Parameter as a Switch Between Dynamical States of a Network in Population Decoding

Jiali Yu, Hua Mao, and Zhang Yi

Abstract—**Population coding is a method to represent stimuli using the collective activities of a number of neurons. Nevertheless, it is difficult to extract information from these population codes with the noise inherent in neuronal responses. Moreover, it is a challenge to identify the right parameter of the decoding model, which plays a key role for convergence. To address the problem, a population decoding model is proposed for parameter selection. Our method successfully identified the key conditions for a nonzero continuous attractor. Both the theoretical analysis and the application studies demonstrate the correctness and effectiveness of this strategy.**

Index Terms—**Continuous attractor, parameter, parameter switch, population decoding.**

I. INTRODUCTION

Population coding is a common scheme used by neural systems for information retrieval [1]–[3]. In the population coding, each neuron has a distribution of responses over a set of inputs, and the responses of many neurons may be combined to determine the input. The population coding is widely applied in the sensor and motor areas of the brain [4], [5]. For instance, in the visual cortical areas, cells are tuned to the moving direction [6]–[8]. Population coding can reduce the uncertainty due to neuronal variability over single-cell coding. Population codes have a lot of computationally desirable properties, such as mechanisms for noise removal, short-term memory, and the instantiation of complex and nonlinear functions [9]. A population code is useful to increase the organism's certainty about the specified feature and to encode the multiple features simultaneously [10].

In [4] and [11], a biological recurrent neural network is adopted for population decoding implementation. This neural network is composed of a population of neurons with bell-shaped tuning curves, and the transfer function of each neuron is nonlinear. The connection weight between two neighboring neurons is symmetric [9]. Wu *et al.* [4] used this network to study the effect of correlation on the Fisher information and compared the performance of three population decoding methods. Deneve *et al.* [11] presented that this class of network could be tuned to implement a maximum likelihood estimator, as long as the network admits smooth hills centered on any point on the neuronal array as the stable activity states. Networks with this property are called continuous attractor networks. Distinguished from separated discrete attractors, their attractors are continuous in this space of activities [12]. A continuous attractor of network is a set of connected stable equilibrium points [13]. Continuous attractor neural network has been used to model short-term memory circuits [14] for continuous stimuli, such as the eye position [15], [16], head direction [17], moving direction [18], [19], path integrator [20]–[22], and cognitive map [23].

The input activity is not always single-peaked, where single-peaked activity means the unidirectional motion. The neural network may receive some stimulations from more than one unknown direction [24]. Most existing research works have focused on decoding the single-peaked input activity. For the single-peaked input activity, the peak position of the continuous attractor of the model is the optimal moving direction [8], [25]. Nevertheless, for the multiple peaked input activity, the output is not multiple peaked any more. In this brief, a model is proposed to read not only the single directional motion but also the multidirectional motion. With multiple peaked input activity, the output of the model is single-peaked, and the peak position is the optimal moving direction.

The major contribution of this brief is the discovery that the parameter configuration determinates the type of continuous attractor contained by a recurrent neural network. An inappropriate parameter may lead to zero continuous attractor. Our model successfully identifies a set of conditions for the parameter selection to achieve a nonzero continuous attractor, which is crucial for decoding.

This brief is organized as follows. Preliminaries are given in Section II. Section III introduces the proposed population model. Continuous attractors of the model are investigated in Section IV. The simulations are given in Section V to illustrate the theory. Section VI gives the application result in the population decoding. Finally, the conclusions are given in Section VII.

II. PRELIMINARIES

The model of general recurrent neural networks can be described by

$$\frac{dx(t)}{dt} = f(x(t)) \tag{1}$$

for $t \geq 0$, where $x \in R^n$, $f : R^n \to R^n$ is some continuous mapping which is local Lipschitz such that given any initial point $x(0) \in R^n$, the trajectory starting from $x(0)$ exists for all $t \geq 0$ and is unique.

Definition 1: A vector $x^* \in R^n$ is called an equilibrium point of (2) if it satisfies $f(x^*) = 0$.

Definition 2: An equilibrium point x^* is said to be stable if given any constant $\epsilon > 0$, there exists a constant $\delta > 0$, such that

$$\|x(0) - x^*\| \leq \delta$$

implies that

$$\|x(t) - x^*\| \leq \epsilon$$

for all $t \geq 0$.

