Corrective feedback control of competing neural network with entire connections

Uramogi Wang [1,*]

[1]Harbin Institute of Technology

Abstract

Continuous persist activity of the competitive network is related to many functions, such as working memory, oculomotor integrator and decision making. Many competition models with mutual inhibition structures achieve activity maintenance via positive feedback, which requires meticulous fine tuning of the network parameters strictly. Negative derivative feedback, according to recent research, might represent a novel mechanism for sustaining neural activity that is more resistant to multiple neural perturbations than positive feedback. Many classic models with only mutual inhibition structure are not capable of providing negative derivative feedback because double-inhibition acts as a positive feedback loop, and lack of negative feedback loop that is indispensable for negative derivative feedback. Here in the proposal, we aim to derive a new competition network with negative derivative feedback. The network is made up of two symmetric pairs of EI populations that the four population are completely connected. We conclude that the negative derivative occurs in two circumstances, in which one the activity of the two sides is synchronous but push-pull-like in the other, as well as the switch of two conditions in mathematical analysis and numerical simulation.

Introduction

Many cortical functions, such as parametric working memory, decision making and neural integrator, are thought to be substantially tied to the persistent neural activity induced by the competitive neural network[1]. The traditional competitive structure is made up of two inhibitory populations connected by mutual inhibition, which is required for competitive activities. To explain working memory in two-interval discriminating choice tasks, Machens and Romo devised a basic model without self -connections to each side[2]. For oculomotor control and the neural integrator, several studies employ a similar paradigm[3]. However, according to Dale's law, two mutual inhibitory populations don't have self-excitation connections, which is ubiquitous in experimental recordings. Therefore, the structure of two mutual inhibition network is insufficiently physiologically rational.

X-J Wang has presented a more biologically sensible competition model[4]. In the model,7200 LIF(Leaky Integrate and Fire) spiking neurons compose a three-population neural architecture network, in which two excitatory neuron populations inhibit each other through a common inhibitory neuron pool. The emphasis of the study is on discrete attractor analysis of the decision-making process, rather than persistent activity. Furthermore, all these models show that the only way to sustain persistent neural activity is via positive feedback mechanisms[5]. Generally speaking, most mutual inhibition structure composes a double-negative connection loop that is equivalent to a positive feedback pathway. Positive feedback inputs might be used to balance the intrinsic leakage activity to maintain the information[6].However, positive feedback is supposed to have precise network parameter, requires a fine tuning of the level of feedback and are extremely susceptible to common perturbations, such as global changes in neuronal or synaptic excitabilities, that disrupt

this tuning.

In the brain, the representation and processing of information have great temporal and spatial dynamics. Furthermore, the presence of noise and physical changes (decline of neurons, change of synaptic strength, etc.) leads to the conclusion that network should be robust against common perturbations to some extent. Therefore, Subkin Lim and Mark.S Goldman[7] introduce corrective feedback control into cortical microcircuit. The structure is made up of one excitatory and one inhibitory population with recurrent connections, consisting of positive feedback pathway and negative feedback pathway in which the connection strengths are equal, but distinct synaptic constants. This accords with the principle of corrective input in control theory: negative derivative feedback arises when recurring excitatory and inhibitory inputs to the population are balanced in strength and offset in time.

We construct a four-population network consisting of two pairs of symmetric EI populations to apply negative derivative feedback to competition network. The structure is divided into three parts. First, the reason why the number of populations is four, with two excitatory and two inhibitory populations. Because Dale's law states that if two neurons have inhibitory connections, they can only be inhibitory neurons. In traditional two-population network with mutual inhibition, the neurons can only be inhibitory and it restricts excitatory connections and activities and is biologically unreasonable. Secondly, the reason why the populations are completely connected. Mutual inhibition may be established in a four-population structure by activating contralateral inhibitory neurons or inhibition of contralateral excitatory neurons(E to I cross and I to E cross). However, if there are no alternative negative feedback routes, simply mutual inhibition connections are not capable of negative derivative feedback. How to achieve those negative feedback pathways? In fact, it is really a complex problem. Therefore, we treat the model as if it were a fully connected network, the most complicated instance with both positive and negative feedback pathways originally. Furthermore, since cortical neurons receive vast quantities of both excitation and inhibition in a broad variety of situations and brain locations, a comprehensive connection is feasible for biological consideration[8].It is also an ideal situation as a mathematical model.

Finally, however, the difficulty is to derive negative derivative feedback from the structure briefly. Due to the complexity of the network structure, pathways of positive feedback and negative feedback intersect within each other, which leads to obscure differentiation and definition of positive or negative feedback pathways. The two EI populations are assumed symmetric for the sake of dimension reduction. Another reason is push-pull like activity needs a symmetric structure to keep the equivalent role of two sides.

The underlying mechanism by which negative derivative feedback might lead to persistent activity in the linked four population structure was investigated. Our analysis is based on the assumption that symmetrical network structure leads to symmetrical activities without any input. Therefore, when the system reaches stability, the activities on both sides are always the same or “opposite” compared with background baseline activity, resulting in two persistent activity modes in which one the activity of the two sides is synchronous but push-pull-like in the other. Finally, using mathematical analysis and numerical simulation, we were able to determine the assumption and infer the transition between two modes.

Method

1.Requirements for negative derivative feedback

The model is a four-population network consists of two same excitatory populations and two inhibitory populations. Connection strengths and time constants are symmetric distributed in two sides. The structure of the network is shown as:

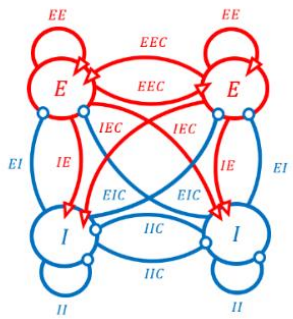

Figure 1. Structure of symmetric competition network with four neural population. All connection strengths and time constants are distributed symmetrically.

As shown in the figure, each population has connections with the other three populations. EE, EI, IE, II represent the connections within each side, and connections EEC, EIC, IEC, IIC represent the connections from one side to the other side. We denote connection strengths and time constants of as J_i and τ_i, with $i = EE, EI, IE, II, EEC, EIC, IEC, IIC$ .The system is described by firing rate model as :

$$\tau_E \dot{R}_{E1} = -R_{E1} + J_{EE} S_{EE1} + J_{EEC} S_{EEC1} - J_{EI} S_{EI1} - J_{EIC} S_{EIC1}$$
$$\tau_E \dot{R}_{E2} = -R_{E2} + J_{EE} S_{EE2} + J_{EEC} S_{EEC2} - J_{EI} S_{EI2} - J_{EIC} S_{EIC2}$$
$$\tau_I \dot{R}_{I1} = -R_{I1} + J_{IE} S_{EI1} + J_{IEC} S_{IEC1} - J_{II} S_{II1} - J_{IIC} S_{IIC1}$$
$$\tau_I \dot{R}_{I1} = -R_{I1} + J_{IE} S_{EI2} + J_{IEC} S_{IEC2} - J_{II} S_{II2} - J_{IIC} S_{IIC2}$$

$$\tau_{EE} \dot{S}_{EE1} = -S_{EE1} + R_{E1} \qquad \tau_{II} \dot{S}_{II1} = -S_{II1} + R_{I1}$$
$$\tau_{EEC} \dot{S}_{EEC1} = -S_{EEC1} + R_{E2} \qquad \tau_{IIC} \dot{S}_{IIC1} = -S_{IIC1} + R_{I2}$$
$$\tau_{EE} \dot{S}_{EE2} = -S_{EE2} + R_{E1} \qquad \tau_{II} \dot{S}_{II2} = -S_{II2} + R_{I1}$$
$$\tau_{EEC} \dot{S}_{EEC2} = -S_{EEC2} + R_{E1} \qquad \tau_{IIC} \dot{S}_{IIC2} = -S_{IIC2} + R_{I2}$$

$$\tau_{IE} \dot{S}_{IE1} = -S_{IE1} + R_{E1} \qquad \tau_{EI} \dot{S}_{EI1} = -S_{EI1} + R_{I1}$$
$$\tau_{IEC} \dot{S}_{IEC1} = -S_{IEC1} + R_{E2} \qquad \tau_{EIC} \dot{S}_{EIC1} = -S_{EIC1} + R_{I2}$$
$$\tau_{IE} \dot{S}_{IE2} = -S_{IE2} + R_{E1} \qquad \tau_{EI} \dot{S}_{EI2} = -S_{EI2} + R_{I1}$$
$$\tau_{IEC} \dot{S}_{IEC2} = -S_{IEC2} + R_{E2} \qquad \tau_{EIC} \dot{S}_{EIC2} = -S_{EIC2} + R_{I2}$$

The system is a linear firing rate model. To analyze the linear network, we use the eigenvector decomposition to decompose the coupled neuronal activities into non-interacting modes (eigenvectors) that can be considered independently.

Define a vector $\vec{R}$ represents the firing rate and synaptic input and a matrix $\overleftrightarrow{Y}$ represents the interaction kernel. Therefore, the linear system activity is given by the equation $\dot{\vec{R}} = \overleftrightarrow{Y}\vec{R}$:

$$\vec{R} = [R_{E1}\ R_{E2}\ R_{I1}\ R_{I2}\ S_{EE1}\ S_{EE2}\ S_{EEC1}\ S_{EEC2}\ S_{IE1}\ S_{IE2}\ S_{IEC1}\ S_{IEC2}\ S_{EI1}\ S_{EI2}\ S_{EIC1}\ S_{EIC2}\ S_{II1}\ S_{II2}\ S_{IIC1}\ S_{IIC2}]^T$$

$$\overleftrightarrow{Y} = \begin{bmatrix}
-\frac{1}{\tau_E} & 0 & 0 & 0 & \frac{J_{EE}}{\tau_E} & \frac{J_{EEC}}{\tau_E} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{J_{EI}}{\tau_E} & -\frac{J_{EIC}}{\tau_E} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\frac{1}{\tau_E} & 0 & 0 & 0 & 0 & \frac{J_{EEC}}{\tau_E} & \frac{J_{EE}}{\tau_E} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{J_{EIC}}{\tau_E} & -\frac{J_{EI}}{\tau_E} & 0 & 0 & 0 & 0 \\
0 & 0 & -\frac{1}{\tau_I} & 0 & 0 & 0 & 0 & 0 & \frac{J_{IE}}{\tau_I} & \frac{J_{IEC}}{\tau_I} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{J_{II}}{\tau_I} & -\frac{J_{IIC}}{\tau_I} & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{\tau_I} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{J_{IEC}}{\tau_I} & \frac{J_{IE}}{\tau_I} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{J_{IIC}}{\tau_I} & -\frac{J_{II}}{\tau_I} \\
\frac{1}{\tau_{EE}} & 0 & 0 & 0 & -\frac{1}{\tau_{EE}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\tau_{EEC}} & 0 & 0 & 0 & -\frac{1}{\tau_{EEC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\tau_{EE}} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{EEC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\tau_{EE}} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{EE}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\tau_{IE}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{IE}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\tau_{IEC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{IEC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{\tau_{IEC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{IEC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\tau_{IE}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{IE}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\tau_{EI}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{EI}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\tau_{EIC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{EIC}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\tau_{EIC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{EIC}} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\tau_{EI}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{EI}} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\tau_{II}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{II}} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\tau_{IIC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{IIC}} & 0 & 0 \\
0 & 0 & \frac{1}{\tau_{IIC}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{IIC}} & 0 \\
0 & 0 & 0 & \frac{1}{\tau_{II}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{II}}
\end{bmatrix}$$

