Summary

Xiangyu Ma
Department of Physics, Hong Kong University of Science and Technology
xmaal@connect.ust.hk

1. Introduction

Multisensory integration and segregation are distinctly important for the survivorship of higher animals. The experimental data shows that the brain process information in a Bayesian way. The model proposed by Wenhao has demonstrate two different neuron groups are relate to the multisensory integration and segregation simultaneously. However, the details of the dynamics of this neuron system remain unknown. In this paper we will discuss the conditions of this neuron network using Bayesian inference to process complementary information.

2. Model

The module in Wenhao's paper is shown in Figure 1, the neural circuit consists of two modules, each of them contains two groups of excitatory neurons, congruent and opposite neurons respectively. The synaptic input of congruent groups and opposite groups are denoted as $\psi_m(y,t)$ and $\bar{\psi}_m(y,t)$ respectively, where $m = 1,2$ is the module index and y represents the preferred heading direction. The recurrent connections is a von Mises function: $J_{rc} = \frac{e^{a_0 cos(y-y')}}{2\pi I_0(a_0)}$, where J_{rc} and a_0 represent the strength and width of the connection. There are two types of reciprocal couplings between different modules. For congruent groups, the coupling function is given by $J_{rp} = \frac{e^{a_0 cos(y-y')}}{2\pi I_0(a_0)}$, however, the coupling function of opposite groups is $J_{rp} = \frac{e^{a_0 cos(y-y'+\pi)}}{2\pi I_0(a_0)}$. The firing rate is determined by synaptic inputs and divisive normalization, $\frac{[\psi_m(y,t)]_+^2}{1+\omega D_m}$, where $[x]_+ = max(x,0)$. D_m denotes normalization pool, $D_m = 1+\omega \sum_{-\pi}^{\pi}([\psi_m]_+^2 + J_{int}[\bar{\psi}]_+^2)$. The neuronal dynamics of congruent groups is given by

$$\begin{aligned}\tau\frac{\partial \psi_m(y,t)}{\partial t} =& -\psi_m(y,t) + \frac{J_{rc}}{D_m}\sum_{-\pi}^{\pi} V(y-y',a_0)\psi_m^2(y') + \frac{J_{rp}}{D_{\bar{m}}}\sum_{-\pi}^{\pi} V(y-y',a_0)\psi_{\bar{m}}^2(y') \\ &+ I_1 V(y-x,a_0/2) + I_b \end{aligned} \tag{1}$$

Where $\bar{m}$ denotes the complementary index. For opposite groups

$$\begin{aligned}\tau\frac{\partial \bar{\psi}_m(y,t)}{\partial t} =& -\bar{\psi}_m(y,t) + \frac{J_{rc}}{\bar{D}_m}\sum_{-\pi}^{\pi} V(y-y',a_0)\bar{\psi}_m^2(y') + \frac{J_{rp}}{\bar{D}_{\bar{m}}}\sum_{-\pi}^{\pi} V(y-y'+\pi,a_0)\bar{\psi}_{\bar{m}}^2(y') \\ &+ I_1 V(y-x,a_0/2) + I_b \end{aligned} \tag{2}$$

3. Basic Functions

Basically we have an intuition to construct a set of von Mises basic functions. First we define $N_n(k)$

$$N_n(k) = \int_{-\pi}^{\pi} e^{kcos\theta} sin^n\theta d\theta$$

Then we integrate by parts and make use of the property of periodical functions, we obtain iterative expression for $N_n(k)$

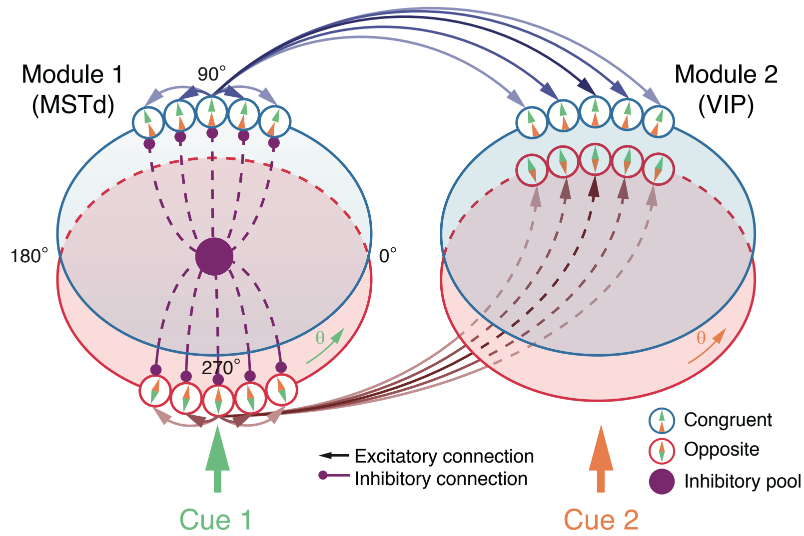

Figure 1: Decentralized module. MSTd and VIP refer to different modules, which consist of two groups of neurons, congruent (blue circles) and opposite (red circle) neurons. Neurons are all connected to the inhibitory pool in each module, which receives the feedforward input respectively (up arrow).

$$
\begin{aligned}
N_n(k) &= \int_{-\pi}^{\pi} e^{kcos\theta} sin^n\theta d\theta \\
&= \frac{n-1}{k}\int_{-\pi}^{\pi} e^{kcos\theta} sin^{n-2}\theta cos\theta d\theta \\
&= \frac{n-1}{k^2}\int_{-\pi}^{\pi} e^{kcos\theta}\left[(n-3)sin^{n-4}\theta cos^2\theta - sin^{n-2}\theta\right] d\theta \\
&= \frac{(n-1)(n-3)}{k^2}N_{n-4}(k) - \frac{(n-1)(n-2)}{k^2}N_{n-2}(k)
\end{aligned}
$$

The first few terms are derived below

$$
\begin{aligned}
N_0(k) &= 2\pi I_0(k) \\
N_2(k) &= \frac{2\pi I_1(k)}{k} \\
N_4(k) &= \frac{6\pi}{k^2}\left[I_0(k) - \frac{2I_1(k)}{k}\right] \\
N_6(k) &= \frac{30\pi}{k^3}\left[-\frac{4}{k}I_0(k) + (1+\frac{8}{k^2})I_1(k)\right]
\end{aligned}
$$

Where $I_n(k)$ is the modified Bessel function of the first kind. $N_n'(k)$ denotes the differential coefficient of $N_n(k)$.

3.1. Recursive Relation

We define $u_0 = C_0 e^{kcos\theta}$ as the zeroth order normalized basic function, and we notice the inner production

$$\langle u_0, u_0 sin\theta \rangle = 0$$

Hence we obtain the first order basic function $u_1 = C_1 u_0 sin\theta$, where C_1 is a normalization factor. Furthermore, we define the second order basic function

$$u_2 = C_2(u_1 sin\theta - u_0\langle u_0, u_1 sin\theta \rangle)$$

Since u_2 is an even function, u_2 is orthogonal to u_1. We can prove $\langle u_2, u_0 \rangle$.

Suppose we have a set of basic functions satisfying the following conditions

$$\begin{aligned}\langle u_i, u_j \rangle &= \delta_{i,j} \\ \langle u_i, u_j sin\theta \rangle &= r_{j-1}\delta_{i,j-1} + r_j \delta_{i,j+1}\end{aligned}$$

Where $i, j = 0, 1, 2...N-1$, we define $r_i = \langle u_i, u_{i+1} sin\theta \rangle$. u_n is given by

$$u_n = C_N(u_{N-1} sin\theta - u_{N-2}\langle u_{N-2}, u_{N-1} sin\theta \rangle)$$

When $i < N$, we can prove

$$\begin{aligned}\langle u_i, u_N \rangle &= C_N \left[\langle u_i, u_{N-1} sin\theta \rangle - \langle u_i, u_{N-2} \rangle < u_{N-2}, u_{N-1} sin\theta \rangle\right] \\ &= C_N \left[r_{N-2}\delta_{i,N-2} + r_{N-2}\delta_{i,N} - r_{N-2}\delta_{i,N-2}\right] \\ &= 0\end{aligned}$$

We have recursive relation

$$u_i sin\theta = r_{i-1}u_{i-1} + r_i u_{i+1}$$

Refer to Completeness of Fourier Expansion by Jeremy Orloff, the completeness of this set of basic function is easy to be proved. We can construct a delta function in this way

$$h(\theta) = c_k e^{k_0(cos\theta-1)}(1 - sin^2\theta)^k \ when \ k \to \infty$$

Where c_k is chosen to make sure $\int_{-\pi}^{\pi} h(\theta)d\theta = 1$. Note that the delta function $h(\theta)$ is a linear combination of $e^{k_0 cos\theta} sin^n\theta$, and $e^{k_0 cos\theta} sin^n\theta$ could be written by a linear combination of our basic functions.