Definition 3: An equilibrium point x^* is said to be asymptotically stable if there exists $\delta > 0$

$$\|x(0) - x^*\| \leq \delta$$

implies that

$$\lim_{t \to +\infty} x(t) = x^*$$

for all $t \geq 0$.

J. Yu is with the School of Mathematics Science, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: yujiali@uestc.edu.cn).

H. Mao and Z. Yi are with the College of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: huamao@scu.edu.cn; zhangyi@scu.edu.cn).

In general, a recurrent neural network may have more than one equilibrium points. The equilibrium points may be isolated or connected. The continuous attractors are related to a set of connected equilibrium points.

Definition 4: A set of equilibrium points CA is called a continuous attractor if it is connected set and each point $x^* \in CA$ is stable.

III. Model

In this brief, we describe a recurrent neural network model with infinite neurons. $a \in R$ is used as an index for the individual neuron, $x(a,t)$ denotes the response of neuron a at time t and is considered as the mean rate at which neuron a produces action potentials or spikes. A neighboring neuron a' can either directly drive neuron a to fire through excitatory synaptic connection $w(a,a')$ or decrease its gain through an inhibitory synaptic connection of strength μ. $w(a,a') > 0$ and $\mu > 0$. μ is the parameter of the network. As shown in Section IV, μ plays a key role to generate a nonzero continuous attractor. For simplicity, all gain interactions have equal strength. The state $x(a,t)$ is governed by

$$\frac{dx(a,t)}{dt} = -x(a,t) + \frac{\int w(a,a')x^2(a',t)da'}{1+\mu\int x^2(a',t)da'} \tag{2}$$

for $t \geq 0$.

The corresponding model with finite neurons can be described by

$$\frac{dx_i(t)}{dt} = -x_i(t) + \frac{\sum_{j=1}^{N} w_{ij}x_j^2(t)}{1+\mu\sum_{j=1}^{N} x_j^2(t)} \tag{3}$$

where $t \geq 0$, $i = 1, \cdots, N$, and N is the total number of neurons. Here, N is a limited number but not an infinite number, so the model (3) involves vectors and matrices rather than continuous functions.

The model (2) can be looked upon as the model of (3) with an infinitely large number of neurons, so that sums over neurons can be replaced by integrals. However, it should be noted that the dynamics of (2) and (3) are quite different. The state space of (2) is infinite, while the state space of (3) is finite.

Consider that the firing rates in the network of all neurons are equal and refer to this uniform rate as X, and suppose that the total synaptic input to all the neurons is the same

$$W_{\text{tot}} = \sum_{j=1}^{N} w_{ij}$$

we can get that

$$\frac{dX}{dt} = -X + \frac{W_{\text{tot}}X^2}{1+\mu N X^2} \tag{4}$$

for $t \geq 0$,

To find the steady states of (4), set the derivative $(dX/dt) = 0$ and solve X. That is

$$X(\mu N X^2 - W_{\text{tot}}X + 1) = 0$$

so if $\mu \leq W_{\text{tot}}^2/4N$, there are three equilibrium points

$$\begin{aligned} X_1 &= 0 \\ X_2 &= \frac{W_{\text{tot}} - \sqrt{W_{\text{tot}}^2 - 4\mu N}}{2\mu N} \\ X_3 &= \frac{W_{\text{tot}} + \sqrt{W_{\text{tot}}^2 - 4\mu N}}{2\mu N}. \end{aligned}$$

Instead, if $\mu > W_{\text{tot}}^2/4N$, there is only one equilibrium point $X_1 = 0$.

Next, let us see whether the three equilibrium points are stable. For a stable equilibrium point, the uniform rate must tend to return to it after a small perturbation. For the system to be stable, each of the deviations must decrease with time. Therefore, we compute

$$\frac{d^2X}{dt^2} = -1 + \frac{2W_{\text{tot}}X}{(1+\mu N X^2)^2}$$

because

$$\begin{aligned} \frac{d^2X}{dt^2}(X_1) = -1 &< 0 \\ \frac{d^2X}{dt^2}(X_2) &> 0 \\ \frac{d^2X}{dt^2}(X_3) &< 0 \end{aligned}$$

then $X_1 = 0$ is stable, X_2 is unstable, and X_3 is stable.