Because the equation is a linear time-invariant(LTI) system , action vector $\vec{R}$ could be

$$\vec{R}(t) = \sum_{i=1}^{n} C_n e^{-\lambda_n t} \vec{v}_n \tag{1}$$

In equation (1), λ_n and $\vec{v}_n$ is the eigenvalue and eigenvector of the matrix $\overleftrightarrow{Y}$ with n dimensions. For the require of persist activity, at least one eigenvalue is equal or close to zero. Analytically to see, we consider the characteristic polynomial of the matrix, which is given by:

$$\begin{aligned} char(x) &= |\vec{Y} - x\overleftrightarrow{E}| \\ &= a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \\ &= (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n) \end{aligned} \tag{2}$$

Here, the matrix $\overleftrightarrow{E}$ denotes the n dimensions identity matrix, λ_n is the root of the characteristic polynomial and a_n is the coefficient of polynomial that represented by the network parameters. In this way, we could connect the eigenvalues with the parameters of the network through the coefficients. This can be shown by:

$$
\begin{aligned}
a_0 &= (-1)^n \prod_{i=1}^{n} \lambda_i \\
a_1 &= (-1)^{n-1} \prod_{i=1}^{n} \lambda_i \sum_{j=1}^{n} \frac{1}{\lambda_j} \\
\frac{a_1}{a_0} &= -\sum_{i=1}^{n} \frac{1}{\lambda_i}
\end{aligned}
\tag{3}
$$

The zero-order coefficient of the polynomial a_0 shows all products of the eigenvalues and determines whether there is a zero-valued eigenvalue. But one eigenvalue is equal or close to zero is not sufficient for $a_0 = 0$, in the case that one eigenvalue λ_0 is close to zero and another eigenvalue has the magnitude larger than $1/\lambda_0$.Therefore , we consider the ratio of a_1/a_0 which represents the sum of reciprocals of the eigenvalues. If one eigenvalue is equal or close to zero, the magnitude of the one reciprocal will be large and therefore the sum a_1/a_0 would be large. We utilize a_1/a_0 represented by the network parameters J_i, τ_i

$$
\begin{aligned}
a_0 = {} & ((J_{EI} + J_{EIC})(J_{IE} + J_{IEC}) - (J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1)) \cdot \\
& ((J_{EI} - J_{EIC})(J_{IE} - J_{IEC}) - (J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1))
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
a_1 = {} & -2((J_{EI} + J_{EIC})(J_{IE} + J_{IEC}) - (J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1)) \cdot \\
& ((J_{EI} - J_{EIC})(J_{IE} - J_{IEC}) - (J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1)) \cdot \\
& (\tau_{EE} + \tau_{EEC} + \tau_{IE} + \tau_{IEC} + \tau_{EI} + \tau_{EIC} + \tau_{II} + \tau_{IIC}) + \\
& \frac{1}{2}[\ ((J_{EE}\tau_{EE} + J_{EEC}\tau_{EEC} + \tau_E)(J_{II} + J_{IIC} + 1) - (J_{EI}\tau_{EI} + J_{EIC}\tau_{EIC})(J_{IE} + J_{IEC})) \cdot \\
& ((J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1) - (J_{EI} - J_{EIC})(J_{IE} - J_{IEC})) + \\
& ((J_{EE}\tau_{EE} + J_{EEC}\tau_{EEC} + \tau_E)(J_{II} + J_{IIC} + 1) - (J_{EI} + J_{EIC})(J_{IE}\tau_{IE} + J_{IEC}\tau_{IEC})) \cdot \\
& ((J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1) - (J_{EI} - J_{EIC})(J_{IE} - J_{IEC})) + \\
& ((J_{EE} + J_{EEC} - 1)(J_{II}\tau_{II} + J_{IIC}\tau_{IIC} - \tau_I) - (J_{EI}\tau_{EI} + J_{EIC}\tau_{EIC})(J_{IE} + J_{IEC})) \cdot \\
& ((J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1) - (J_{EI} - J_{EIC})(J_{IE} - J_{IEC})) + \\
& ((J_{EE} + J_{EEC} - 1)(J_{II}\tau_{II} + J_{IIC}\tau_{IIC} - \tau_I) - (J_{EI} + J_{EIC})(J_{IE}\tau_{IE} + J_{IEC}\tau_{IEC})) \cdot \\
& ((J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1) - (J_{EI} - J_{EIC})(J_{IE} - J_{IEC})) + \\
& ((J_{EE}\tau_{EE} - J_{EEC}\tau_{EEC} + \tau_E)(J_{II} - J_{IIC} + 1) - (J_{EI}\tau_{EI} - J_{EIC}\tau_{EIC})(J_{IE} - J_{IEC})) \cdot \\
& ((J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1) - (J_{EI} + J_{EIC})(J_{IE} + J_{IEC})) + \\
& ((J_{EE}\tau_{EE} - J_{EEC}\tau_{EEC} + \tau_E)(J_{II} - J_{IIC} + 1) - (J_{EI} - J_{EIC})(J_{IE}\tau_{IE} - J_{IEC}\tau_{IEC})) \cdot \\
& ((J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1) - (J_{EI} + J_{EIC})(J_{IE} + J_{IEC})) + \\
& ((J_{EE} - J_{EEC} - 1)(J_{II}\tau_{II} - J_{IIC}\tau_{IIC} - \tau_I) - (J_{EI}\tau_{EI} - J_{EIC}\tau_{EIC})(J_{IE} - J_{IEC})) \cdot \\
& ((J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1) - (J_{EI} + J_{EIC})(J_{IE} + J_{IEC})) + \\
& ((J_{EE} - J_{EEC} - 1)(J_{II}\tau_{II} - J_{IIC}\tau_{IIC} - \tau_I) - (J_{EI} - J_{EIC})(J_{IE}\tau_{IE} - J_{IEC}\tau_{IEC})) \cdot \\
& ((J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1) - (J_{EI} + J_{EIC})(J_{IE} + J_{IEC}))\ ]
\end{aligned}
\tag{5}
$$

Generally, there are two ways to ensure that a_1/a_0 to be large. The denominator $a_0 = 0$ or the denominator a_0 is finite, but the numerator a_1 is much larger than the denominator , $a_1 \gg a_0$. $a_0 = 0$ corresponds positive feedback mechanism with fine tuning parameters. Here we consider the second constraint $a_1 \gg a_0$ with finite a_0 and show how it corresponds to negative derivative feedback requirement : input from positive and negative feedback pathway is balance in strength and offset in time, especially for faster negative feedback dynamics.

To analyze the expressions above, recognize our assumption that symmetrical network structure leads to symmetrical activities when the system is stable. That means, two sides of the network will be homogeneous with the same firing rate or heterogeneous with "opposite" firing rate. In other words, they will have similar requirements of mathematical expressions for persist activity. In latter competitive condition, due to the firing rate could not be negative, there must be a baseline activity caused by isotropy step-like background input. In the way, the stable activity of four population network is equivalent to the stable activity of two same EI couples in two modes. In mode1, inputs from the contralateral side EEC, IIC, EIC, IEC act as input in the same side, while in mode 2 , inputs from the contralateral side act as inputs in same side, but have the opposite characteristic of excitatory and inhibitory relationship.

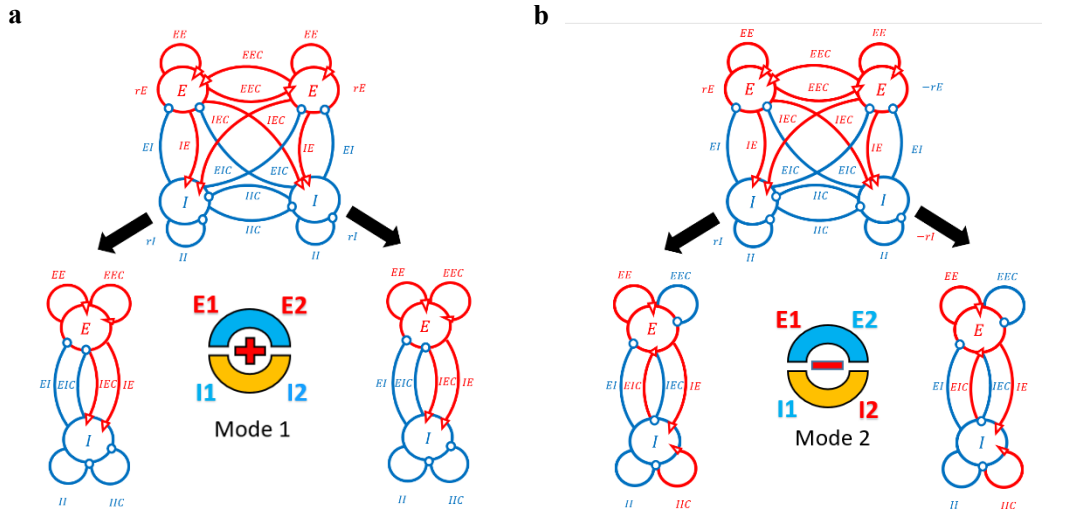

Figure 2. Two Modes of symmetrical activities when the system is stable.**(a)** Mode1. Two sides of the populations have the same activity while the "cross" connections act as ipsilateral connections. **(b)** Mode2. Two sides of the populations have the "opposite" activity while the "cross" connections act as ipsilateral connections with reverse relationship of excitatory and inhibitory.