The first few terms are derived below

$$\begin{aligned}u_0 &= c_0 e^{kcos\theta} \\ u_1 &= c_1 e^{kcos\theta} sin\theta \\ u_2 &= c_2 e^{kcos\theta}\left[sin^2\theta - \frac{N_2(2k)}{N_0(2k)}\right] \\ u_3 &= c_3 e^{kcos\theta}\left[sin^3\theta - \frac{N_4(2k)}{N_2(2k)} sin\theta\right]\end{aligned}$$

The coefficient c_i $i = 0, 1, 2...$ are listed

$$
\begin{aligned}
c_0 &= \frac{1}{\sqrt{N_0(2k)}} \\
c_1 &= \frac{1}{\sqrt{N_2(2k)}} \\
c_2 &= \frac{1}{\sqrt{N_4(2k) - \frac{N_2^2(2k)}{N_0(2k)}}} \\
c_3 &= \frac{1}{\sqrt{N_6(2k) - \frac{N_4^2(2k)}{N_2(2k)}}}
\end{aligned}
$$

Furthermore, we have

$$
r_i = \frac{c_i}{c_{i+1}} \; i = 0, 1, 2...
$$

3.2. Convolution

However, the convolution of two basic functions is hard to calculate. First we define

$$
\begin{aligned}
M_{m,n}(\theta, k_1, k_2) &= \int_{-\pi}^{\pi} e^{k_1 cos(\theta - \theta')} sin^m(\theta - \theta') e^{k_2 sin\theta'} sin^n \theta' d\theta' \\
T_{m,n}(\theta, k_1, k_2) &= \int_{-\pi}^{\pi} e^{k_1 cos(\theta - \theta')} e^{k_2 sin\theta'} sin^m(\theta - \theta') cos(\theta - \theta') sin^n \theta' d\theta'
\end{aligned}
$$

Our calculation base on the approximation

$$
M_{0,0}(\theta, k_1, k_2) = \int_{-\pi}^{\pi} e^{k_1 cos(\theta - \theta')} e^{k_2 sin\theta'} d\theta' \approx \frac{2\pi I_0(k_1) I_0(k_2)}{I_0(k_3)} e^{k_3 cos\theta}
$$

Where $A(k) = \frac{I_1(k)}{I_0(k)}$, $k_3 = A^{-1}(A(k_1)A(k_2))$. Take derivative of $M_{0,0}$ with respect to k_1, we have $T_{0,0} \approx e^{k_3 cos\theta} \left[F'(k_1) + F \frac{\partial k_3}{\partial k_1} cos\theta \right]$, where $F(k_1, k_2) = \frac{2\pi I_0(k_1) I_0(k_2)}{I_0(k_3)}$, $\frac{\partial k_3}{\partial k_1} = \frac{A'(k_1)A(k_2)}{A'(k_3)}$. We can derive convolution systematically. Take derivative of $M_{m,0}$

$$
\frac{\partial M_{m,n}}{\partial \theta} = m T_{m-1,n} - k_1 M_{m+1,n}
$$

That allow us to find $M_{m+1,n}$. We take derivative of $T_{m,n}$ with respect to θ, we obtain

$$
\frac{\partial T_{m,n}}{\partial \theta} = -k_1 T_{m+1,n} + m M_{m-1,n} - (m+1) M_{m+1,n}
$$

We will find $T_{m+1,n}$. We start from $T_{0,0}$ and $M_{0,0}$, consider the symmetry we have

$$
M_{m,n}(\theta, k_1, k_2) = M_{n,m}(\theta, k_2, k_1)
$$

For $T_{m,n}$, we have

$$
\begin{aligned}
T_{m,n}(k_1, k_2) &= \int_{-\pi}^{\pi} e^{k_1 cos(\theta - \theta')} e^{k_2 sin\theta'} sin^m(\theta - \theta') cos(\theta - \theta') sin^n \theta' d\theta' \\
&= cos\theta T_{n,m}(k_2, k_1) + sin\theta M_{m,n+1}(k_1, k_2)
\end{aligned}
$$

That allow us to find $T_{0,n}$ $n = 1, 2...$ starting from $T_{0,0}$.

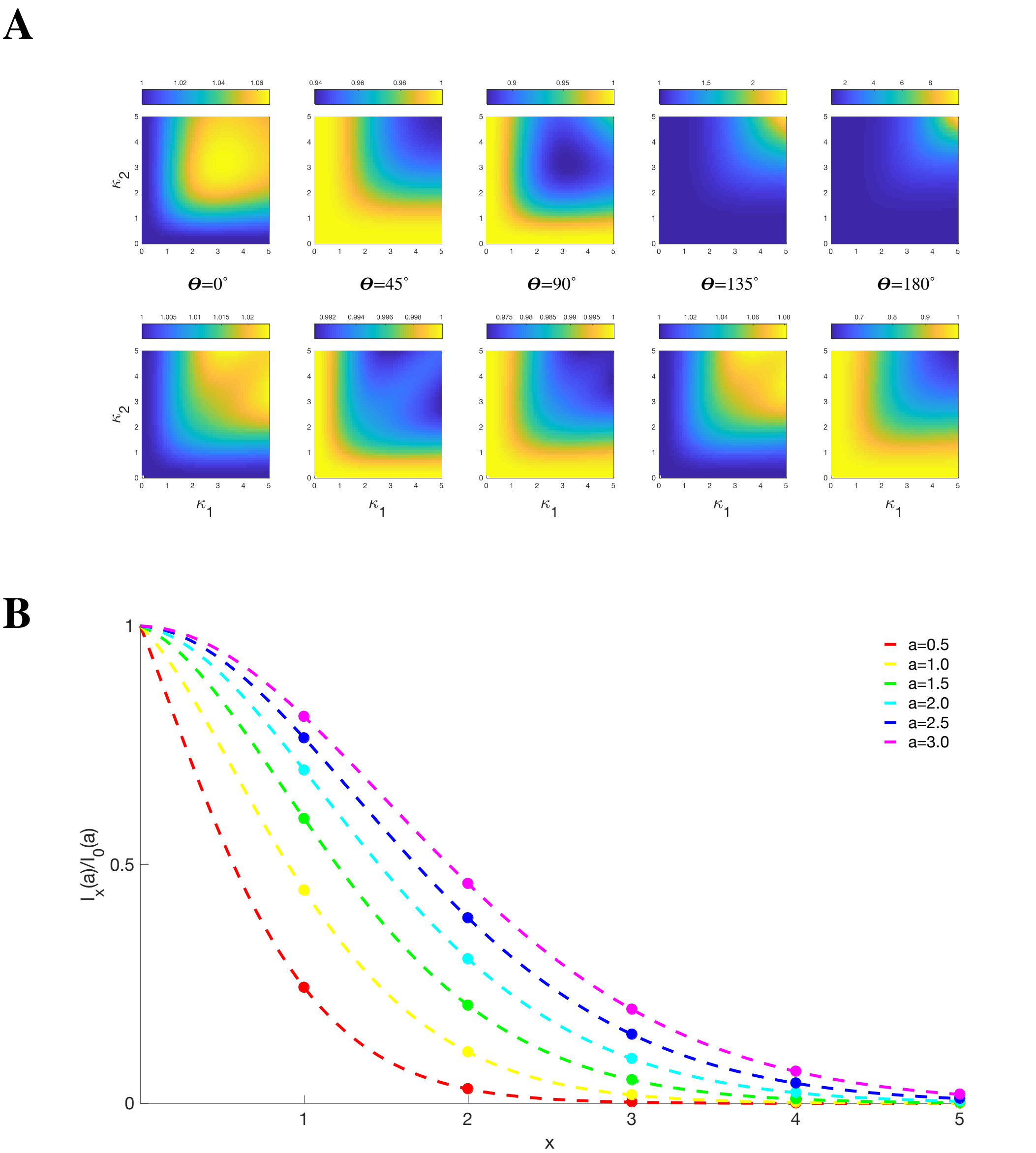

Figure 2: Illustration of approximations and Fourier basis. A: the ratio of estimated value to actual value for previous approximation (first row) and modified approximation (second row) given two concentration (κ_1, κ_2) and angular disparity (θ), color indicates the magnitude of the ratio. B: Function $I_x(a)/I_0(a)$ for different width a.

