the training sequence is 1(25ms), 2(40ms). The histogram (blue) in panel B is the distribution of the formation of super-synapse projecting to HVC2 over 100 trials.
Figure 3-16 shows the raster plot of the spikes of HVC 2. For Figure 3-16A, HVC2 only receive the lateral input, according to the single neuron firing probability, at this time, HVC2 has elevated firing probability, and it provide the base for synapse growth from Nif2 to HVC2. With the presence of noise, we can see the spikes of HVC 2 are bounded in the next time window which spans from 40ms to 60ms. This feature facilitates the growth of the synapse from Nif 2 to HVC 2 if symbol 2 arrives at HVC 2 at a proper time. For Figure 3-16B where there is symbol 2, because the firing of Nif2 is right before all the spontaneous activities of HVC2, thus LTP will make the synaptic strength from Nif2 to HVC2 grow until it reaches the threshold to become a super-synapse, which is shown in inset C. Then the firing of HVC2 will be stable and intense. The noise amplitude is 0.04; slightly higher than the noise we used in recognition. The above learning process is 100 trials with the same training sequence.

3.7.2 Encoding from training sequence to sensory wiring

We have shown how the synapse grows governed by STDP mechanism. Since there is one Nif-HVC pair to connect except the starter neuron, it is straightforward for the only synapse to grow. When there are more neurons in the synfire chain to connect, we need axon withdraw to prune those undesired synaptic connections.
Figure 3-17: The minimal model involving axon withdrawing. Except the starter HVC neuron (HVCs), there are all-to-all weak connections from Nif neurons to HVC neurons, which is denoted by dashed line. Which connection becomes the super-synapse depends on the symbol sequence training the network, and the other one withdraws.

Figure 3-17 is the example of the synfire chain with three HVC neurons. There are three HVC neurons HVC 1, HVC 2, HVC 3 and three Nif neurons Nif 1, Nif 2, Nif 3. The training sequence is supposed to be 1 2 3.
Figure 3-18: The top row is the raster plot of HVC neuron 3, the middle row is the raster plot of HVC neuron 2, and the bottom row is the raster plot of HVC neuron 5. Each raster plot is based on 100 trials. Column A is the raster plot of HVC neurons trained by sequence 1(25ms). Column B is the raster plot of HVC neurons trained by sequence 1(25ms) 2(40ms). Column C is the raster plot of HVC neurons trained by sequence 1(25ms) 2(40ms) 3(60ms). The histograms (blue) in panel B and C are the distributions of the formation of super-synapses projecting to HVC2 and HVC3 over 100 trials.

Figure 3-18A is the spontaneous activities of HVC 2 and HVC 3. We can see the spikes of HVC 2 is bounded in the time window as we see in the two-neuron model, the number of spikes of HVC 3 is much less than the number of spikes of HVC 2. This means HVC 2 is ready to connect with the Nif neuron fires at the beginning of that time window, but HVC 3 is not. Given the input from Nif2, the synapse from Nif 2 to HVC 2 will grow. After some trials, when it becomes the super-synapse, HVC 2 could propagate spikes reliably to HVC3, the spikes distribution of HVC 3 will be restricted in the next time window. Figure 3-18B shows the spontaneous activity of HVC 3 when the training sequence is 1 2. Figure 3-18C is the evolution of the four pairs of synapse, if the training sequence is 1 2 3. The synapses selectively grow according to different training sequence. Nif 2 connects to HVC 2 and Nif 3 connects to HVC 3.
The synapse from Nif 2 to HVC 3 does not grow prominently because the LTP change per pair decays exponentially as time difference of the pre-neuron spike and the post-neuron spike increases. Although Nif 2 fires before HVC 3, their distributions are in two adjacent time windows, the growing rate is much less than in the same time window. Therefore the synapse from Nif 2 to HVC 3 grows much slower than the synapse from Nif 2 to HVC 2, and will withdraw from HVC 3 when the latter reach the withdraw threshold. The synapse from Nif 3 to HVC 2 will not grow because most of spikes of HVC 2 happen before Nif 3, so LTD suppresses its growth. The above process takes about 100 trials. The noise amplitude is 0.04.
However is the training sequence is 1 3 2, Nif 2 will connect to HVC 3 instead of HVC 2 and Nif 3 will connect to HVC 2 instead of HVC 2, which is shown in Figure 3-19A. In our scheme, only one super-synapse is allowed for each Nif neuron, so which one is final target depends on the training sequence provided. Column A is trained by sequence 1(25ms), 3(40ms), 2(60ms), as shown in the top panel, there are four cells separated by dashed lines, and the x-axis is HVC neuron index and y-axis is Nif neuron index. We can see Nif3 chooses HVC2 as its
preferred target and the synapse continues growing until it becomes a super-synapse, while the synapse from Nif3 to HVC3 does not grow, because the firing of Nif3 is much earlier than the spontaneous activities of HVC3, thus the synapse does not get enough enhancement from LTP to resist the homogenous decaying or the growth is more less than the growth from Nif3 to HVC2. When the latter reach the super-synapse threshold, it finally withdraws. After the super-synapse from Nif3 to HVC2 is formed, the lateral input gives HVC3 sparse but stable spontaneous activities, when there is an input from Nif2(60ms), which is earlier than most of spontaneous activities of HVC3, the synapse will keep strengthening until it reach the super-synapse threshold. Nif2 cannot grow to HVC2, because most of firing takes before the input from Nif2 (60ms). Column B shows the training process by sequence 1(25ms), 2(40ms), 3(60ms). Based on the same mechanism, the super-synapses are from Nif2 to HVC2 and from Nif3 to HVC3 as shown in the bottom panel.