Mathematical analyze has proved our assumption above. In fact , the expression a_0 of four population is the product of $a_{0,1}$ and $a_{0,2}$ which denote the first coefficient of the characteristic polynomial in figure3(a) and figure3(b).In the same way, a_1 is the product of $a_{1,1}$ and $a_{1,2}$.

$$a_0 = a_{0,1} \cdot a_{0,2} \tag{6}$$

$$a_1 = a_{1,1} \cdot a_{0,2} + a_{1,2} \cdot a_{0,1} \tag{7}$$

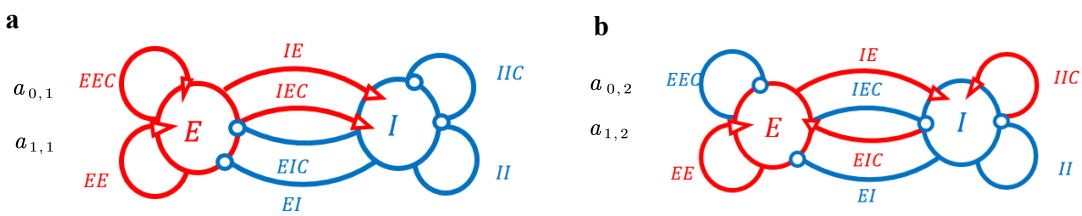

Figure 3. (a) and **(b)** are networks with of multiple feedback pathways in two conditions .

Expressions of $a_{0,1}$, $a_{0,2}$, $a_{1,1}$ and $a_{1,2}$ is :

$$a_{0,1} = (J_{EI} + J_{EIC})(J_{IE} + J_{IEC}) - (J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1) \tag{8}$$

$$a_{0,2} = (J_{EI} - J_{EIC})(J_{IEC} - J_{IE}) - (J_{EE} - J_{EEC} - 1)(J_{IIC} - J_{II} + 1) \tag{9}$$

$$
\begin{aligned}
a_{1,1} = & (J_{EI}+J_{EIC})(J_{IE}+J_{IEC})-(J_{EE}+J_{EEC}-1)(J_{II}+J_{IIC}+1)\cdot \\
& (\tau_{EI}+\tau_{EIC}+\tau_{IE}+\tau_{IEC}+\tau_{EE}+\tau_{EEC}+\tau_{II}+\tau_{IIC}) \\
& +(J_{EE}\tau_{EE}+J_{EEC}\tau_{EEC}+\tau_{E})(J_{II}+J_{IIC}+1)+(J_{EE}+J_{EEC}-1)(J_{II}\tau_{II}+J_{IIC}\tau_{IIC}-\tau_{I}) \\
& -(J_{EI}\tau_{EI}+J_{EIC}\tau_{EIC})(J_{IE}+J_{IEC})-(J_{EI}+J_{EIC})(J_{IE}\tau_{IE}+J_{IEC}\tau_{IEC})
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
a_{1,2} = & (J_{EI}-J_{EIC})(J_{IEC}-J_{IE})-(J_{EE}-J_{EEC}-1)(J_{IIC}-J_{II}+1)\cdot \\
& (\tau_{EI}+\tau_{EIC}+\tau_{IE}+\tau_{IEC}+\tau_{EE}+\tau_{EEC}+\tau_{II}+\tau_{IIC}) \\
& +(J_{EE}\tau_{EE}-J_{EEC}\tau_{EEC}+\tau_{E})(J_{IIC}-J_{II}+1)+(J_{EE}-J_{EEC}-1)(J_{IIC}\tau_{IIC}-J_{II}\tau_{II}-\tau_{I}) \\
& -(J_{EI}\tau_{EI}-J_{EIC}\tau_{EIC})(J_{IEC}-J_{IE})-(J_{EI}-J_{EIC})(J_{IEC}\tau_{IEC}-J_{IE}\tau_{IE})
\end{aligned} \tag{11}
$$

Generally, there are two ways to ensure that a_1/a_0 to be large. The denominator $a_0=0$ or the denominator a_0 is finite, but the numerator a_1 is much larger than the denominator , $a_1 \gg a_0$. $a_0=0$ corresponds positive feedback mechanism and the $a_1 \gg a_0$ with finite a_0 corresponds negative derivative feedback mechanism . For negative derivative feedback, we have $a_0 \ll O(J^4)$, $a_1 \sim O(J^4)$ while $a_{1,}a_0>0$ for stability .

Combined with (6) and (7), this corresponds to :

$$a_{0,1} \ll O(J^2)\ ,\ \ a_{1,1} \sim O(J^2)\ ,\ \ a_{1,1}>0\ ,\ a_{0,2}>0 \ \text{ or} \tag{12}$$

$$a_{0,2} \ll O(J^2)\ ,\ \ a_{1,2} \sim O(J^2)\ ,\ \ a_{1,2}>0,\ a_{0,1}>0 \tag{13}$$

Notice that the first three terms in (12)and (13) correspond the negative derivative feedback requirement of two models in **figure3** separately, while $a_{0,2}>0$ for condition 1 and $a_{0,1}>0$ for condition2 determined the relationship of the two interaction symmetric populations in one four population model, which could also be the “switch” of the two conditions.

We first consider mode 1 .The first three terms of (12) means the leading terms $O(J^2)$ of a_1, a_0 should satisfy:

$$(J_{EI}+J_{EIC})(J_{IE}+J_{IEC})-(J_{EE}+J_{EEC})(J_{II}+J_{IIC}) \ll O(J^2) \tag{14}$$

$$
\begin{aligned}
& ((J_{EI}+J_{EIC})(J_{IE}+J_{IEC})-(J_{EE}+J_{EEC})(J_{II}+J_{IIC}))\cdot \Sigma_i \tau_i \\
& +\ J_{EE}J_{II}(\tau_{EE}+\tau_{II})+J_{EEC}J_{IIC}(\tau_{EEC}+\tau_{IIC})+J_{EEC}J_{II}(\tau_{EEC}+\tau_{II})+J_{EE}J_{IIC}(\tau_{EE}+\tau_{IIC}) \\
& -\ J_{EI}J_{IE}(\tau_{EI}+\tau_{IE})-J_{EIC}J_{IEC}(\tau_{EIC}+\tau_{IEC})-J_{EIC}J_{IE}(\tau_{EIC}+\tau_{IE})-J_{EI}J_{IEC}(\tau_{EI}+\tau_{IEC}) \sim O_+(J^2) \\
& i=EE,EI,IE,II,EEC,EIC,IEC,IIC
\end{aligned} \tag{15}
$$

However , the expressions are too complicated. Noting that the first line of (15) is $\ll O(J^2)$ according to (14), we have

$$
\begin{aligned}
& +\ J_{EE}J_{II}(\tau_{EE}+\tau_{II})+J_{EEC}J_{IIC}(\tau_{EEC}+\tau_{IIC})+J_{EEC}J_{II}(\tau_{EEC}+\tau_{II})+J_{EE}J_{IIC}(\tau_{EE}+\tau_{IIC}) \\
& -\ J_{EI}J_{IE}(\tau_{EI}+\tau_{IE})-J_{EIC}J_{IEC}(\tau_{EIC}+\tau_{IEC})-J_{EIC}J_{IE}(\tau_{EIC}+\tau_{IE})-J_{EI}J_{IEC}(\tau_{EI}+\tau_{IEC}) \sim O_+(J^2)
\end{aligned} \tag{16}
$$

A(16) shows negative feedback (the second line) has faster dynamics than positive feedback (the first line). Furthermore, to find τ_+ and τ_- that represents the time constant of positive feedback pathway and negative feedback pathway, combined with (14), the expression (16) is equal to

$$
\begin{aligned}
& (J_{EE}+J_{EEC})(J_{II}+J_{IIC})\left(\frac{J_{EE}\tau_{EE}+J_{EEC}\tau_{EEC}}{J_{EE}+J_{EEC}}+\frac{J_{II}\tau_{II}+J_{IIC}\tau_{IIC}}{J_{II}+J_{IIC}}\right)- \\
& (J_{EI}+J_{EIC})(J_{IE}+J_{IEC})\left(\frac{J_{EI}\tau_{EIC}+J_{EI}\tau_{EIC}}{J_{EI}+J_{EIC}}+\frac{J_{IE}\tau_{IE}+J_{IEC}\tau_{IEC}}{J_{IE}+J_{IEC}}\right) \sim O_+(J^2)
\end{aligned} \tag{17}
$$

Therefore,(14) and (17) is equivalent to

$$
\begin{aligned}
& (J_{EI}+J_{EIC})(J_{IE}+J_{IEC})/(J_{EE}+J_{EEC})-(J_{II}+J_{IIC}) \ll O(J)\ \ i.e., \\
& (J_{EI}+J_{EIC})(J_{IE}+J_{IEC})/(J_{EE}+J_{EEC})(J_{II}+J_{IIC}) \sim 1
\end{aligned} \tag{18}
$$

$$(J_{EE}+J_{EEC})\left(\frac{J_{EE}\tau_{EE}+J_{EEC}\tau_{EEC}}{J_{EE}+J_{EEC}}+\frac{J_{II}\tau_{II}+J_{IIC}\tau_{IIC}}{J_{II}+J_{IIC}}\right)-$$
$$(J_{EI}+J_{EIC})(J_{IE}+J_{IEC})/(J_{II}+J_{IIC})\left(\frac{J_{EI}\tau_{EIC}+J_{EI}\tau_{EIC}}{J_{EI}+J_{EIC}}+\frac{J_{IE}\tau_{IE}+J_{IEC}\tau_{IEC}}{J_{IE}+J_{IEC}}\right)>0 \tag{19}$$

Denote $J_{+,1}=(J_{EE}+J_{EEC})$ as positive feedback to the excitatory population itself , and $J_{-,1}=(J_{EI}+J_{EIC})(J_{IE}+J_{IEC})/(J_{II}+J_{IIC})$ as negative feedback onto the excitatory population through the inhibitory population. The expression is consistent to $J_{+,1}=J_{-,1}$,$J_{+,1}\tau_{+,1}>J_{-,1}\tau_{-,1}$, the requirement for negative derivative feedback in previous research. $J_{+,1}=J_{-,1}$ is the balance input of positive and negative feedback pathways, $J_{+,1}\tau_{+,1}>J_{-,1}\tau_{-,1}$ is the faster dynamics of negative feedback pathways than positive feedback pathways. $\tau_{+,1}$,$\tau_{-,1}$ are strength-weighted average time constant of positive and negative feedback pathways.

What's more important ,the expression of τ_+ and τ_- have provide us a perspective of synapse integration, that is, different synapses in a same direction(start from which population and arrive to which population) could be integrated as one. The new connection strength is the sum of the original connections, and the time constant of the new connection is the average of original time constants with a strength-weighted scaling factor.