3.3. Approximation

Actually this approximation of von Mises convolution may not as accurate as we expect. For instance

$$Solution = \int_{-\pi}^{\pi} e^{2kcos\theta} sin^2\theta d\theta = N_2(2k) = \frac{\pi I_1(2k)}{k}$$

$$Approximation \approx -M_{1,1}(0,k,k) = \frac{2\pi k_3 I_0^2(k) e^{k_3}}{k^2 I_0(k_3)}$$

Where $k_3 = A^{-1}[A(k)^2]$. The percentage error is close to 40% when $2 < k < 3$. It works perfectly when k is approaching infinity.

There are two way to solve the problem. One is to modified this approximation. We know the integration

$$\int_{-\pi}^{\pi} e^{k_1 cos(\theta-\theta')} e^{k_2 cos\theta'} d\theta' = 2\pi I_0(\sqrt{k_1^2 + k_2^2 + 2k_1 k_2 cos\theta})$$

Previously, the Bessel function is approximated by

$$I_0(\sqrt{k_1^2 + k_2^2 + 2k_1 k_2 cos\theta}) \approx \frac{I_0(k_1) I_0(k_2)}{I_0(k_3)} e^{k_3 cos\theta} \ where \ k_3 = A^{-1}[A(k_1)A(k_2)]$$

For convenience, we take the log of both sides. Note $k = \sqrt{k_1^2 + k_2^2 + 2k_1 k_2 cos\theta}$. The following paragraph will compare our approximation with $ln(I_0(k))$.

In order to have a von Mises form approximation, suppose $ln(I_0(k))$ could be approximated by $d + hcos(f\theta)$.

We expand $ln(I_0(k))$, consider first three terms only

$$ln(I_0(k)) \approx \frac{b_0}{2} + b_1 cos\theta + b_2 cos2\theta$$

.

Where $b_i = \frac{1}{\pi}\langle ln(I_0(k)), cos(i\theta)\rangle \ i = 0,1,2$

$$\frac{b_0}{2} + b_1 cos\theta + b_2 cos2\theta \approx d + hcos(f\theta)$$

Multiply both sides by $cos(i\theta)$, integrate over θ, we get the solution

$$f = \sqrt{\frac{4b_2 + b_1}{b_2 + b_1}}$$
$$h = \frac{\pi b_1 (1 - f^2)}{2f sin(\pi f)}$$
$$d = \frac{b_0}{2} - \frac{b_1(1 - f^2)}{2f^2}$$

Figure 2 A shows this approximation is more accurate than previous one if we still want a von Mises form basic functions. Given k_1 and k_2, we calculate f h d respectively. $F = \frac{e^d}{2\pi I_0(k_1) I_0(k_2)}$ and $k_3 = h$ represent the coefficient and concentration. So the convolution of two von Mises functions is still a von Mises function.

$$\int_{-\pi}^{\pi} V(\theta - \theta', k_1) V(\theta', k_2) d\theta' \approx F(k_1, k_2) V(f\theta, k_3)$$

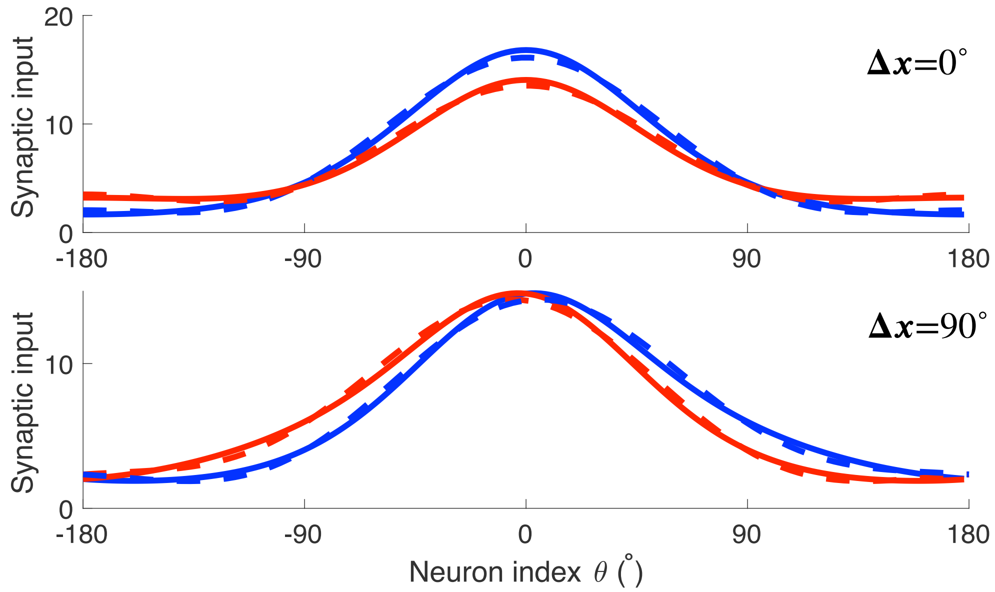

Figure 3: Bump profiles (synaptic inputs) of congruent and opposite groups in module 1 without noise. The blue and red colors represent congruent and opposite groups respectively. Solid lines: simulation results, dash lines: analytic results by using projection method.

The other way is to switch to Fourier basic functions. When we consider applying Fourier basis to a von Mises function, the coefficient of each order will be $\frac{I_n(a)}{I_0(a)}$, where a is the width of that von Mises function. Figure 2 B shows that when a is small, we don't need take high order perturbation into consideration. Since we are handling the case when concentration is comparatively small ($a \approx 1.5$), it's more suitable for us to choose Fourier series when applying projection method. For Fourier basis, the zeroth-order perturbation may refer to a constant (shift mode), the first-order perturbation will refer to height mode ($cos(\theta)$) and position mode ($sin(\theta)$), and the second-order perturbation refer to width mode ($cos(2\theta)$) and skewness mode ($sin(2\theta)$) around the point $a = 1.5$.