We can see that the network (4) always has a zero stable state $X_1 = 0$ and parameter μ decides the type of the stable steady state. When $\mu \leq W_{\text{tot}}^2/4N$, network (4) possesses another nonzero stable point X_3, while if $\mu > W_{\text{tot}}^2/4N$, network (4) only possesses one zero stable steady state but no nonzero stable point.

In Section IV, we study the stability of the full model (2), not only regarding (4).

IV. Continuous Attractors

This section studies the continuous attractors of the model (2). What happens when all rates are near the equilibrium point but are perturbed by independent amount? We use linearization techniques to check it.

Suppose the weight between the neuron a and a' has the following form:

$$w(a,a') = w(a-a') = W \cdot \exp\left(-\frac{(a-a')^2}{2d^2}\right)$$

where W and d are both positive constants.

The initial state of network has Gaussian shape as

$$x(a,0) = X(0) \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right) \tag{5}$$

where $X(0)$ is a nonnegative constant, and s is a free value, which is also the network representation of stimulus. The peak of activity in this neural system can be localized at any value of s within a range and this value is set by the initial conditions.

Suppose each trajectory of network (2) can be represented as

$$x(a,t) = X(t) \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right) \tag{6}$$

for all $t \geq 0$, where $X(t)$ is some differentiable function.

Theorem 1: The network (2) always possesses the zero stable state

$$O = \{0 | a, s \in R\}.$$

Proof: O is a connected point set and it makes the right-hand side of (2) be zero, so O is an equilibrium.

From (6), we can see that $X(t)$ satisfies

$$\dot{X}(t) = -X(t) + \frac{\sqrt{\pi}dWX^2(t)}{1+\sqrt{2\pi}d\mu X^2(t)} \tag{7}$$

for all $t \geq 0$.

The linearization of (7) at $X(t) = 0$ is given by

$$\frac{d[X(t)-0]}{dt} = -[X(t)-0].$$

Then

$$X(t) = X(0) \cdot e^{-t}$$

thus

$$X(t) \to 0$$

as $t \to +\infty$. Then, there exists $\delta > 0$ if

$$|X(0) - 0| \le \delta$$

then

$$|x(a, 0) - 0| \le \delta$$

we have

$$\lim_{t\to+\infty} x(a, t) = \lim_{t\to+\infty} X(t) \exp\left(-\frac{(a-s)^2}{4d^2}\right) = 0.$$

Therefore, the trajectories converge to O asymptotically, that is to say O is asymptotically stable. The proof is complete. ∎

Theorem 2: When parameter μ satisfies that

$$0 \le \mu < \frac{\sqrt{\pi}}{4\sqrt{2}} dW^2$$

besides the zero continuous attractor, network (2) also possesses a nonzero continuous attractor

$$CA = \left\{ X^* \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right) \middle| a, s \in R \right\}$$

where

$$X^* = \frac{\sqrt{\pi} dW + \sqrt{\pi d^2 W^2 - 4\sqrt{2\pi} d\mu}}{2\sqrt{2\pi} d\mu}.$$

Proof: Clearly, CA is a connected set.

Given each $s \in R$, for $\{x^*(a, s) | a \in R\} \in CA$, $x^*(a, s)$ can be written as

$$x^*(a, s) = X^* \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right).$$

Substitute $x^*(a, s)$ into the right-hand side of (2), it follows that:

$$\begin{aligned}
-x^*(a, s) &+ \frac{\int w(a, a') x^{*2}(a', s) da'}{1 + \mu \int x^{*2}(a', s) da'} \\
&= -X^* \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right) + W X^{*2}(t) \\
&\quad \cdot \frac{\int \exp\left(-\frac{(a-a')^2 + (a'-s)^2}{2d^2}\right) da'}{1 + \mu X^{*2}(t) \cdot \int \exp\left(-\frac{(a'-s)^2}{4d^2}\right) da'} \\
&= \left(-X^* + \frac{\sqrt{\pi} dW X^{*2}}{1 + \sqrt{2\pi} d\mu X^{*2}}\right) \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right).
\end{aligned}$$

Because

$$-X^* + \frac{\sqrt{\pi} dW X^{*2}}{1 + \sqrt{2\pi} d\mu X^{*2}} = 0. \tag{8}$$

Then

$$-x^*(a, s) + \frac{\int w(a, a') x^{*2}(a', s) da'}{1 + \mu \int x^{*2}(a', s) da'} = 0$$

for $a \in R$. By Definition 1, $\{x^*(a, s) | a \in R\}$ is an equilibrium.