Next, we consider the last term of A(12) , $a_{0,2}>0$. For $a_{0,2}>0$, combined with the balance condition(14), we have:

$$J_{EIC}J_{IE}+J_{EI}J_{IEC}<J_{EEC}J_{II}+J_{EE}J_{IIC} \tag{20}$$

The expression (16) shows the mutual activation of two sides $J_{EEC}J_{II}+J_{EE}J_{IIC}$ is larger than the mutual inhibition of two sides,$J_{EIC}J_{IE}+J_{EI}J_{IEC}$, which reflects the characteristic of condition 1, that is, the activities on both sides are the same.

In the same way, for condition2 $a_{0,2}\ll O(J^2)$,$a_{1,2}\sim O(J^2)$,$a_{1,2}>0$, $a_{0,1}>0$ yields

$$(J_{EE}-J_{EEC})(J_{II}-J_{IIC})-(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})\sim O(J^2) \tag{21}$$

$$(J_{EE}-J_{EEC})(J_{II}-J_{IIC})\left(\frac{J_{EE}\tau_{EE}-J_{EEC}\tau_{EEC}}{J_{EE}-J_{EEC}}+\frac{J_{II}\tau_{II}-J_{IIC}\tau_{IIC}}{J_{II}-J_{IIC}}\right)$$
$$(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})\left(\frac{J_{EI}\tau_{EI}-J_{EIC}\tau_{EIC}}{J_{EI}-J_{EIC}}+\frac{J_{IE}\tau_{IE}-J_{IEC}\tau_{IEC}}{J_{IE}-J_{IEC}}\right)\sim O_+(J^2) \tag{22}$$
$$J_{EE}\neq J_{EEC},J_{II}\neq J_{IIC},J_{EI}\neq J_{EIC},J_{IE}\neq J_{IEC}$$

Or equivalently ,

$$(J_{EE}-J_{EEC})(J_{II}-J_{IIC})/(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})\sim 1 \tag{23}$$

$$(J_{EE}-J_{EEC})\left(\frac{J_{EE}\tau_{EE}-J_{EEC}\tau_{EEC}}{J_{EE}-J_{EEC}}+\frac{J_{II}\tau_{II}-J_{IIC}\tau_{IIC}}{J_{II}-J_{IIC}}\right).$$
$$(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})/(J_{II}-J_{IIC})\left(\frac{J_{EI}\tau_{EI}-J_{EIC}\tau_{EIC}}{J_{EI}-J_{EIC}}+\frac{J_{IE}\tau_{IE}-J_{IEC}\tau_{IEC}}{J_{IE}-J_{IEC}}\right)>0 \tag{24}$$
$$J_{EE}\neq J_{EEC},J_{II}\neq J_{IIC},J_{EI}\neq J_{EIC},J_{IE}\neq J_{IEC}$$

The condition (23) and (24) are similar to $J_{+,2}=J_{-,2}$,$J_{+,2}\tau_{+,2}>J_{-,2}\tau_{-,2}$. $J_{+,2}=(J_{EE}-J_{EEC})$ denotes of net positive feedback that self-inhibition connections are subtracted for excitatory population, while $J_{-,2}=(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})/(J_{II}-J_{IIC})$ is net negative feedback that contralateral side "opposite" connections are subtracted, and divided by self-inhibition that are subtracted the self-excitatory for the inhibitory population. $J_{+,2}\tau_{+,2}>J_{-,2}\tau_{-,2}$ shows in synaptic integration,

strength-weighted time constant that positive one $\tau_{+,2}$ is larger than negative one $\tau_{-,2}$.

As for the last term in (13), $a_{0,2} > 0$, combined with (17),we have

$$J_{EIC}J_{IE} + J_{EI}J_{IEC} > J_{EEC}J_{II} + J_{EE}J_{IIC} \tag{25}$$

The expression (19) shows the mutual activation of two sides $J_{EEC}J_{II} + J_{EE}J_{IIC}$ is less than the mutual inhibition of two sides, $J_{EIC}J_{IE} + J_{EI}J_{IEC}$, which reflects the characteristic of condition 2, that is, the activities of two sides are competitive and be "opposite" with each other on the baseline activity.

To make a conclusion , we derive the two conditions of the four-population network to have negative derivative feedback. $J_+ = J_-$, $J_+\tau_+ > J_-\tau_-$. And the switch of two conditions, the relationship between mutual activation term of two sides $J_{EIC}J_{IE} + J_{EI}J_{IEC}$ and the mutual inhibition term of two sides, $J_{EIC}J_{IE} + J_{EI}J_{IEC}$. $J_{EIC}J_{IE} + J_{EI}J_{IEC} < J_{EEC}J_{II} + J_{EE}J_{IIC}$ means condition 1 and $J_{EIC}J_{IE} + J_{EI}J_{IEC} > J_{EEC}J_{II} + J_{EE}J_{IIC}$ means condition 2.

In addition, synapse integration has provided us a perspective to extend the model. As a corollary, we assume there are ni number of sub-connections in the whole connection i with $i = EE, EEC, EI, EIC, IE, IEC, II, IIC$ and we denote the sub-connections as ij with $j = 1, 2, .., ni$. These sub-connections could be integrated by the following expressions and we can analyze the model as before.

$$J_i = \sum_{j=1}^{ni} J_{ij} \qquad \tau_i = \frac{\sum_{j=1}^{ni} \tau_{ij} J_{ij}}{\sum_{j=1}^{ni} J_{ij}} \tag{26}$$

2.Requirements for positive feedback

Recalling the restrictions of the polynomial coefficients for persist activity $a_1/a_0 \to \infty$. Generally, there are two ways to ensure that $a_1/a_0 \to \infty$. The denominator $a_0 = 0$ or the denominator a_0 is finite, but the numerator a_1 is much larger than the denominator , $a_1 \gg a_0$. We have proved the second condition, a_0 is finite with $a_1 \gg a_0$ corresponds negative derivative feedback. Here, we will prove that how $a_0 = 0$ corresponds positive feedback.

According to (6), $a_0 = 0$ is equal to $a_{0,1} = 0$ or $a_{0,2} = 0$, but according to their expressions (8),(9), $a_{0,1} = 0$ and $a_{0,2} = 0$ are mutually exclusive, so $a_{0,1} = 0$ means $a_{0,2} \neq 0$ and $a_{0,2} = 0$ means $a_{0,1} \neq 0$.Combined with symmetric reduction in previous section and requirements for stability, $a_{0,2}$ and $a_{0,1}$ should > 0 in two modes.

From the analysis above, for mode 1 that each side have mutual activation than inhibition, we have

$$(J_{EI} + J_{EIC})(J_{IE} + J_{IEC}) = (J_{EE} + J_{EEC} - 1)(J_{II} + J_{IIC} + 1) \tag{27}$$

$$(J_{EI} - J_{EIC})(J_{IE} - J_{IEC}) > (J_{EE} - J_{EEC} - 1)(J_{II} - J_{IIC} + 1) \tag{28}$$

Or equivalently,

$$(J_{EE}+J_{EEC})-\frac{(J_{EI}+J_{EIC})(J_{IE}+J_{IEC})}{(J_{II}+J_{IIC}+1)}=1 \tag{29}$$

$$J_{EEC}J_{II}+J_{EE}J_{IIC}>J_{EIC}J_{IE}+J_{EI}J_{IEC} \tag{30}$$

The form (30) shows the mutual activation is larger than mutual inhibition, corresponds the requirements of mode1. Denote $W_{pos,1}=(J_{EE}+J_{EEC})-(J_{EI}+J_{EIC})(J_{IE}+J_{IEC})/(J_{II}+J_{IIC}+1)$ as the strength of pure positive feedback in mode1. (29) shows the difference of positive feedback and negative feedback that could precisely cancel the neural leakiness.

For mode2, we have

$$(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})=(J_{EE}-J_{EEC}-1)(J_{II}-J_{IIC}+1) \tag{31}$$

$$(J_{EI}+J_{EIC})(J_{IE}+J_{IEC})>(J_{EE}+J_{EEC}-1)(J_{II}+J_{IIC}+1) \tag{32}$$

Or equivalently,

$$(J_{EE}-J_{EEC})-\frac{(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})}{(J_{II}-J_{IIC}+1)}=1 \tag{33}$$

$$J_{EEC}J_{II}+J_{EE}J_{IIC}<J_{EIC}J_{IE}+J_{EI}J_{IEC} \tag{34}$$

The form (34) shows the mutual inhibition is larger than mutual activation, corresponds the requirements of mode2. Denote $W_{pos,2}=(J_{EE}-J_{EEC})-(J_{EI}-J_{EIC})(J_{IE}-J_{IEC})/(J_{II}-J_{IIC}+1)$ as the strength of pure positive feedback in mode2. (33) shows the difference of net positive feedback that self-inhibition connections are subtracted for excitatory population, and net negative feedback that contralateral side "opposite" connections are subtracted , could precisely cancel the neural leakiness.

3. Measurement of negative derivative feedback and positive feedback

To derive a measurement of the amounts of negative derivative feedback and positive feedback, we examine the expression for the longest time constant of decay of network activity ,τ_{net}. Mathematically, the longest time constant is determined by the smallest eigenvalue of the system: $\tau_{net}=-1/\lambda$.Combine with (3), if one eigenvalue is sufficiently small, then $\tau_{net}\approx a_1/a_0$.Recall that $a_0=a_{0,1}\cdot a_{0,2}$, $a_1=a_{1,1}\cdot a_{0,2}+a_{1,2}\cdot a_{0,1}$,$\tau_{net}\approx a_1/a_0=a_{1,1}/a_{0,1}+a_{1,2}/a_{0,2}$ .For modes 1 that $a_{0,1}\ll O(J^2)$,we have $a_1/a_0\approx a_{1,1}/a_{0,1}$.The reason of we ignore second term $a_{1,2}/a_{0,2}$ is that $a_{0,1}\ll O(J^2)$ and $a_{0,2}\ll O(J^2)$ are mutually exclusive in expressions. In fact, $a_{0,1}\ll O(J^2)$ will lead to $a_{0,2}\sim O(J^2)$ and in this way $a_{1,1}/a_{0,1}\gg a_{1,2}/a_{0,2}$. In the same way, for mode 2 we have the expression $a_1/a_0\approx a_{1,2}/a_{0,2}$.

Concretely ,combined with the expressions we have analyzed in previous sections, the expressions of decay time constant of mode 1and mode 2 are :

$$\tau_{net,1}=\frac{J_{der,1}(\tau_{+,1}-\tau_{-,1})+C_1}{1-W_{pos,1}} \tag{35}$$

$$\tau_{net,2}=\frac{J_{der,2}(\tau_{+,2}-\tau_{-,2})+C_2}{1-W_{pos,2}} \tag{36}$$

Where $J_{der,1}=J_{+,1}=J_{-,1}$, $J_{der,2}=J_{+,2}=J_{-,2}$ and C_1,C_2 are constants of low orders.