4. Projection Method

Suppose the solution to steady state equations is a linear combination of Fourier series

$$\psi_{1c} = u_0 + u_1 cos(y_1 - s_1) + u_2 cos2(y_1 - s_1) + u_3 sin2(y_1 - s_1)$$

For congruent groups, the steady state equation is given by

$$\begin{aligned}\psi(y_1) =& \frac{\rho J_{rc}}{D_1}\int dy_2 V(y_1 - y_2, a_0)\psi^2(y_2) + \frac{\rho J_{rp}}{D_2}\int dy_2 V(y_1 - y_2, a_0)\bar{\psi}^2(y_2) \\ &+ I_1 V(y_1 - x_1, \frac{a_0}{2}) + I_b\end{aligned} \tag{3}$$

Referring to inhibition pool, D_m could be replace by integration

$$D_m = 1 + \omega \int \rho[u_m^2(x,k) + J_{int}\bar{u}_m^2(x,k)]dx$$

Furthermore we define

$$B(n,k) = \frac{I_n(k)}{I_0(k)}$$

$D_m = 1 + \pi\omega\rho[2u_0^2 + u_1^2 + u_2^2 + u_3^2 + J_{int}(2\bar{u}_0^2 + \bar{u}_1^2 + \bar{u}_2^2 + \bar{u}_3^2)]$. Multiply the steady state equation's both sides by 1 $cos(y_1 - s_1)$ $sin(y_1 - s_1)$ $cos2(y_1 - s_1)$ $sin2(y_1 - s_1)$ and integrate over y_1, we obtain different dynamic equations of each mode respectively.

$$
\begin{aligned}
u_{10} =& \frac{\rho J_{rc}}{D_1}\left[u_{10}^2 + \frac{u_{11}^2}{2} + \frac{u_{12}^2}{2} + \frac{u_{13}^2}{2}\right] + \frac{\rho J_{rp}}{D_2}\left[u_{20}^2 + \frac{u_{21}^2}{2} + \frac{u_{22}^2}{2} + \frac{u_{23}^2}{2}\right] + \frac{I_1}{2\pi} + I_b \\
u_{11} =& \frac{\rho J_{rc}}{D_1}B(1,a_0)\left[2u_{10}u_{11} + u_{11}u_{12}\right] + \frac{I_1 B(1,a_0/2)}{\pi}cos(x_1 - s_1) \\
&+ \frac{\rho J_{rp}}{D_2}B(1,a_0)\left[(2u_{20}u_{21} + u_{21}u_{22})cos(s_2 - s_1) - u_{21}u_{23}sin(s_2 - s_1)\right] \\
0 =& \frac{\rho J_{rc}}{D_1}B(1,a_0)u_{11}u_{13} + \frac{I_1 B(1,a_0/2)}{\pi}sin(x_1 - s_1) \\
&+ \frac{\rho J_{rp}}{D_2}B(1,a_0)\left[(2u_{20}u_{21} + u_{21}u_{22})sin(s_2 - s_1) + u_{21}u_{23}cos(s_2 - s_1)\right] \\
u_{12} =& \frac{\rho J_{rc}}{D_1}B(2,a_0)\left[2u_{10}u_{12} + \frac{u_{11}^2}{2}\right] + \frac{I_1 B(2,a_0/2)}{\pi}cos2(x_1 - s_1) \\
&+ \frac{\rho J_{rp}}{D_2}B(2,a_0)\left[(2u_{20}u_{22} + \frac{u_{21}^2}{2})cos2(s_2 - s_1) - 2u_{20}u_{23}sin2(s_2 - s_1)\right] \\
u_{13} =& \frac{\rho J_{rc}}{D_1}B(2,a_0)2u_{10}u_{13} + \frac{I_1 B(2,a_0/2)}{\pi}sin2(x_1 - s_1) \\
&+ \frac{\rho J_{rp}}{D_2}B(2,a_0)\left[(2u_{20}u_{22} + \frac{u_{21}^2}{2})sin2(s_2 - s_1) + 2u_{20}u_{23}cos2(s_2 - s_1)\right]
\end{aligned}
$$

Firing rate $R_m(s_1, u_0, u_1, u_2, u_3) = \frac{\psi_m^2}{D_m}$, the definition of $\hat{s}$ in wenhao's paper is given by

$$\hat{s} = arg(\sum_{-\pi}^{\pi} R_m e^{j\theta})$$

Hence $\hat{s}$ can be expressed by the coefficients

$$\hat{s} = atan2[(2u_0u_1 + u_1u_2)sin(s) + u_1u_3cos(s), (2u_0u_1 + u_1u_2)cos(s) - u_1u_3sin(s)]$$

We compare analytic results with simulation in Figure 3. $\Delta x = x_2 - x_1$ is the angular disparity of two cues. When $\Delta x > 90°$ we only need to reverse the order of congruent and opposite groups because of the symmetry. All parameters are given in the following section except here we set $J_{int} = 0.5$. The time step size is 0.01τ and each time period is set equal to 100τ to make sure the solution has already converged. We are using Euler method to solve dynamic equations in our simulation.

Next we need take noise into consideration. The dynamics of congruent groups of neurons in module 1 in the presence of noise can by approximated by

$$
\begin{aligned}
\tau\frac{\partial\psi(y_1)}{\partial t} =& -\psi(y_1) + \frac{J_{rc}}{D_1}\int_{-\pi}^{\pi}V(y_1 - y_2, a_0)\psi^2(2) + \frac{J_{rp}}{D_2}\int_{-\pi}^{\pi}V(y_1 - y_2, a_0)\bar{\psi}^2(y_2) \\
&+ I_1 V(y - x, a_0/2) + I_b + \sqrt{F_0 I_1 V(y - x, a_0/2)}\xi_1 + \sqrt{F_0 I_b}\varepsilon_1
\end{aligned}
$$

Consider the dynamics of displacement mode and multiply both sides by $sin(y_1 - s_1)$, integrate over y_1

$$\tau\frac{\partial}{\partial t}\delta s_1 = -\delta s_1 + \frac{\rho J_{rc}}{D_1 u_{11}} B(1,a_0)\left[2u_{10}u_{11} + u_{11}u_{12}\right]\delta s_1$$
$$+ \frac{\rho J_{rp}}{D_2 u_{11}} B(1,a_0)\left[(2u_{20}u_{21} + u_{21}u_{22})cos(s_2 - s_1) - u_{21}u_{23}sin(s_2 - s_1)\right]\delta s_2$$
$$+ \frac{\sqrt{F_0 I_1}}{\pi u_{11}}\int\sqrt{V(y-x,a_0/2)}sin(y_1 - s_1)\xi_1 dy_1 + \frac{\sqrt{F_0 I_b}}{\pi u_{11}}\int sin(y_1 - s_1)\varepsilon_1 dy_1$$

The equivalent noise temperature is given by

$$T_1 = \frac{F_0}{2\pi^2\rho u_{11}^2}\left[I_1\int V(y_1 - x_1, a_0/2)sin^2(y_1 - s_1)dy_1 + I_b\int sin^2(y_1 - s_1)dy_1\right]$$
$$= \frac{F}{2\pi^2\rho u_{11}^2}\left[(\frac{I_1}{2} + \pi I_b) - \frac{I_1}{2}B(2,a_0/2)cos2(x_1 - s_1)\right]$$

And T_2

$$T_2 = \frac{F_0}{2\pi^2\rho u_{21}^2}\left[(\frac{I_2}{2} + \pi I_b) - \frac{I_2}{2}B(2,a_0/2)cos2(x_2 - s_2)\right]$$

Consider the dynamics $\tau\frac{\partial x_i}{\partial t} = -\Sigma_j g_{ij}x_j + \xi_j$ where $\langle\xi_i(t)\xi_j(t')\rangle = 2T_i\delta_{ij}\delta(t = t')$ and $\langle\xi_i(t)\rangle = 0$. Diagonalizing the matrix, $g_{ij} = \Sigma_k S_{ik}\lambda_k S^{-1}_{kj}$. Then

$$\tau\frac{\partial}{\partial t}\Sigma_j S^{-1}_{ij}x_j = -\lambda_i\Sigma_j S^{-1}_{ij}x_j + \Sigma_j S^{-1}_{ij}\xi_j$$

The solution is

$$x_i(t) = \Sigma_j S_{ij}\int_{-\infty}^{t}\frac{dt'}{\tau}exp[-\frac{\lambda_j}{\tau}(t - t')]S^{-1}_{jk}\xi_k(t')$$

Consider the moments of the distribution of $x_i(t)$

$$\langle x_i^2(t)\rangle = \Sigma_{j_1\ldots j_{2k}k_1\ldots k_{2k}}S_{ij_1}\ldots S_{ij_{2n}}\int_{-\infty}^{t}\frac{dt_1}{\tau}exp[-\frac{\lambda_{j_1}}{\tau}(t - t_1)]$$
$$\ldots\int_{-\infty}^{t}\frac{dt_{2n}}{\tau}exp[-\frac{\lambda_{j_{2n}}}{\tau}(t - t_{2n})]S^{-1}_{j_1k_1}\ldots S^{-1}_{j_{2n}k_{2n}}\langle\xi_{k(t_1)}\ldots\xi_{k_{2n}}(t_{2n})\rangle$$