Next, we prove that each equilibrium of CA is stable.

The linearization of (7) at $X(t) = X^*$ gives

$$\frac{d[X(t) - X^*]}{dt} = \left(-1 + \frac{2\sqrt{\pi} dW X^*}{(1 + \sqrt{2\pi} d\mu X^{*2})^2}\right) \cdot [X(t) - X^*]$$

then

$$X(t) - X^* = \exp\left(\left(-1 + \frac{2\sqrt{\pi} dW X^*}{(1 + \sqrt{2\pi} d\mu X^{*2})^2}\right) t\right) \cdot \left(X(0) - X^*\right).$$

It follows from (8) that:

$$-1 + \frac{2\sqrt{\pi} dW X^*}{(1 + \sqrt{2\pi} d\mu X^{*2})^2} < 0$$

then

$$\lim_{t\to+\infty} X(t) = X^*.$$

Thus, there exists δ for $a, s \in R$, and suppose

$$|X(0) - x^*| \le \delta$$

then

$$|x(a, 0) - x^*(a, s)| \le \delta$$

we have

$$\begin{aligned}
\lim_{t\to+\infty} x(a, t) &= \lim_{t\to+\infty} X(t) \exp\left(-\frac{(a-s)^2}{4d^2}\right) \\
&= X^* \exp\left(-\frac{(a-s)^2}{4d^2}\right) \\
&= x^*(a, s).
\end{aligned}$$

Therefore, the nonzero state $\{x^*(a, s) | a \in R\}$ is asymptotically stable. Thus, CA is a nonzero continuous attractor. The proof is complete. ∎

V. Simulations

In this section, simulations are carried out to further illustrate the theory established in Section IV.

Example 1:

$$\frac{dx(a, t)}{dt} = -x(a, t) + \frac{2\int \exp\left(-\frac{(a-a')^2}{2}\right) \cdot x^2(a', t) da'}{1 + 0.5 \int x^2(a', t) da'} \tag{9}$$

for $t \ge 0$.

Denote $W = 2$, $d = 1$, and $\mu = 0.5$. It can be checked that $0 \le 0.5 \le \sqrt{\pi/2}$. By Theorem 2, the network has a zero continuous attractor O and nonzero continuous attractors

$$CA = \left\{ \frac{2\sqrt{\pi} + \sqrt{4\pi - 2\sqrt{2\pi}}}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(a-s)^2}{4}\right) \middle| a, s \in R \right\}.$$

The two continuous attractors are plotted in Fig. 1. The red low plane is the zero continuous attractor, while the blue higher one is the nonzero continuous attractor CA, which has Gaussian shape.

To illustrate the stability of CA, suppose that the trajectory of the network (9) has the form

$$x(a, t) = X(t) \cdot \exp\left(-\frac{(a-s)^2}{4}\right) \quad a \in R \tag{10}$$

for $t \ge 0$. Substituting (10) into (9), it follows that $X(t)$ must satisfy the following equation:

$$\frac{dX(t)}{dt} = -X(t) + \frac{2\sqrt{\pi} X^2(t)}{1 + 0.5\sqrt{2\pi} X^2(t)} \tag{11}$$

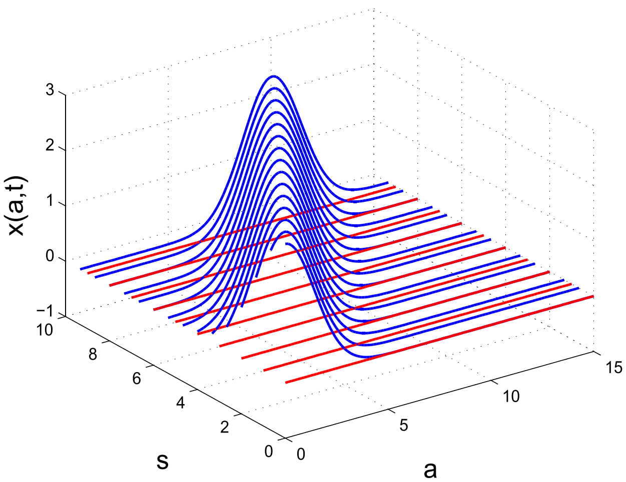

Fig. 1. Continuous attractors of (9) for ten values of s. The red low plane is the zero continuous attractor, while the blue higher surface is the nonzero continuous attractor, which has Gaussian shape.