4.Stability Analysis

In the above section , we analyze the requirements for negative derivative feedback with one eigenvalue is equal to zero. In general, we should also promise that other eigenvalues of the system

less than zero to keep the system stable. That means , In the Routh stability criterion, if all the coefficients in the first column of the Rouse table (shows bellow) are positive, then the system is stable:

$$
\begin{array}{lllllllll}
x^{n} & a_{n} & a_{n-2} & a_{n-4} & a_{n-6} & \cdot & \cdot & \cdot & \cdot \\
x^{n-1} & a_{n-1} & a_{n-3} & a_{n-5} & a_{n-7} & \cdot & \cdot & \cdot & \cdot \\
x^{n-2} & b_{1} & b_{2} & b_{3} & b_{4} & \cdot & \cdot & \cdot & \cdot \\
x^{n-3} & c_{1} & c_{2} & c_{3} & c_{4} & \cdot & \cdot & \cdot & \cdot \\
x^{n-4} & d_{1} & d_{2} & d_{3} & d_{4} & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & &
\end{array}
$$

$$
\begin{array}{lll}
b_{1}=\dfrac{-1}{a_{n-1}}\begin{vmatrix} a_{n} & a_{n-2} \\ a_{n-1} & a_{n-3} \end{vmatrix}, & c_{1}=\dfrac{-1}{b_{1}}\begin{vmatrix} a_{n-1} & a_{n-3} \\ b_{1} & b_{2} \end{vmatrix}, & d_{1}=\dfrac{-1}{c_{1}}\begin{vmatrix} b_{1} & b_{2} \\ c_{1} & c_{2} \end{vmatrix} \quad \cdot\ \cdot \\
b_{2}=\dfrac{-1}{a_{n-1}}\begin{vmatrix} a_{n} & a_{n-4} \\ a_{n-1} & a_{n-5} \end{vmatrix}, & c_{2}=\dfrac{-1}{b_{1}}\begin{vmatrix} a_{n-1} & a_{n-5} \\ b_{1} & b_{3} \end{vmatrix}, & d_{2}=\dfrac{-1}{c_{1}}\begin{vmatrix} b_{1} & b_{3} \\ c_{1} & c_{3} \end{vmatrix} \quad \cdot\ \cdot \\
b_{3}=\dfrac{-1}{a_{n-1}}\begin{vmatrix} a_{n} & a_{n-6} \\ a_{n-1} & a_{n-7} \end{vmatrix}, & c_{3}=\dfrac{-1}{b_{1}}\begin{vmatrix} a_{n-1} & a_{n-7} \\ b_{1} & b_{4} \end{vmatrix}, & d_{3}=\dfrac{-1}{c_{1}}\begin{vmatrix} b_{1} & b_{4} \\ c_{1} & c_{4} \end{vmatrix} \quad \cdot\ \cdot \\
\cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot \\
\cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot
\end{array}
\tag{23}
$$

The first two rows of a Routh table consist of the coefficients of the characteristic polynomial $char(x)=a_0+a_1x+a_2x^2+...+a_nx^n$.The first row consists of the coefficients of odd number a_1, a_3,a_5.. The second row consists of the coefficients of even number a_0,a_2,a_4,a_6.. Further, we need to exclude the selected eigenvalue because this eigenvalue is equal to 0, and other eigenvalues less than 0. Denote the selected eigenvalue as λ_1, the characteristic polynomial is equal to :

$$
\begin{aligned}
char(x)=&(x-\lambda_1)\left(x^{n-1}+(a_{n-1}+\lambda_1)x^{n-2}+(a_{n-2}+\lambda_1(a_{n-1}+\lambda_1))x^{n-3}\right. \\
&\left.+(a_{n-3}+\lambda_1(a_{n-2}+\lambda_1(a_{n-1}+\lambda_1)))x^{n-4}+....\right)+a_{const}
\end{aligned}
\tag{24}
$$

Because λ_1 is one root of the characteristic polynomial $char(x)=0$, the reminder constant term a_{const} is equal to zero. In particular , when the chosen eigenvalue λ_1 approach to zero, the condition comes to

$$
newchar(x)=x^{n-1}+a_{n-1}x^{n-2}+a_{n-2}x^{n-3}+...+a_2x+a_1=0 \tag{25}
$$

$newchar(x)$ is the new characteristic polynomial of the reminding eigenvalue. To promise all reminding eigenvalue less than zero, we use Routh table for $newchar(x)$.

From the symmetry dimension reduction, the four-population model and the two-population model in two modes are equivalent in one same persist activity, therefore, they would have similar requirements for negative derivative feedback and requirements for stability. However, in one hand, the other coefficients of two-population model and four-population model has more differences whether in total numbers and expressions, and we are unable to give an analytical expression of the stability equivalence of the two models. In the other hand, even for the stability requirements of the two-population model , the expression is going to become really complicated and is unable to be analytical discussed. Therefore, we apply Routh analysis to numerical simulation.

5.Robustness against perturbations

cording to previous analysis, negative derivative feedback enables robustness against many perturbations. Here we will analyze the influence of the perturbations in the aspect of changes in the intrinsic neural gains, changes in specific neural sub-population's gains, changes in excitatory or inhibitory synapses and inactivation of the neurons.

Except that the changes of synapses could be represented by the scaling of the connection strength, other perturbations such as changes in neural gains and inactivation of neurons, could be either represented by the scaling of the connection strength. The reason is that the neural gains of one population is linear to the connections it receives, and the inactivation of one population is negatively linear to all the connections it has. We denote a multiplicative factor S_{ij} with $i,j = E_1, E_2, I_1, I_2$ to present fluctuations of one synaptic connection strength. After perturbations, the connection strength become $S_{ij} J_{ij}$.

Firstly, we consider the perturbations that causes symmetric changes of synapses and would not break the symmetric structure of the network. Therefore, the basic form of the coefficients (4) and (5) is still available for the new system with perturbations. The simplest condition is that intrinsic neural gains change in the entire network, corresponds to a uniform multiplicative factor S to every connection. Under the perturbation, a_0 and a_1 expressions of (4) and (5) become $a_0' = S^4 a_0$ and $a_1' = S^4 a_1$ when connection strength J are large and the coefficients a_0, a_1 are represented by the highest order terms. Because they are scaled similarly, a_1/a_0 will not change for the requirement of negative derivative feedback.

Similarly, changes with excitatory and inhibitory characteristics are symmetrical either, such as inactivation or gains change of E/I populations and synapses. Because the analysis method is similar, we derive a table to represent the three kinds of perturbations in this case. They are changes of E/I neural population gains , changes of E/I connections and inactivation of E/I populations.

Table1. Analysis of the symmetrical perturbations with E/I characteristic.

Perturbation Type	Perturbation Expression	Explanation of the expression	Expression of a_0' and a_1'	Expression of a_1'/a_0'
Changes of E/I neural gains	$S_E J_{Em}, S_E J_{EmC}$ $S_I J_{In}, S_I J_{InC}$, $m,n = E,I$	$S_E J_{Em}$ and $S_I J_{In}$ denote connections that E or I populations gained from other populations under the perturbation.	$a_0' = S_I{}^2 S_E{}^2 a_0$ $a_1' = S_I{}^2 S_E{}^2 a_1$ With large J s	$\frac{a_1'}{a_0'} = \frac{a_1}{a_0}$
Changes of E/I synapses	$S_E J_{mE}, S_E J_{mEC}$ $S_I J_{nI}, S_I J_{nIC}$, $m,n = E,I$	$S_E J_{mE}$ and $S_I J_{nI}$ denote the E or I synapses that originated from E or I populations under the perturbation.	$a_0' = S_I{}^2 S_E{}^2 a_0$ $a_1' = S_I{}^2 S_E{}^2 a_1$ With large J s	$\frac{a_1'}{a_0'} = \frac{a_1}{a_0}$
Inactivation of E/I populations	$S_E{}^2 J_{EE}, S_E{}^2 J_{EEC}$ $S_I{}^2 J_{II}$, $S_I{}^2 J_{IIC}$ $S_E S_I J_p$ $p = EI, IE,$ EIC, IEC	In inactivation perturbations, neural gains and E/I synapses has changed together, so the square or multiple scaling factor shows the two aspects of the changes: neural gains and synapses.	$a_0' = S_I{}^4 S_E{}^4 a_0$ $a_1' = S_I{}^4 S_E{}^4 a_1$ With large J s	$\frac{a_1'}{a_0'} = \frac{a_1}{a_0}$

The table above shows a robustness against perturbations of symmetrical E/I changes, because the final expression a_1'/a_0' is same to a_1/a_0 with large J s and the coefficients a_0, a_1 are represented by the highest order terms.

Secondly ,we consider the three kinds of perturbations in specific sub-populations: changes in neural gains of one specific sub-population, changes in synapses from one specific sub-population and inactivation of one specific sub-population. Because the change of one specific population has disturbed the symmetrical structure of the network, the a_0 and a_1 expressions (4) and (5) would not make sense at all. Here we derive the expression a_0 of the network without symmetrical structure.

$$\begin{aligned} a_0 = & ((J_{E1E1}-1)(J_{I1I1}+1)-J_{E1I1}J_{I1E1})((J_{E2E2}-1)(J_{I2I2}+1)-J_{E2I2}J_{I2E2})+ \\ & (J_{E1I2}J_{I1E2}-J_{E1E2}J_{I1I2})(J_{E2I1}J_{I2E1}-J_{E2E1}J_{I2I1})- \\ & ((J_{I1I1}+1)J_{E1E2}-J_{I1E2}J_{E1I1})((J_{I2I2}+1)J_{E2E1}-J_{I2E1}J_{E2I2})- \\ & ((J_{E1E1}-1)J_{I1I2}-J_{I1E1}J_{E1I2})((J_{E2E2}-1)J_{I2I1}-J_{I2E2}J_{E2I1})- \\ & ((J_{I1I1}+1)J_{E1I2}-J_{I1I2}J_{E1I1})((J_{E2E2}-1)J_{I2E1}-J_{I2E2}J_{E2E1})- \\ & ((J_{E1E1}-1)J_{I1E2}-J_{E1E2}J_{I1E1})((J_{I2I2}+1)J_{E2I1}-J_{I2I1}J_{E2I2}) \end{aligned} \tag{26}$$

The expression of a_1 is too complex to show in the section, but it is closely related to the expression of a_0 ,that is ,every highest-order item of a_0 and a_1 is an one to one correspondence. The difference is every item in a_1 is multiplicative with a sum of a part of the time constants. Therefore, the conclusion is that if each highest order term of a_0 has a common product of the all connection strength scaling factors ΠS , then each highest order term of a_1 also has a common product factor ΠS and in the condition we have $a_0' = \Pi S a_0, a_1' = \Pi S a_1$ when J s are large. So, the key point is if each highest term in (20) could be multiplied with a common product of the scaling factors under these three kinds of perturbations. To show influence of different perturbations, we made a table similar to Table1, and take the sub-population $E1$ as an example.

