**1. Reduced Spiking Neural Network (RSNN)**

**1.1 RSNN Methods**

***Network Definition.***Xiaojing-Wang [1] proposed a biophysically realistic cortical network model for a binary visual discrimination experiment. It’s a common way to reduce the model by treating the net input to a neuron in a large homogeneous population as a Gaussian random process [2]. Therefore, for a binary visual decision-making task, the mean activity of a (homogeneous) population (left/right) can be represented by a single unit. Based on the well-developed simplified two-variable version of a biophysically realistic cortical network model of decision making[2], I extended it to a four-variable version (See Figure [2).](#_bookmark2) The 4 units stand for 4 directions, where 1/2/3/4 indicate 180/90/0/270 deg. There is a self- to-self excitation for each unit, and each unit will send inhibition to all the other three units.

Basically, let 𝑟 be the firing rate of a leaky integrate-and-fire (LIF) neuron receiving a noisy input current. Then, the firing rate could be described by the equation [1](#_bookmark1) [[3,](#_bookmark13) [7,](#_bookmark17) [8]](#_bookmark18) shown below.

Diagram

Description automatically generated

### Figure 2: simplified four-variable version of a biophysically realistic cortical network model of decision making

where Φ is a function of the total synaptic input current . is the membrane time constant. is the spiking threshold for the membrane voltage, is the reset voltage, is the refractory period, is the membrane potential SD, and = . Instead of using the equation 1, the two-variable version used Abbott and Chance function {Abbott, 2005 #4} for . Then, they could use several first-order dynamical equations to model the firing rate of the model. Here, I used the same simplification but for four variables. The model considers four excitatory neural assemblies, populations 1, 2, 3 and 4, standing for left, up, right, and down, that compute with each other through a shared pool of inhibitory neurons. The firing rate is shown in the equation 2.

In the equation [2,](#_bookmark3) I let 𝑟1, 𝑟2, 𝑟3, and 𝑟4 be firing rates of E and I populations, and the total synaptic input current 𝐼𝑖 and the resulting firing rate 𝑟𝑖 of the neural population 𝑖 obey the following input- output relationship (𝐹 𝐼 curve).It captures the current-frequency function of a leaky integrate-and-fire neuron.

−

Moreover, the synaptic drive is defined as below:

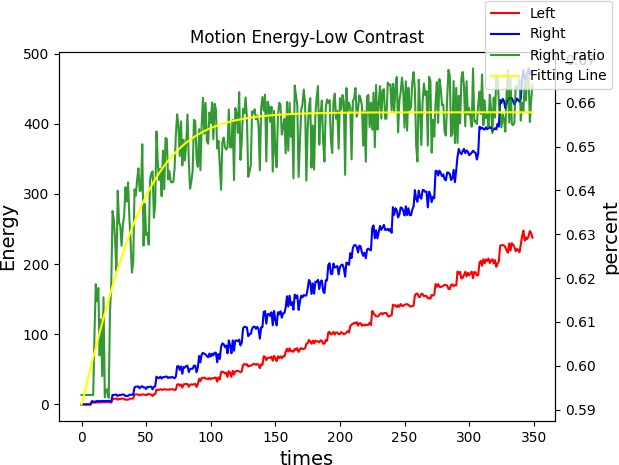
The net current into each population directions 1/2/3/4 indicate180/90/0/270 deg:

There are three critical parameters in this model, which are , and , where is the connection weight for the self- to-self excitatory connection weight, is the self-to-opposite inhibitory connection weight, and is the self-to-orthogonal inhibitory connection weight. Since unit 1 (180 deg) to unit 3 (0 deg) are opposite direction, and 1 (180 deg) to both unit 2 (90 deg) and unit 4 (270 deg) are orthogonal direction, the, and are set accordingly in the equation [7.](#_bookmark5)

An important contribution here I made is designing the input

of the RSNN. Basically, we computed the motion energy of the input stimulus. After using the Naka-Rushton function [[9](#_bookmark19)] to fit the motion energy, we design the input based on the motion energy. The details could be found in the section [4.](#_bookmark9)

***Input to RSNN*.** The input to the RSNN is the motion energy based on the spatiotemporal energy models [[2](#_bookmark12)]. It’s a classical method in capturing the perception of motion, based on the outputs of quadrature pairs of filters. The first step is to compute the Fourier Transform of the stimulus, then numerically fitting the distribution of the frequencies across the time. As shown in the Figure [7,](#_bookmark10) for low contrast stimuli, the ratio difference immediately emerges at very short du- rations. But for high-contrast stimuli, the ratio difference gradually increases. This may explain why we see the two directions get tangled only for high-contrast stimuli. We take the temporal profile of surround influence [4] as the convolution kernel to represent the size.

 Chart

Description automatically generated

1. **(b)**

### Figure 7: Motion Energy (a) High Contrast (b) Low Contrast

***Non-Decision Time.*** There is a well-established theory on the non-decision time in the perceptual decision-making. Therefore, although the time 0ms is when the stimulus was given, the network will receive the input util 100ms. Notice that here the input as 100ms is corresponding to the motion energy computed at time 0ms. Then, from 0ms to 100ms, we smooth the firing rate, leading to a smooth decision curve.

**1.2 RSNN Results**

After training the model accordingly, wcould obtain the firing rate for low (contrast=0.05) and high (contrast=0.99) contrast conditions, as shown in Figure 3a and Figure 3b. The parameters are listed as follows: = 0.3103, = -0.007, = -0.048. There is a stronger self-to-opposite inhibitory connection weight. As we can see, both in the low and high contrast condition, there is a competition among the four groups of neurons at the beginning, and then, following with the wining of the right direction. However, only in the high contrast condition could we observe the stronger opposite direction (left) comparing to the orthogonal directions (up/down). In terms of the decision making, we could put the firing rate to a softmax function to derive the psychophysical curve shown in Figure 3c and Figure 3d. That’s a good match to the behavioral data regarding the axis extraction.

1. Wang, X.J., *Probabilistic decision making by slow reverberation in cortical circuits.* Neuron, 2002. **36**(5): p. 955-68.

2. Wong, K.-F. and X.-J. Wang, *A Recurrent Network Mechanism of Time Integration in Perceptual Decisions.* The Journal of Neuroscience, 2006. **26**(4): p. 1314-1328.