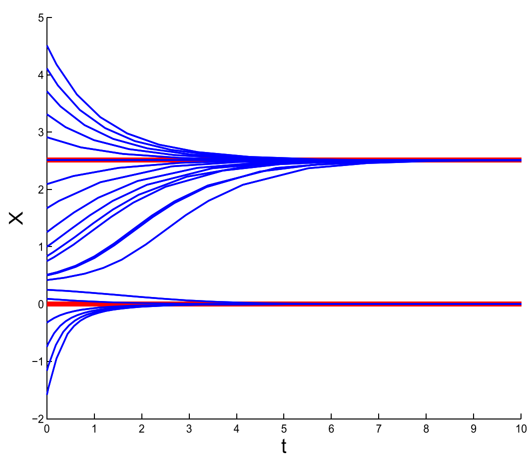

Fig. 2. Stability of (11). There are two stable equilibrium points: $X_1^* = 0$ and $X_2^* = 2.5106$.

for $t \geq 0$. This 1-D equation has two stable equilibrium points: 1) $X_1^* = 0$ and 2) $X_2^* = 2.5106$. Fig. 2 shows the stability of these two equilibrium points of (11).

It clearly indicates that if $x(a, 0)$ is close to CA, then the continuous attractor CA attracts trajectories $x(a, t)$ asymptotically.

Example 2:

$$\frac{dx(a,t)}{dt} = -x(a,t) + \frac{2\int \exp\left(-\frac{(a-a')^2}{2}\right) \cdot x^2(a',t)da'}{1 + 100\int x^2(a',t)da'} \tag{12}$$

for $t \geq 0$.

Here, $\mu = 100$ and $100 > \sqrt{\pi/2}$. The network only has one zero continuous attractor, which is plotted in Fig. 3.

To illustrate the stability of the zero attractor, suppose that the trajectory of the network (12) has the form of

$$x(a,t) = X(t) \cdot \exp\left(-\frac{(a-s)^2}{4}\right) \quad a \in R \tag{13}$$

for $t \geq 0$. Substituting (13) into (12), it follows that $X(t)$ must satisfy the following equation:

$$\frac{dX(t)}{dt} = -X(t) + \frac{2\sqrt{\pi}X^2(t)}{1 + 100\sqrt{2\pi}X^2(t)} \tag{14}$$

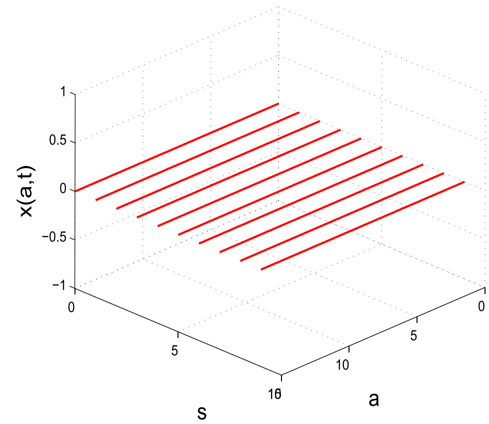

Fig. 3. Zero continuous attractors of (12).