Table2. Analysis of the perturbations of a specific population. (Take the population $E1$ as an example).

Perturbation Type	Perturbation Expression	Explanation of the expression	Expression of a_0' and a_1'	Expression of a_1'/a_0'
Changes of $E1$ neural gains	$S_{E1}J_{E1m}$ $m=E2,I1,I2$	$S_{E1}J_{E1m}$ denote connections that $E1$ population gained from the three other populations under the perturbation.	$a_0'=S_{E1}a_0$ $a_1'=S_{E1}a_1$ With large J s	$\frac{a_1'}{a_0'}=\frac{a_1}{a_0}$
Changes of synapses from $E1$	$S_{E1}J_{mE1}$ $m=E2,I1,I2$	$S_{E1}J_{E1m}$ denote the synapses that originated from $E1$ population under the perturbation.	$a_0'=S_{E1}a_0$ $a_1'=S_{E1}a_1$ With large J s	$\frac{a_1'}{a_0'}=\frac{a_1}{a_0}$
Inactivation of $E1$ population	${S_{E1}}^2J_{E1E1}$ $S_{E1}J_{E1p},S_{E1}J_{qE1}$ $p,q=E2,$ $I1,I2$	In inactivation perturbations, neural gains and synapses of one population both changes. $S_{E1}J_{E1p},S_{E1}J_{qE1}$ denote the changes of neural gains of $E1$ and synapses from $E1$. ${S_{E1}}^2J_{E1E1}$ shows the two aspects of influences together.	$a_0'={S_{E1}}^2a_0$ $a_1'={S_{E1}}^2a_1$ With large J s	$\frac{a_1'}{a_0'}=\frac{a_1}{a_0}$

From the expressions above, we conclude that every perturbation of a specific subpopulation could not disturb the form of a_1/a_0 so the network is robust to the perturbations.

Result

1. Persist activity and ramp output

In above sections, we analyze the requirements for negative derivative feedback. To make a

conclusion, negative derivative feedback arises in two firing modes in the completely symmetric structure, mode one is that the firing rates of two sides are exactly the same, modes two shows that two sides have the "opposite" activity on the baseline of background activity. To show the results, firstly, we simulate the basic firing patterns with two different types of external input, the graded pulse-like and step-like input. As shown in the figure4 (A) and (C), two different modes are both capable for maintaining persist activity after the pulse-like input, as a result of the negative derivative feedback mechanism. And the final neural activity has the same proportion as the graded input. In the figure4 (B) and (D), the model have the ramp-like activity under step-like input, and the ramp could be acted as an integrator of the step input with time.

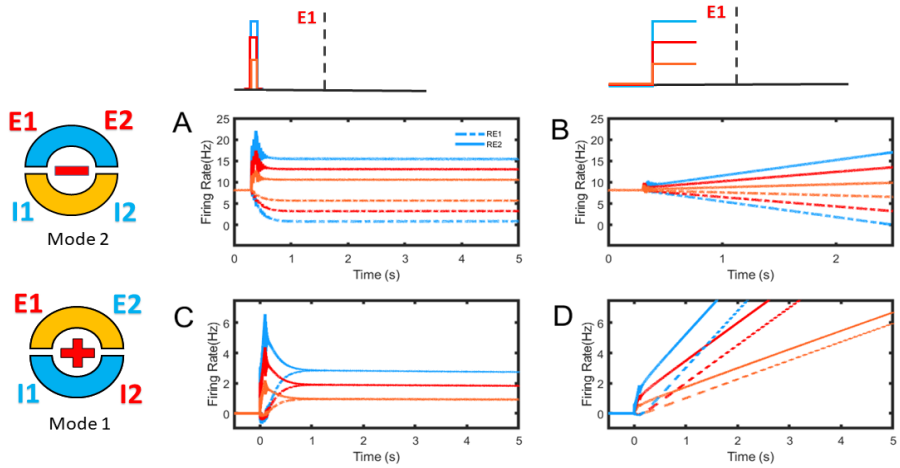

Figure4. The activity of two modes under the graded pulse-like and step-like input. Two modes are shown as a ring at left, while the plus and minor mean the mutual excitation and inhibition relationship of two sides. Shape of external input are shown on the top.**(A) and (C):** The two sides could maintain persist neural activity after external input, and when they become stable, the activity of two sides are same in mode1 and "opposite" in mode2 on the baseline of background. However, before reaching stability, the two modes are not symmetrical: opposite or same. **(B) and (D):** The ramp activity under step-like input, the ramp shows the time integration of the step input. There is an asymmetric bias because of the persist action of asymmetric external input.

2. Symmetric activity and asymmetric bias

Another important feature of the networks is symmetry. Symmetry means complete same or opposite at the baseline of background activity, corresponding to the two firing modes we discussed in the previous sections. Generally speaking, symmetrical structure will lead to symmetrical activity, but with no influence of external input. Only in this way, the dynamic characteristics will completely depend on the internal structural parameters such as connection weight and time constant. In other words, on the one hand, if the external input is asymmetric, the network will have asymmetric activity during the action time of the asymmetric external input. Until it decays to zero or is removed, the system will eventually become stable and symmetrical. (**Fig.4 A,C Fig.5 A,D**). On the other hand, if the external input is symmetrical, the network activity

will always be symmetrical whether it reaches stability or not(**Fig.5 B,C,E,F**). Because the two factors that determine the dynamic characteristics of the system - external input and internal structure are both symmetrical. In addition, symmetric external input also means two kinds of patterns, be completely same or be opposite.

Here we show how two activity modes respond to two kinds of symmetric pulse-like input. For mode 2 that both sides inhibit each other and the inhibition is same, final effect caused by positive unilateral input to E1 is equivalent to negative unilateral input to E2. As a result, the final overall effect of symmetrical opposite external inputs is double of each unilateral input (**Fig 5.A,B**).However, if the external inputs are symmetrically same, effects of two unilateral input are opposite and will offset each other because of the equivalent mutual inhibition. Therefore, finally the activity comes to zero.(**Fig 5.C**)

In the same way, for mode 1 that both sides excite each other, the symmetrical same input will double the unilateral effect (**Fig 5.D,E**) and symmetrical opposite input will offset each other because the equivalent mutual excitation. (**Fig 5.F**)

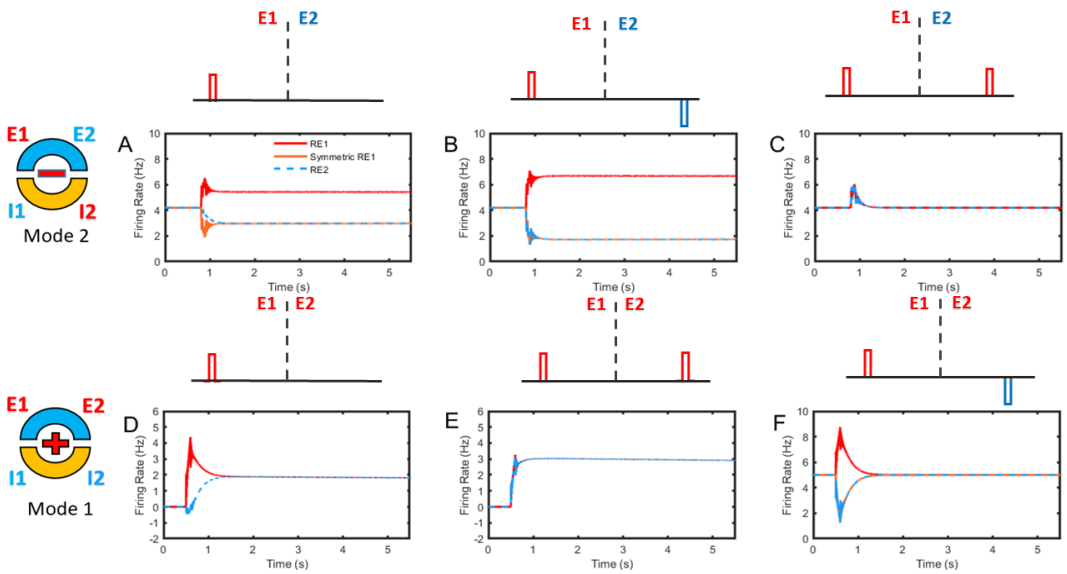

Figure4. The reaction activity of two modes responds to the symmetrical pulse-like external inputs of two patterns. **A,D:** The activity caused by unilateral external input. The network has asymmetric activity during the action time of the asymmetric external input. **B,E:** External input is consistent with the intrinsic structure of two modes. Effect of each unilateral input is same so the overall activity is sum to be double as the activity caused by unilateral input of each side. **C,F:** Each external unilateral input causes opposite activity, and finally offset each other.

Unlike pulse-like external input, the action time of asymmetric step-like input is continuous. As a result, the asymmetric bias of the activities RE1 and RE2 will exist during the simulation (**Fig5. B,D Fig 6.A,D**).But if the step input is symmetric, the activity will keep symmetric all the time. Next we consider the reaction activity of two modes responds to two symmetric step-like input.

For mode 2, because of the equivalent mutual inhibition, ramp output caused by positive

unilateral step input to E1 is equivalent to negative unilateral step input to E2. The symmetric opposite step-like input, consistent with the intrinsic structure, will double the slope of the ramp activity. (**Fig 6.A,B**). Similarly, if the step inputs are symmetrical same, there are no increasing activity because same unilateral step inputs lead to opposite slope of the ramp and they offset each other (**Fig 6.C**). It is worth noting that the offset of increasing activity does not mean decay to zero, the activity before offset happens will be maintain. The question is, what's the source of this part of activity? In fact, as shown before, ramp of RE1 and RE2 caused by unilateral step input is asymmetrical, they have opposite slope but different original states, lead to an asymmetrical bias(**Fig 5.B, Fig 6.A**). The offset only take place in the slope, but not the activity caused by different original states.

Physically to see, every synapse has it time constant and delay. When the symmetric same input be injected to symmetrical E1 and E2, they will instantaneously trigger the same activity and at beginning mutual inhibition has not yet occur because of the synaptic time delay. After that, mutual inhibition happens and the increasing activity offset each other by the opposite slope but the activities before the mutual inhibition happens are left and maintained. This kind of persist activity in mode2 act as the same persist activity of mode1 caused by symmetrical same pulse like input(**Fig 4.E**). And it is also the underlying mechanism of background activity in compete networks.