The number of ways the pairing of the noise terms can be done is $(2n)!/(2!^n n!) = (2n-1)!!$. Hence

$$\langle x_i^2(t)\rangle = (2n-1)!!\{\Sigma_{j_1j_2k_1k_2}S_{ij_1}S_{ij_2}\int_{-\infty}^{t}\frac{dt_1}{\tau}exp[-\frac{\lambda_{j_1}}{\tau}(t - t_1)]$$
$$\int_{-\infty}^{t}\frac{dt_2}{\tau}exp[-\frac{\lambda_{j_2}}{\tau}(t - t_2)]S^{-1}_{j_1k_1}S^{-1}_{j_2k_2}2T_{k_1}\delta_{k_1k_2}\delta(t_1 - t_2)\}^n$$

After integrating

$$\langle x_i^2(t)\rangle = (2n-1)!!\left[\Sigma_{ijk}S_{ij}S_{ik}\frac{2T_l}{(\lambda_j + \lambda_k)\tau}S^{-1}_{jl}S^{-1}_{kl}\right]^n$$

Hence the distribution of $x_i(t)$ is a Gaussian with variance

$$\sigma_i^2 = \Sigma_{ijk} S_{ij} S_{ik} \frac{2T_l}{(\lambda_j+\lambda_k)\tau} S_{jl}^{-1} S_{kl}^{-1}$$

For the matrix $\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$

Where

$$G_{11} = 1 - \frac{\rho J_{rc}}{D_1 u_{11}} B(1,a_0)\left[2u_{10}u_{11} + u_{11}u_{12}\right]$$
$$G_{22} = 1 - \frac{\rho J_{rc}}{D_2 u_{21}} B(1,a_0)\left[2u_{20}u_{21} + u_{21}u_{22}\right]$$
$$G_{12} = -\frac{\rho J_{rp}}{D_2 u_{11}} B(1,a_0)\left[(2u_{20}u_{21} + u_{21}u_{22})cos(s_2-s_1) - u_{21}u_{23}sin(s_2-s_1)\right]$$
$$G_{21} = -\frac{\rho J_{rp}}{D_1 u_{21}} B(1,a_0)\left[(2u_{10}u_{11} + u_{11}u_{12})cos(s_1-s_2) - u_{11}u_{13}sin(s_1-s_2)\right]$$

The variance and concentration

$$\sigma_1^2 = \frac{T_1}{(G_{11}+G_{22})\tau} + \frac{T_1 G_{22}^2 + T_2 G_{12}^2}{(G_{11}G_{22} - G_{12}G_{21})(G_{11}+G_{22})\tau}$$
$$\kappa_1 = \frac{1}{\sigma_1^2}$$

5. Bayesian Inference

However, this network doesn't behave in a Bayesian way. We focus on the case that the stimulus strengths are equal, that is $I_1 = I_2 = I$. We consider the competition between congruent and opposite group in module 1. The solution is symmetric with respect to the perception displacement, that is $s_1 - x_1 = x_2 - s_2$. The matrix then becomes

$$G_{11} = 1 - \frac{\rho J_{rc}}{D} B(1,a_0)\left[2u_0 + u_2\right]$$
$$\bar{G}_{11} = 1 - \frac{\rho J_{rc}}{\bar{D}} B(1,a_0)\left[2\bar{u}_0 + \bar{u}_2\right]$$
$$G_{12} = -\frac{\rho J_{rp}}{D} B(1,a_0)\left[(2u_0 + u_2)cos(s_2-s_1) - u_3 sin(s_2-s_1)\right]$$
$$\bar{G}_{12} = -\frac{\rho J_{rp}}{\bar{D}} B(1,a_0)\left[(2\bar{u}_0 + \bar{u}_2)cos(\bar{s}_1-\bar{s}_2) - \bar{u}_3 sin(\bar{s}_1-\bar{s}_2)\right]$$

We have $G_{11} = G_{22},\ G_{12} = G_{21},\ \bar{G}_{11} = \bar{G}_{22},\ \bar{G}_{12} = \bar{G}_{21}$ because of symmetry. The noise temperature are given below

$$T = \frac{F_0}{2\pi^2 \rho u_1^2}\left[\frac{I_1}{2}(1 - B(2,a_0/2)cos2(x_1-s_1)) + \pi I_b)\right]$$
$$\bar{T} = \frac{F_0}{2\pi^2 \rho \bar{u}_1^2}\left[\frac{I_1}{2}(1 - B(2,a_0/2)cos2(x_1-\bar{s}_1)) + \pi I_b)\right]$$

Finally we obtain

$$\sigma_1^2 = \frac{G_{11}T}{(G_{11}^2 - G_{12}^2)\tau}$$
$$\kappa_1 = \frac{1}{\sigma_1^2} = \frac{\tau}{T}(G_{11} - \frac{G_{12}^2}{G_{11}}) \approx \frac{\tau}{T}$$

Hence $\frac{\kappa_1}{\bar{\kappa}_1} \approx \frac{u_1^2}{\bar{u}_1^2}$. The information of position and variance is decoded from the second and the third equations.

$$1 = HJ_{rc} + HJ_{rp}cos\Delta s + \frac{F}{u_1}cos(x_1 - s_1) \tag{4}$$

$$0 = HJ_{rp}sin\Delta s + \frac{F}{u_1}sin(x_1 - s_1) \tag{5}$$

Where $H = \frac{\rho}{D}(2u_0 + u_2)B(1, a_0),\ F = \frac{IB(1, a_0/2)}{\pi}$, under the weak input limit H can be treated as a constant.

When $\Delta s = \frac{\pi}{2}$

$$1 = HJ_{rc} + \frac{F}{u_1}$$

$$0 = HJ_{rp} + \frac{F}{u_1}(x_1 - s_1)$$

Where $(x_1 - s_1) \to 0$. We have $(s_1 - x_1)_{\Delta s = \frac{\pi}{2}} = \frac{HJ_{rp}}{1 - HJ_{rc}}$. When $\Delta x = 0$, we know $x_1 = s_1,\ x_2 = s_2$, the equations become

$$1 = HJ_{rc} + HJ_{rp} + \frac{F}{u_1}$$

Since HJ_{rp} is small, the ratio of u_1 to $\bar{u}_1$ can be approximated by

$$\left(\frac{u_1}{\bar{u}_1}\right)_{\Delta x=0} = \frac{1 - HJ_{rc} + HJ_{rp}}{1 - HJ_{rc} - HJ_{rp}} \approx 1 + \frac{2}{1 - HJ_{rc}}HJ_{rp} = 1 + 2(s_1 - x_1)_{\Delta s = \frac{\pi}{2}}$$

$$\left(\frac{\kappa_1 - \bar{\kappa}_1}{\kappa_1 + \bar{\kappa}_1}\right)_{\Delta x=0} = \frac{\left(\frac{u_1^2}{\bar{u}_1^2}\right)_{\Delta x=0} - 1}{\left(\frac{u_1^2}{\bar{u}_1^2}\right)_{\Delta x=0} + 1} \approx 2(s_1 - x_1)_{\Delta s = \frac{\pi}{2}}$$

That is, under weak input limit, if we are going to fit the variance in this network, the angle decoded from population vector will be twice as large as the real angle. This ratio is always greater than 2 when input strength is in normal range.

5.1. Assumption

Here we define the length of population vector $\hat{A} = mod(\frac{1}{N}\sum_{-\pi}^{\pi} R_i e^{j\theta})$ $= \frac{\rho}{ND}u_1\sqrt{(2u_0 + u_2)^2 + u_3^2}$, $\hat{A}$ is used to rescale the noise variance. We set $J_{int} = 1$ and rewrite the noise term $\sqrt{F_0 I_b \hat{A}}\varepsilon_1$.