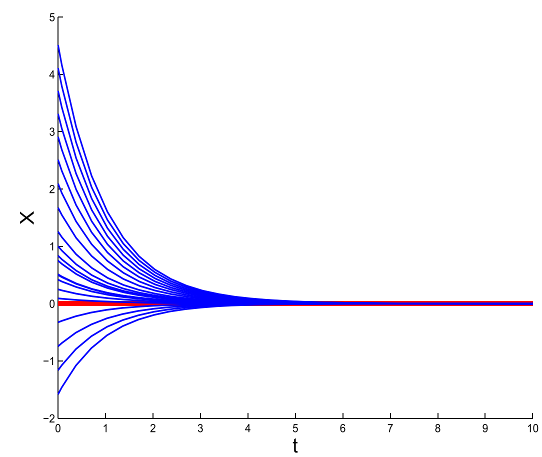

Fig. 4. Stability of (14). All points are asymptotically attracted to zero point.

for $t \geq 0$. This 1-D equation has only one stable equilibrium point: $X^* = 0$. Fig. 4 shows the stability of the zero point of (14).

No matter the initial state $x(a, 0)$ is close or not to zero continuous attractor O, O attracts all trajectories asymptotically.

VI. Application in Population Decoding

In this section, we use the continuous attractor of the population decoding model (2) to read the input signal.

The neural system may receive some stimulations coming from some unknown directions. In the first case, imagine that we present an object moving in one unknown direction s and get the responses of all neurons [Fig. 5 (open circles)]. The responses of neurons can be seen as the initial activity of the network $x(a, 0) = f(a, s) + n(a)$, where

$$f(a,s) = p \cdot \exp\left(-\frac{(a-s)^2}{4d^2}\right)$$

where $n(a)$ is the noise. The noise can be any non-Poisson noises, such as Gaussian white noise, 0-1 normal noise, Rayleigh noise, and Weibull noise. The population decoding model can also be thought as an optimal nonlinear noise filter, as it essentially removes the noise from the inputs.

The network has $N = 60$ neurons with preferred directions uniformly distributed between 0° and 360°. If the parameter of the network satisfies the condition in Theorem 2, the activity of the network converges over time to a single-peaked nonzero continuous

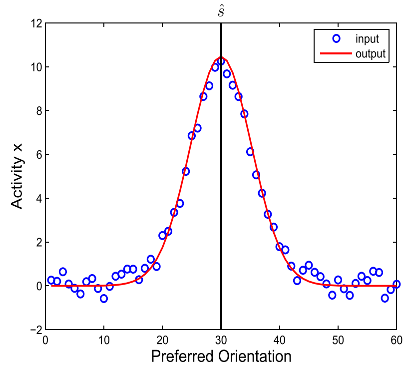

Fig. 5. One peaked population decoding with Gaussian white noise. The open circles are a noisy pattern of activity in response to an object moving at some unknown direction. The activity of 60 cells is plotted according to the preferred direction. The noise was drawn from a random distribution. The red curve is the continuous attractor of the neural network. The continuous attractor fits a smooth hill through the noisy hill. It removes the noise from the noisy hill. The estimate of the moving direction is 180°.

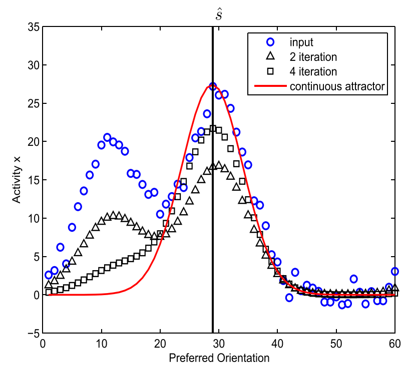

Fig. 6. Reading two-peaked population codes with Gaussian white noise. The continuous attractor (red curve) fits a single-peaked smooth hill through the initial activities. The preferred direction is the peak position.

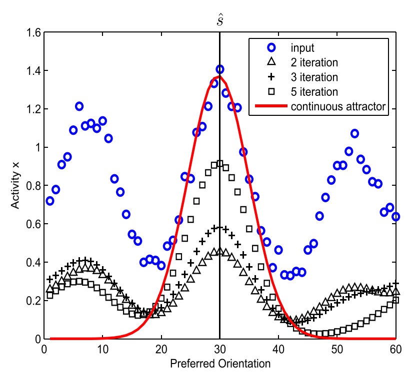

Fig. 7. Three-peaked population decoding with the 0-1 normal noise.

attractor (red curve). The peak position is the estimate of the direction of motion.