For mode 1 with mutual excitation on both sides, the positive step input with symmetry will double the slope.(**Fig 6.D,E).** For symmetrical opposite step-like input, each unilateral step-like input will cause ramp activities of opposite slope and they will be offset by mutual excitation, but initial opposite activity before the happening of mutual excitation will be maintained(**Fig 6.F)**.

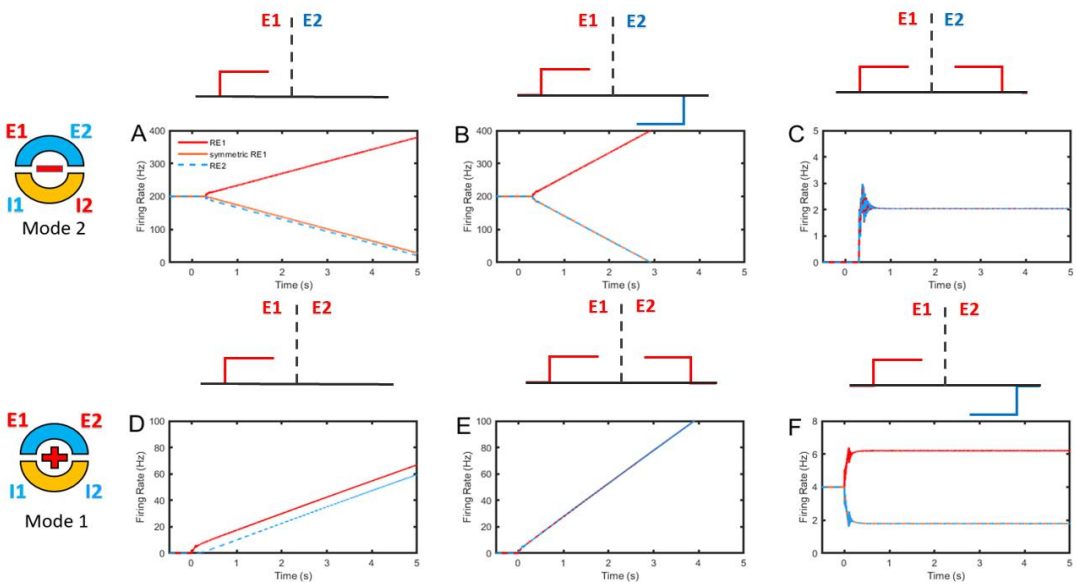

Figure6. The reaction activity of two modes responds to the symmetrical step-like external inputs of two patterns. **A,D:** The activity caused by unilateral external input. The network has asymmetric activity during all time because the continuous action of asymmetrical step-like input. **B,E:** External input is consistent with the

intrinsic structure of two modes. Slope of overall activity is twice as the activity caused by each unilateral step input. **C,F:** Each unilateral external input causes opposite slope to other side, and finally offset each other. Therefore, the overall activity will not increase but the original activity that happens before mutual reaction and offset is left and maintained.

3.Difference in firing pattern compared with two-population model

Our analyze of the network has contained an important dimension reduction. If the system is stable, in two modes, the network activity is equivalent to the activity of two same network that consists of two populations. Mathematical analysis has shown the correlation of four population model and two population model, that is, the requirements for persist activity are same (the same a_1/a_0 in two modes). To show the equivalence and correlation, we compare the firing patterns of two models. The network parameters of two population model are according to the dimension reduction of the four-population model in two modes, separately.

The result shows the equivalence persist activity of the two networks in two modes(**Fig7.A,B**). Equivalence means that they have all same firing rate and synaptic inputs but only in stable state. In other words, for persist activities, if we set the firing rates RE1 for four population model and rE for two population model as the same value, then they will have same firing rates of inhibitory populations RI1 and rI, the same synaptic input SEE and sEE, etc. However, the equivalence of persist activity not means they are actually same all the time. They have different unstable original states if triggered by same external inputs, and will lead to different firing patterns when stable. Conversely, if the two model has same firing patterns when stable, that means the external inputs to the two model are not necessarily equal.

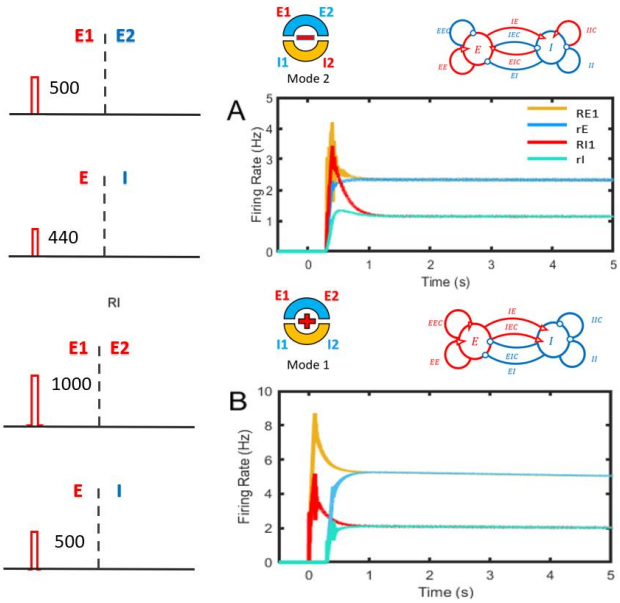

Figure7. The comparison of four population model and two population model. They have equivalence firing rates and synaptic inputs but different original unstable states and triggered by different external inputs. **A.** Two model's firing rate of mode2. **B.** Two model's firing rate of mode1.

4.Roubtness and perturbations

We have analyzed why negative derivative feedback could be robust to such perturbations. In

the section, we will test the model with symmetric perturbations in **Table1** and shows the difference of positive feedback with non-robustness. To make the simulation result more intuitive, we take mode2 to represent the perturbations and robustness, while these figure shows the activity of E1 and E2 with no changes(blue) and perturbations(orange). We first consider changes of intrinsic neural gains that all populations receives more or less synaptic input. We set the perturbation as 5% increase, the result shows positive feedback is unable to keep persist activity and positive feedback has been over strengthened(**Fig8 A'**). If the difference positive feedback and negative feedback strengths is equal to1, the difference is greater than 1 after strengths are increasing as the same multiple. But for negative derivative feedback that positive feedback and negative feedback are balanced in strength, the relationship has not been changed.

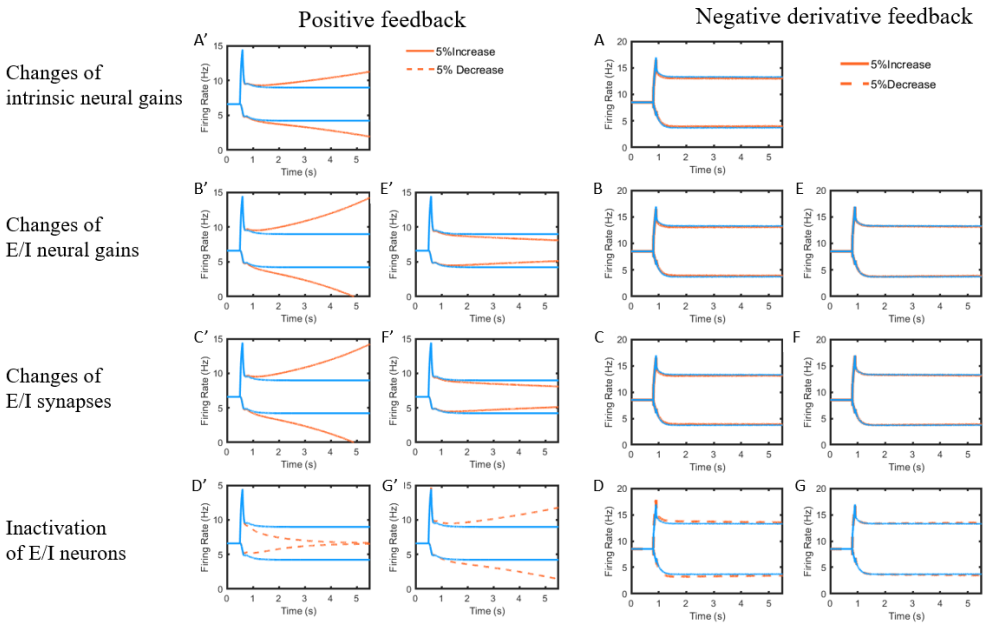

Figure8 The reaction of positive feedback and negative feedback model to the symmetric perturbations.

Secondly, we consider the neural gains changes and synaptic changes of excitatory or inhibitory populations (left column for excitatory changes and right column for inhibitory changes). The two kinds of perturbation have similar mathematical expressions, as a result, they have the same effect to the activities(**Figure 8. B',C',E',F') .** Obviously, the positive feedback model has no robustness, neither excitatory or inhibitory changes will disturb the fine tuning. In contrast, negative derivative feedback has robustness to the perturbation because the balance condition (See in Method) has not changed.

Finally, we consider inactivation of excitatory or inhibitory populations. For positive feedback model, the inactivation will obviously disturb the fine tuning and lead to perturbations. For negative feedback model, activity has not changed and the result shows the robustness.

The asymmetric perturbation is mainly about changes of sub-population. According to mathematical analysis in Method, perturbations of one subpopulation has not changed basic a_1/a_0 expression for negative derivative feedback, but has changed a_0 and disturbed $a_0 = 0$ for positive feedback .

To show the robustness and perturbations, we simulate the activity under the changes of E1 and I1 neural gains, changes of synapses of E1 and I1 and inactivation of E1 and I1.All results shows that negative derivative feedback has robustness against these perturbations, but positive feedback is disturbed by these perturbations and could not maintain persist activity again.