The noise temperature

$$T = \frac{\sqrt{(2u_0 + u_2)^2 + u_3^2}}{2\pi N D u_1}F_0 I_b$$

$$\bar{T} = \frac{\sqrt{(2\bar{u}_0 + \bar{u}_2)^2 + \bar{u}_3^2}}{2\pi N \bar{D} \bar{u}_1}F_0 I_b$$

Consider weak input limit, $u_0 \approx \bar{u}_0$

$$
\begin{aligned}
\kappa \approx& \frac{\pi\tau ND}{F_0 I_b u_0} u_1 \\
\bar{\kappa} \approx& \frac{\pi\tau N\bar{D}}{F_0 I_b \bar{u}_0} \bar{u}_1
\end{aligned}
$$

c_0 denotes the coefficient of u_1 and $\bar{u}_1$, the concentration is proportional to the coefficient of the height mode.

$$\kappa = c_0 u_1$$

$$\bar{\kappa} = c_0 \bar{u}_1$$

From (4) and (5) we have

$$
\begin{aligned}
\left(\frac{F}{u_1}\right)^2 =& (1-HJ_{rc})^2 + (HJ_{rp})^2 - 2(1-HJ_{rc})HJ_{rp}cos\Delta s \\
u_1^2 =& \frac{F^2}{(1-HJ_{rc})^2 + (HJ_{rp})^2 - 2(1-HJ_{rc})HJ_{rp}cos\Delta s}
\end{aligned}
$$

Actually HJ_{rp} is small, and $cos\Delta s = cos[\Delta x - 2(s_1 - x_1)] \approx cos\Delta x$. We expand u_1^2 and $tan(s_1 - x_1)$ then get

$$
\begin{aligned}
tan(s_1 - x_1) =& \frac{HJ_{rp}sin\Delta s}{1-HJ_{rc}-HJ_{rp}cos\Delta s} \\
\approx& \frac{HJ_{rp}sin\Delta s}{1-HJ_{rc}} \\
\approx& \frac{HJ_{rp}sin\Delta x}{1-HJ_{rc}+HJ_{rp}cos\Delta x}
\end{aligned}
$$

$$
\begin{aligned}
u_1^2 \approx& \frac{F^2}{(1-HJ_{rc})^2} + \frac{2F^2HJ_{rp}}{1-HJ_{rc})^3} cos\Delta s \\
=& \frac{F^2}{(1-HJ_{rc})^4}\left[(1-HJ_{rc})^2 - 2HJ_{rp}(1-HJ_{rc})cos(\pi-\Delta s)\right] \\
\approx& \frac{F^2}{(1-HJ_{rc})^4}\left[(1-HJ_{rc})^2 - 2HJ_{rp}(1-HJ_{rc})cos(\pi-\Delta x) + (HJ_{rp})^2\right]
\end{aligned}
$$

Finally we obtain

$$tan(s_1 - x_1) \approx \frac{HJ_{rp}sin\Delta x}{1-HJ_{rc}+HJ_{rp}cos\Delta x}$$

$$\kappa^2 \approx \frac{c_0^2F^2}{(1-HJ_{rc})^4}\left[(1-HJ_{rc})^2 - 2HJ_{rp}(1-HJ_{rc})cos(\pi-\Delta x) + (HJ_{rp})^2\right]$$

5.1.1. Only Cue 1

When $I_1 = I$ and $I_2 = 0$, this network is asymmetric. The bump in module 2 is much weaker. We ignore the reciprocal coupling term then (4) and (5) in module 1 become

$$1 \approx HJ_{rc} + \frac{F}{u_1} cos(x_1 - s_1)$$

$$0 \approx \frac{F}{u_1} sin(x_1 - s_1)$$

The solution to this problem is

$$s_1 = x_1$$
$$u_1 = \frac{F}{1 - HJ_{rc}}$$

The concentration decoded from module 1 is

$$\kappa_1 = \frac{c_0 F}{(1 - HJ_{rc})}$$

5.1.2. Only Cue 2

When $I_2 = I$ and $I_1 = 0$, (4) and (5) in module 1 become

$$1 = HJ_{rc} + HJ_{rp} \frac{u_{12}}{u_{11}} cos\Delta s$$
$$0 = HJ_{rp} \frac{u_{12}}{u_{11}} sin\Delta s$$

We obtain

$$s_1 = s_2 \approx x_2$$
$$u_{11} = \frac{HJ_{rp}}{1 - HJ_{rc}} u_{12}$$

From previous discussion we know the solution is

$$u_{12} = \frac{F}{1 - HJ_{rc}}$$

The concentration decoded from module 1 is

$$\kappa_{12} = c_0 u_{11} = \frac{c_0 F HJ_{rp}}{(1 - HJ_{rc})^2}$$

5.1.3. Combined Cues

Compare the results, we have

$$\kappa e^{js_1} = \kappa_1 e^{jx_1} + \kappa_{12} e^{jx_2}$$

Where j is the imaginary unit.
The second and the third equations of opposite group are

$$1 = HJ_{rc} + HJ_{rp} cos\Delta\bar{s} + \frac{F}{\bar{u}_1} cos(x_1 - \bar{s}_1) \quad (6)$$
$$0 = -HJ_{rp} sin\Delta\bar{s} + \frac{F}{\bar{u}_1} sin(x_1 - \bar{s}_1) \quad (7)$$

Similarly, we have

$$\bar{\kappa} e^{j\bar{s}_1} = \bar{\kappa}_1 e^{jx_1} + \bar{\kappa}_{12} e^{j(x_2 + \pi)}$$

This network can apply Bayesian inference for a wide range of values of the inputs.

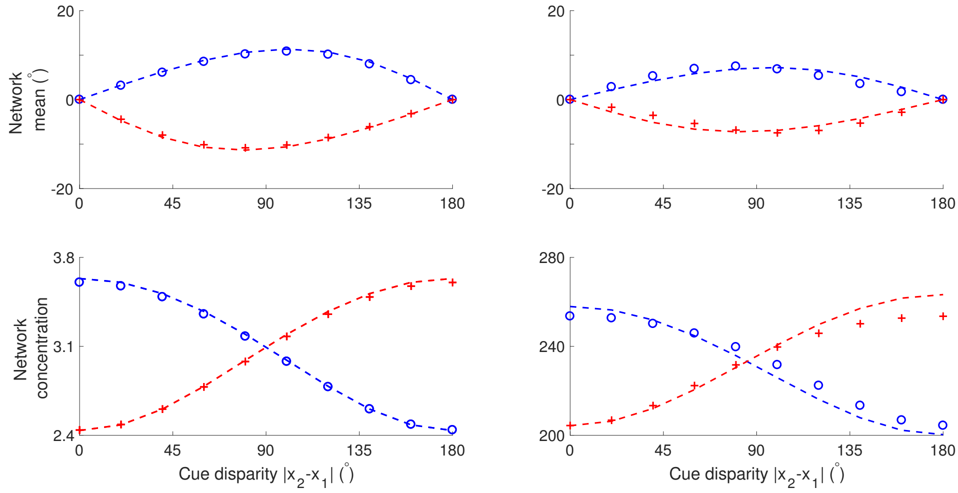

Figure 4: Weak input (left column) and strong input (right column). Symbols: network results; dash lines: Bayesian prediction. The blue and red colors represent congruent and opposite groups in module 1 respectively.

5.2. Results

The parameters are listed. Each network consist of $N = 180$ congruent and opposite neurons respectively. The connection width $a_0 = 3$ and the time step is 0.01τ using Euler method where τ is rescaled to 1. The strength of background input is $I_b = 1$. The strength of divisive normalization $\omega = 3\times10^{-4}$. F_0 in the noise term is the Fano factor which is set to 0.5.

In simulation we fix $x_1 = 0$, that is, $\Delta x = x_2 - x_1 = x_2$. J_c is the minimal recurrent strength and we choose $J_{rc} = 0.3J_c$ and $J_{rp} = 0.5J_{rc}$. J_c can be found by solving the dynamic equations, which is given by

$$J_c = \sqrt{\frac{8\pi I_0(a_0/2)^2\omega(1+J_{int})}{I_0(a_0)\rho}}$$

U_0 is the rescaled input strength $U_0 = \frac{J_c e^{a_0/2}}{2\pi\omega(1+J_{int})I_0(a_0/2)}$. We compare the case of weak input $I = 0.01U_0$ (left column) and strong input $I = 0.7U_0$ (right coulmn) in module 1. The results are derived by solving dynamic equations. Figure 4 shows this network could implement Bayesian inference under weak input limit. When external inputs are strong, the prediction from this network is close to the Bayesian way.