3.7.3 Learning Rate

There are several factors that can affect the learning rate. The most prominent factor is the maximum percentage change per spike pair. The standard maximum LTP change is 5% and the standard maximum LTD change is 5.25%. Increasing these two factors can greatly speed up the growth of synapses. As a comparison, it takes about 200 trials for Nif2-HVC2 to reach the super-synapse threshold for standard value, while it takes 30 trials to reach the super-synapse threshold for 10 times of standard value. The other three factors can affect learning in similar way: the noise level, the lateral connection and the homogeneous decaying factor. All of them will determine the spontaneous activity of HVC neuron. High level spontaneous activity makes synaptic change easier to happen, and it takes longer time to reach super-synapse when the spontaneous activity is low.
3.8 Conclusion

1) Two-compartment model is used to model synfire chain neuron. This will give intrinsic platform to integrate sensory information with local network activities, which can be used as sequence recognition.

2) A-current can prominently enlarge the coincident detection region of single neuron because dendrite gives more room for A-current to play a role without firing.

3) STDP can encode the training sequence into specific wiring from Nif (sensory) to HVC (synfire chain). Noise plays an essential role to elicit postsynaptic spike from UP state in order to STDP works.

4) Symbolic sequence is generated from song spectrogram, which borrowed from speech recognition. The method is not robust to the natural variation of song
BIBLIOGRAPHY


44. P. D. Roberts and C. C. Bell, Biological Cybernetics 87, 392 (2002).


50. E. M. Izhikevich, Cerebral Cortex 17, 2443 (2007).


77. Lapicque L., J. Physiol. Paris 9, 622 (1907).


104. R. Legenstein and W. Maass, Neural Networks 20, 323 (2007).


VITA

Linli Wang

Education

- **The Pennsylvania State University** 2003-2009
  Ph.D, Physics  December 2009
  Major GPA: 3.7/4.
  Academic advisor: Professor Dezhe Jin
  Awards: University Fellowship
- **The University of Hong Kong** 1999-2002
  Master of Philosophy, Physics
  Awards: University studentship
- **University of Science & Technology of China** 1994-1999
  Bachelor of Science, Physics
  Major GPA: 3.7/4.
  Awards: University Scholarship

Teaching experience

- Physics instructor for engineering college (summer 2009)
- 6 months online course assignment design experience (2007)
- Judge of Penn State University graduate exhibition 2008
- Judge of Pennsylvania Junior Academy of Science 2009

Skills

- **Computational and theoretical physics**
  Probability theory and mathematical statistics, Numerical analysis, Computational physics, Mathematical methods of physics, Linear algebra, Statistical mechanics, Time series analysis, etc.
- **Computational physics with Fortran, C++ programming**
  OOP programming on modeling biological neural network composed of hundreds of neurons.
  Parallel optimization of serial code; Integrated computing with Matlab, Modular design.
- **Matlab programming**
  Spectrogram analysis; Sound signal analysis; Speech recognition; Pattern recognition; Vector quantization; Hidden Markov model; Matching pursuit decomposition; Integrated computing with C++,
- **Linux user and 1 year server administration experience**
  Set up network file system and auto mount file system;
  Manage two high performance computing servers and five workstations;
  Write a shell script to do 7 day auto-rotating backup system;
- **PC system experience with Microsoft Office Suite, VBA programming**
- **Neuroscience research**
  Model the sequence recognition behavior and sensory learning of songbird

Presentation

- “A computational model of song selectivity in songbird nucleus HVC”
  The 2006 Society for Neuroscience Annual Conference, Atlanta, GA, Linli Wang and Dezhe Jin
- “A learning mechanism of song selectivity in songbird nucleus HVC”
  The 2007 Society for Neuroscience Annual Conference, San Diego, CA, Linli Wang and Dezhe Jin