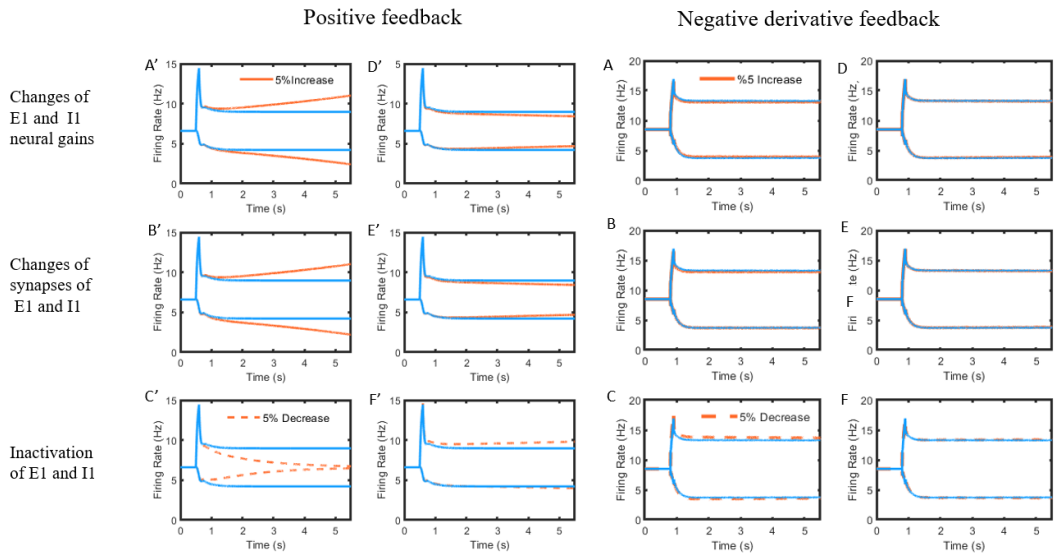

Figure9 The reaction of positive feedback and negative feedback model to the asymmetric perturbations of one specific sub-population.

5.Simple verification of synapse integration

In our analysis, if there are many connections from one specific population to the other specific population, the multiple connections could be integrated as one new connection. The integrated connection strength is the sum of the original connections, and the time constant of the integrated connection is the average of original time constants with a strength-weighted scaling factor. However , the method of synapse integration is only capable for our analysis of the coefficient a_1 and a_0 for the existence of persist activity. In order to examine the consequence of synapse integration on other aspects, we simulated the following model. There are two connections from I1 to E1, $EI1$ and $EI2$. The sum of connection strengths of $EI1$ and $EI2$ is equal to non-synapse integration model, $J_{EI1}+J_{EI2}=J_{EI}$ and the synaptic time constant of EI is the strength weighted average of $EI1$ and $EI2$: $\tau_{EI}=(J_{EI1}\tau_{EI1}+J_{EI2}\tau_{EI2})/(J_{EI1}+J_{EI2})$ .We examine the firing patterns of the two model with the same external input, and the result shows synapse integration really make sense because the two model act as nearly same. This shows that synapse integration is not only a simplification method for the analysis of the existence of persist activity, but also reveals some underlying dynamic relationships.

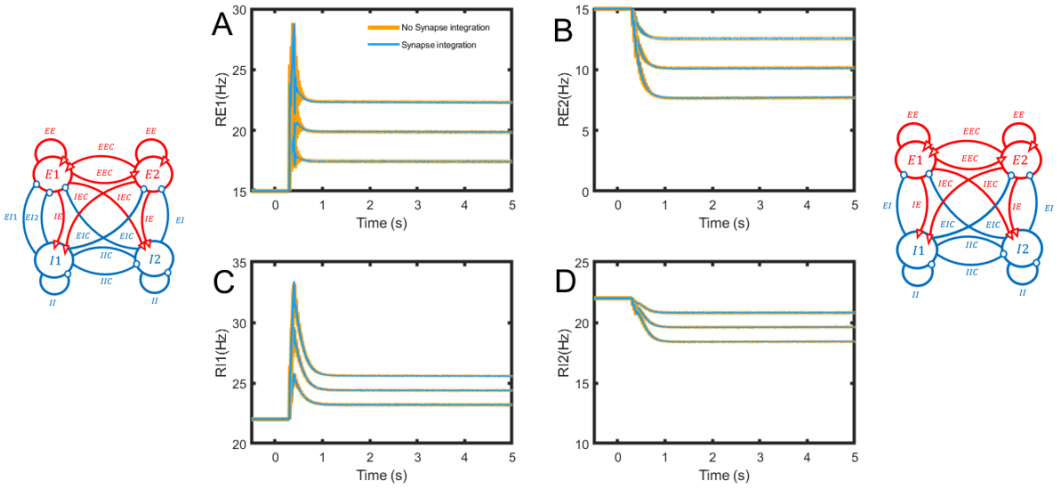

Figure10 The comparison of the activity of synapse-integration network and no synapse-integration network. **A,B,C,D**: The activity of E1,E2,I1,I2 reacted to the same external input to each model. The result shows integrated synapse act as really the original synapse in all aspect of the activities.

Discussion

Our result shows a new competitive neural network structure, which is composed of two symmetrical EI populations, and all populations are completely connected. The structure provides not only feasibility for the negative derivative feedback that has robustness to many common perturbations but also more biological rationality.

We have an important assumption of the network that without the influence of external input, the symmetrical structure will lead to the symmetrical activities, because the final state of the network is determined by the structural parameters. The assumption could provide us a perspective of dimension reduction that four-population network with persist activity is equivalent to two same two-population networks on both sides.

Mathematical analysis further shows that this dimension reduction will lead to two activity modes : mode 1 shows the neural activities on both sides are symmetrical same, and mode 2 shows neural activities on both sides are symmetrically opposite. Because the firing rate cannot be negative, “opposite” means that the activity has a baseline is not equal 0. **Fig6.C** shows that for modes2, the activity baseline can be caused by the same positive step inputs on both sides . The switch of two modes is the relationship of excitation and inhibition among two sides. For mode 1, mutual activation on both sides is greater than mutual inhibition, and for mode 2, mutual inhibition on both sides is greater than mutual activation.

As a verification, the simulation shows the equivalence in persist activity of four population model and two population model **(Fig.7)**. They have equivalence firing rates and synaptic inputs but different original unstable states. Therefore, to trigger the equivalent persist activity of two models, they need different external inputs.

The symmetry dimension reduction can provide us a simplification to analyze negative derivative feedback because two-population networks and four-population networks have the

same requirements for negative derivative feedback. To illustrate the requirements in specific, negative derivative feedback needs balance in positive and negative feedback strengths, but lager arrival time for positive feedback than negative feedback. (14),(15) show the expressions that $a_1/a_0 \to \infty$ and $a_1 \gg a_0$ with infinite a_0. However, the expression is too complicated to make a conclusion. We simplify the expressions to the new form: $J_{+,1} = J_{-,1}, \tau_{+,1} > \tau_{-,1}$ for mode 1 and $J_{+,2} = J_{-,2}, \tau_{+,2} > \tau_{-,2}$ for mode2, which correspond the traditional model that one direction(from and to which population) just has one connection. The simplification has also suggested us a perspective of synapse integration that multiple connections in one direction could be integrated as one. The integrated connection strength is the sum of the original connections, and the time constant of the integrated connection is the average of original time constants with a strength-weighted scaling factor. Therefore, we could extend the four-population model with multiple connections between two specific sub-populations. The simulation **Fig10.** shows that multiple connections and the integrated synapse act as the same. However, the biological basis of synaptic integration and the mathematical principle behind it still need to be confirmed in the future experiments and analysis.

One of the characteristics of negative derivative feedback is it has robustness to many common perturbations. In nerver system, common perturbations are changes of neural gains, changes of synapses and inactivation of neurons. Combined with the model structure, we divide these perturbations into symmetric perturbations caused by changes of E/I population on both sides and asymmetric perturbations caused by changes of one sub-population. All the perturbations could be mathematically represented by a multiplicative scaling factor to connection strengths. Analyze shows these perturbations could not changes the basic requirements for negative derivative feedback, simulations also show the strong robustness to these perturbations as verification. However, by comparison, positive feedback has no robustness to any kind of perturbations.

External inputs have many influences to activities. The persist activity is proportional to graded pulse-like external input, and the slope of ramp-like activity is proportional to graded step-like input. If the external input is asymmetric, the activity of the system is also asymmetric during its action. The asymmetric bias is clear when the network is activated by step-like unilateral external input**(Fig.6 A,D)**. If the external input is symmetric, the activity will always be symmetric. What's more, the symmetric same external input will double the activity caused by unilateral input in mode 1 because mutual activation, but in mode2, because of the equivalent mutual inhibition, the effect of each unilateral input will offset each other. In contrast, the symmetric opposite external input will double the activity caused by unilateral input in mode 2 it is consistent with the mutual inhibition structure, but effect of each unilateral input will offset, because of the equivalent mutual activation will weaken the difference between two sides.

Symmetry is important for competition network. In addition to the symmetric model, the question is, what's the requirements for negative derivative feedback of asymmetric four-population model? Because it could not be dimensionally reduced, a_0 expression (25) shows the restrictions to find

positive feedback pathway and negative feedback pathways on a global perspective. When there are more than two populations, positive feedback and negative feedback pathways are not well defined, so we could not draw an intuitive conclusion.

Positive feedback is used in many traditional competition networks to maintain the activity, such as working memory and neural integrator. Our model is still applicable to traditional positive feedback models by symmetric dimensional reduction, and still has more biological rationality at this level. Besides, for some cases, negative derivative feedback may not be applicable. The model has a wide range of practicability, but how to apply the model to experiments and explain physiological recordings are still needs further exploration.

Reference

[1] Major.G , D.Tank. Persistent neural activity: prevalence and mechanisms[J]. *Current Opinion in Neurobiology*, 2004, **14**(6):675-684.

[2] Machens, Christian.K , Romo. Flexible Control of Mutual Inhibition: A Neural Model of Two-Interval Discrimination.[J]. *Science*, 2005.

[3] Aksay.E , Baker.R. Anatomy and discharge properties of pre-motor neurons in the goldfish medulla that have eye-position signals during fixations.[J]. *Journal of Neurophysiology*, 2000, 84(2):1035-1049.

[4] Wang X.J Probabilistic Decision Making by Slow Reverberation in Cortical Circuits[J]. *Neuron*,2002.

[5] Goldman, Compte A , Wang X J . Neural Integrator Models[J]. *Encyclopedia of Neuroscience*, 2015:165-178.

[6] Durstewitz D . Self-organizing neural integrator predicts interval times through climbing activity.[J]. *Journal of Neuroscience*, 2003, 23(12):5342-5353.

[7] Lim , Goldman M.S . Balanced cortical microcircuitry for maintaining information in working memory[J]. *Nature Neuroscience*, 2013, 16(9):1306-U196.

[8] Haider.B , D.A.Mccormick. Rapid neocortical dynamics: cellular and network mechanisms.[J]. *Neuron*, 2009, 62(2):171-189.

[9] D. Green, J. Swets, *Signal Detection Theory and Psychophysics* (J. Wiley, New York, 1966)

[10] Romo.R , A.Hernández, Zainos.A . Neuronal correlates of a perceptual decision in ventral premotor cortex.[J]. *Neuron*, 2004, 41(1):165-173.

[11] Fung.C , Wong.K , Wu.S . A Moving Bump in a Continuous Manifold: A Comprehensive Study of the Tracking Dynamics of Continuous Attractor Neural Networks[J].*Neural computation*, 2008.