6. Causal Inference

Let's extend our work to generic type. The prior, which given in Wenhao's paper, is $p(s_1,s_2) \propto M(s_1-s_2,\kappa_s)$. We rewrite the prior of congruent and opposite groups

$$p(s_1,s_2) = \frac{p_0}{2\pi}M(s_1,s_2|\pm\kappa_s) + \frac{1-p_0}{(2\pi)^2}$$

The posterior can be calculated by marginalizing $p(s_1,s_2|x_1,x_2)$. We integrate over s_2 yielding

$$\begin{aligned} p(s_1|x_1,x_2) =& (2\pi)^2 p(x_1|s_1)\int_{-\pi}^{\pi} p(x_2|s_2)p(s_1,s_2)ds_2 \\ \approx& p_0 C M(s_1|x_1,\kappa_1)M(s_1|x_2,\pm\kappa_s) + (1-p_0)M(s_1|x_1,\kappa_1) \end{aligned}$$

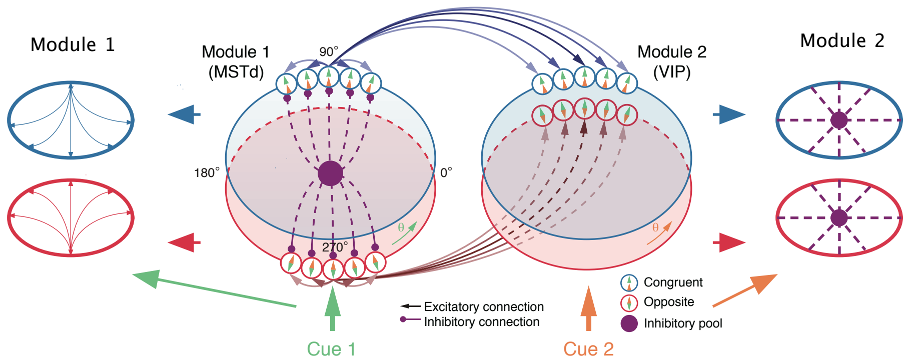

Figure 5: New modules in the second layer. No connection between two groups of neurons in each module. The inputs and feedforward inputs from first layer are rescaled by $1-p_0$ and p_0 respectively. Each groups of neurons have their own inhibition pool and there's no reciprocal coupling between two modules.

Where C is norm constant, normalization pool $D' = 1 + \omega \sum\limits_{-\pi}^{\pi} [\psi]_+^2$. By adding another module, we can implement the second component as an independent part. The dynamics of congruent and opposite groups is given by

$$\begin{aligned}\tau \frac{\partial \psi(y,t)}{\partial t} =& -\psi(y,t) + \frac{J_{rc}}{D'} \sum_{-\pi}^{\pi} V(y-y', a_0)\psi^2(y') + p_0 \frac{c_k}{D} \sum_{-\pi}^{\pi} cos(y-y')\psi^2(y') \\ &+ (1-p_0) I V(y-x, a_0/2) + I_b \end{aligned} \tag{8}$$

Where $\frac{\psi^2(y')}{D}$ is the firing rate with kernel $c_k cos(y-y')$ from lower layer, that is, the module introduced by Wenhao's paper. In Figure 5, we use synaptic inputs as the feedforward inputs in the upper layer. Without reciprocal connections from other modules, read out the concentration $\kappa'_m(\bar{\kappa}'_m)$ from population vectors in these modules. The final output will be the weighted sum of two. That is, for the case of combined cues, the final output of congruent group in module 1 will become

$$p_0 \kappa e^{js_1} + (1-p_0)\kappa'_1 e^{jx_1}$$

The first component is a complex number since it contains position information. Meanwhile, κ'_1 doesn't change when this network only receive input 1.

$$p_0 \kappa_1 e^{jx_1} + (1-p_0)\kappa'_1 e^{jx_1}$$

Only cue 2, the final output from congruent group in module 1 is

$$p_0 \kappa_{12} e^{jx_2}$$

From previous discussion, we know this network still behave in a Bayesian way.

6.1. Dynamics

Similarly, we project (8) onto height and position modes. Consider the dynamics of the congruent group in module 1.

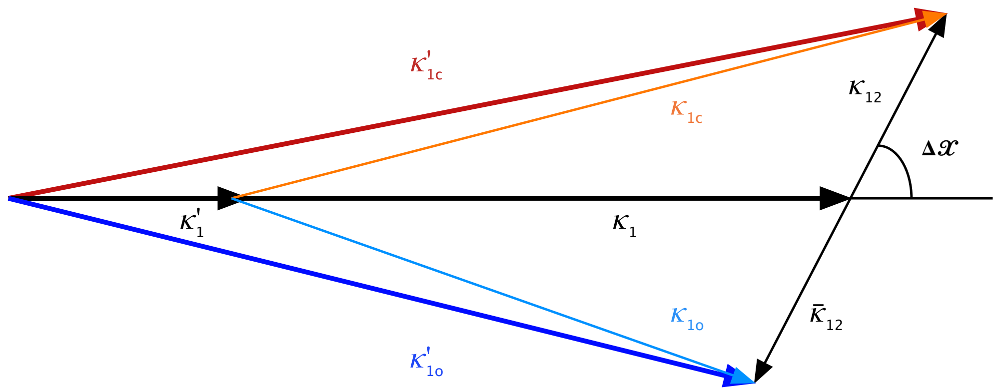

Figure 6: Geometric interpretation of vector space. Colored vectors are the sum of other vectors (black). Outputs from first layer: κ_{1c} (congruent) and κ_{1o} (opposite), second layer: κ'_{1c} (congruent) and κ'_{1o} (opposite). Note that the vectors (black) have been rescaled by p_0 and $1-p_0$.

$$1 = H'J_{rc} + (1-p_0)\frac{F}{u'_1}cos(x_1 - s'_1) + p_0\frac{c_k\rho u_1}{Du'_1}\left[(2u_0+u_2)cos(s_1-s'_1) - u_3 sin(s_1-s'_1)\right] \quad (9)$$

$$0 = (1-p_0)\frac{F}{u'_1}sin(x_1 - s'_1) + p_0\frac{c_k\rho u_1}{Du'_1}\left[(2u_0+u_2)sin(s_1-s'_1) + u_3 cos(s_1-s'_1)\right] \quad (10)$$

Where $H' = \frac{\rho J_{rc}}{D'}(2u'_0+u'_2)B(1,a_0)$, $F = \frac{IB(1,a_0/2)}{\pi}$, consider weak input limit, we know the solution is

$$u'_1 = (1-p_0)\frac{F}{1-H'J_{rc}}cos(x_1-s'_1) + p_0\frac{c_kHu_1}{J_{rc}B(1,a_0)(1-H'J_{rc})}cos(s_1-s'_1)$$
$$0 = (1-p_0)\frac{F}{1-H'J_{rc}}sin(x_1-s'_1) + p_0\frac{c_kHu_1}{J_{rc}B(1,a_0)(1-H'J_{rc})}sin(s_1-s'_1)$$

That is, for combined case, we have $u'_1e^{js'_1} = (1-p_0)\frac{F}{1-H'J_{rc}}e^{jx_1} + p_0\frac{c_kHu_1}{J_{rc}B(1,a_0)(1-H'J_{rc})}e^{js_1}$. When cue 1 is applied, this expression doesn't change except the second term u_1, which refers to the single cue case (only cue 1). On the other hand, when cue 2 is applied, the first term vanishes, the second term u_1 refers to the single cue case (only cue 2). Under this circumstance we won't have Bayesian inference unless $\kappa \propto u_1$ is still held.

Here we still consider using $\hat{A}$ to rescale the noise variance. The noise temperature doesn't change since they have the same expression. The noise variance can be calculate

$$\sigma^2 = \frac{T_1}{[1-\frac{\rho J_{rc}}{D'}(2u'_0+u'_2)B(1,a_0)]\tau}$$

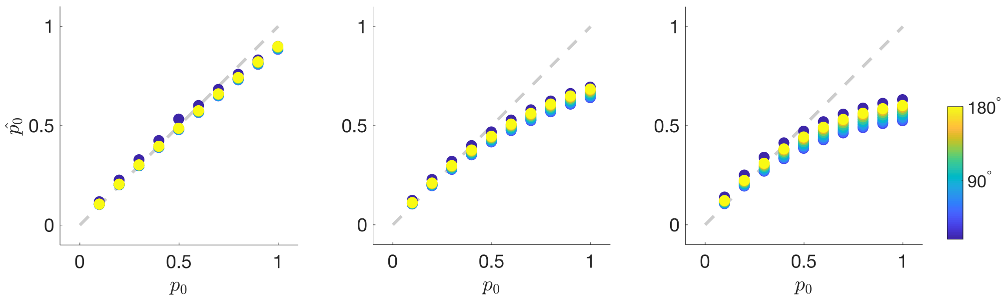

Figure 7: Outputs from second layer with varying input strength. From left to right, $I = 0.01U_0,\ 0.05U_0,\ 0.09U_0$. Colormap: color indicates the disparity Δx of two cues.

Hence we get concentration

$$\kappa_1' = \frac{1}{\sigma^2} \approx \frac{\tau}{T_1} \approx c_0 u_1'$$

This approximation is under weak input limit, c_0 is defined in previous section. We consider weak inputs and strong inputs, that is, $I = 0.01U_0$ and $I = 0.7U_0$, and compare the results with Bayesian prediction in Figure 8. So far the outputs from the second layer will behave in Bayesian way under weak input limit. Although the prior is the sum of two von Mises functions, it won't get two peaks since the position disparity between inputs and feedforward inputs is actually small. Figure 6 shows in vector space, causal inference can be achieved by taking advantage of the properties of this network.

6.2. Vector Space

The vectors, which are correspond to prior distributions, contain both position and variance information. Under weak input limit, we can set $c_k = \frac{J_{rc}+J_{rp}}{1-\sqrt{1-4\rho I_b(J_{rc}+J_{rp})}}$. Another interesting story is about causal inference, when J_{rc} is small, we have $\frac{|\kappa_1|}{|\kappa_1'|} \approx \frac{p_0}{1-p_0}$. If we extract information from two layers, the probability of common source p_0 could calculate from the vector diagram.

$$\hat{p}_0 = 1 - \frac{|\kappa_{1c}'|coss_1' - |\kappa_{1c}|coss_1}{\lambda|\kappa_{1o}'|cos\bar{s}_1' + |\kappa_{1c}'|coss_1'}(\lambda+1)$$

Where $\lambda = -\frac{|\kappa_{1c}'|sins_1'}{|\kappa_{1o}'|sin\bar{s}_1'}$, $1-\hat{p}_0$ corresponds to the probability of two cues coming from the different source. Hence opposite groups are important to distinguish whether two cues are the same or different.

The correlation between prediction $\hat{p}_0$ and given p_0 has been shown in Figure 7.

6.3. Information Segregation

In Wenhao's paper, the definition of the disparity information is given by

$$p_d(s_1|x_1,x_2) \propto p(s_1|x_1)/p(s_1|x_2)$$

However, the prior has changed. $p_d \propto \frac{M(s_1-x_1,\kappa_1}{p_0 CM(s_1-x_1,\kappa_{2s})+(1-p_0)}$ is different from that in Wenhao's paper. In order to implement information segregation, we need to modify our model.

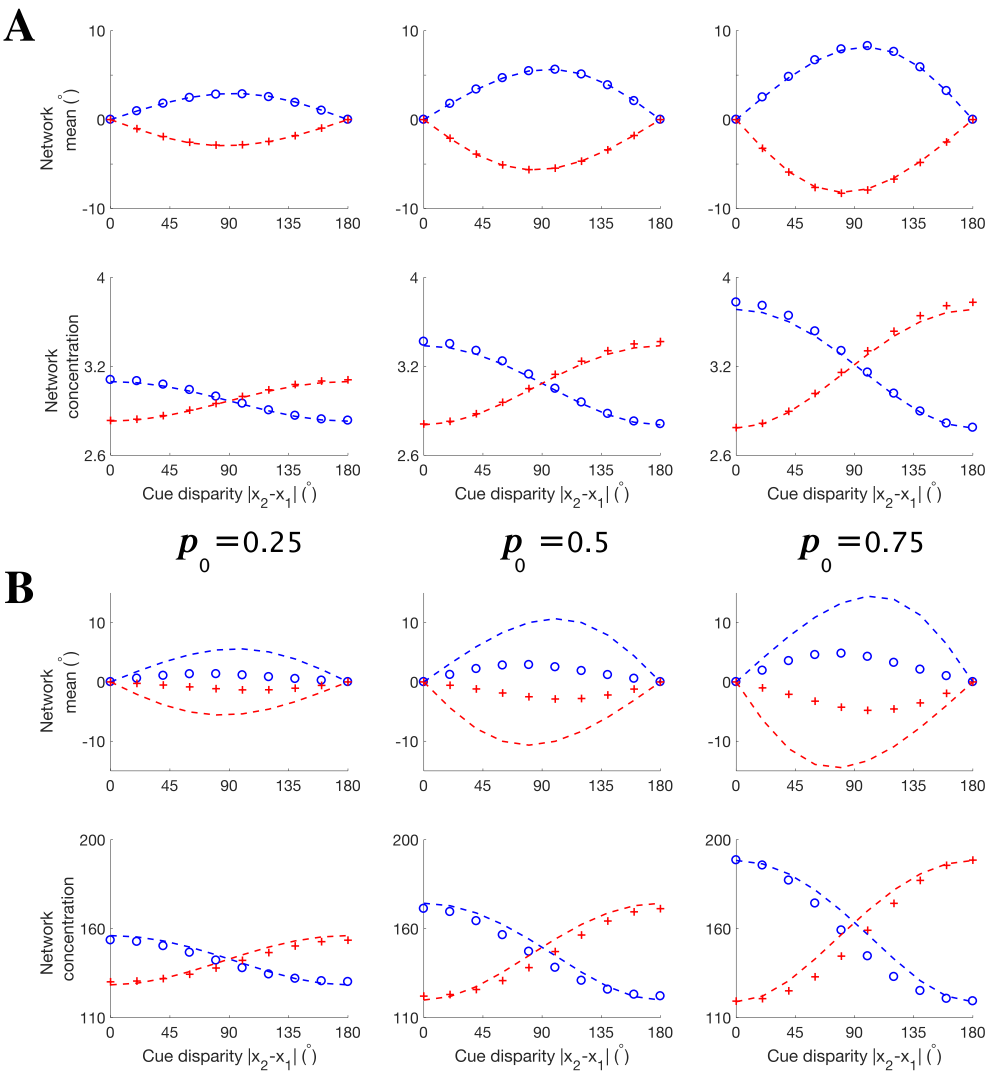

Figure 8: Outputs from second layer with varying probability p_0. Illustration of the population response of congruent and opposite groups in module 1 applying weak inputs (A) and strong inputs (B). Symbols: network results; dash lines: Bayesian prediction. The blue and red colors represent congruent and opposite groups in module 1 respectively.

Although p_d is hard to obtain from this neuron network, however, this ratio is equivalent to p_d^{-1}, which contains the same amount of information of cue segregation. p_d^{-1} can be written by

$$p_d^{-1} = p_0 C M(s_1 - x_2, \kappa_{2s}) M(s_1 - x_1 + \pi, \kappa_1) + (1 - p_0) M(s_1 - x_1 + \pi, \kappa_1)$$

Where C is norm constant. That is, the input's direction in module 1 will be shifted by π for opposite group. Actually, it is relatively easy to prove that the modified model have the same structure as that we define in previous section. The only difference is that position of the outputs from the second layer will be shifted by π.