Next, we study the case more than one directions of motion are perceived simultaneously at the same location. Then, the activity pattern is multiple peaked population response. No matter how many peaks the input is, as long as we choose an appropriate value of parameter, the output of the population decoding model (2) can converge to the nonzero continuous attractor, which is single-peaked. The peak position is the estimate of the direction of motion. Two-peaked and three-peaked population decoding results are shown in Figs. 6 and 7 separately. The input activities can be seen as the initial state of the network (blue circle curve), the iteration of (2) caused the network activity to relax to a continuous attractor (red curve), which is single-peaked.

If the parameter is not in the range in Theorem 2, the output of the network is always the zero state, which is useless to read the input. Therefore, the value of the parameter is crucial for decoding the population activities.

VII. Conclusion

Population decoding is crucial for the analysis of neural data as well as the understanding of neural computations mechanism. In this brief, we reveal that an inappropriate parameter may lead to zero continuous attractor. To resolve the problem, a model is proposed to read population codes, identify the correct parameter, and guarantee the nonzero continuous attractors. We believe our mechanism represent the standard way for parameter selection. Furthermore, the analysis of this brief may benefit many of neural computing applications and contribute to a better understanding of population decoding in the neural systems.

References

[1] S.-I. Amari and H. Nakahara, "Difficulty of singularity in population coding," *Neural Comput.*, vol. 17, no. 4, pp. 839–858, 2005.

[2] H. S. Seung and H. Sompolinsky, "Simple models for reading neuronal population codes," *Proc. Nat. Acad. Sci. USA*, vol. 90, no. 22, pp. 10749–10753, 1993.

[3] J. Yu, H. Tang, and H. Li, "Dynamics analysis of a population decoding model," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 24, no. 3, pp. 498–503, Mar. 2013.

[4] S. Wu, S.-I. Amari, and H. Nakahara, "Population coding and decoding in a neural field: A computational study," *Neural Comput.*, vol. 14, no. 5, pp. 999–1026, 2002.

[5] S.-I. Amari, "Dynamics of pattern formation in lateral-inhibition type neural fields," *Biol. Cybern.*, vol. 27, no. 2, pp. 77–87, 1977.

[6] J. H. Maunsell and D. C. Van Essen, "Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation," *J. Neurophysiol.*, vol. 49, no. 5, pp. 1127–1147, 1983.

[7] W. M. Usrey and R. C. Reid, "Synchronous activity in the visual system," *Annu. Rev. Physiol.*, vol. 61, pp. 435–456, Mar. 1999.

[8] R. S. Zemel, P. Dayan, and A. Pouget, "Probabilistic interpretation of population codes," *Neural Comput.*, vol. 10, no. 2, pp. 403–430, 1998.

[9] A. Pouget, P. Dayan, and R. S. Zemel, "Information processing with population codes," *Nature Rev. Neurosci.*, vol. 1, no. 2, pp. 125–132, 2000.

[10] W. J. Ma and A. Pouget, "Population codes: Theoretic aspects," in *Encyclopedia of Neuroscience*, vol. 7. New York, NY, USA: Elsevier, Jun. 2009, pp. 749–755.

[11] S. Deneve, P. E. Latham, and A. Pouget, "Reading population codes: A neural implementation of ideal observers," *Nature Neurosci.*, vol. 2, no. 8, pp. 740–745, 1999.

[12] J. Yu, Z. Yi, and L. Zhang, "Representations of continuous attractors of recurrent neural networks," *IEEE Trans. Neural Netw.*, vol. 20, no. 2, pp. 368–372, Feb. 2009.

[13] J. Yu, Z. Yi, and J. Zhou, "Continuous attractors of Lotka–Volterra recurrent neural networks with infinite neurons," *IEEE Trans. Neural Netw.*, vol. 21, no. 10, pp. 1690–1695, Oct. 2010.

[14] G. Mongillo, O. Barak, and M. Tsodyks, "Synaptic theory of working memory," *Science*, vol. 319, no. 5869, pp. 1543–1546, 2008.

[15] H. S. Seung, "Continuous attractors and oculomotor control," *Neural Netw.*, vol. 11, nos. 7–8, pp. 1253–1258, 1998.

[16] H. S. Seung, "How the brain keeps the eyes still," *Proc. Nat. Acad. Sci. USA*, vol. 93, no. 23, pp. 13339–13344, 1996.

[17] K. Zhang, "Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory," *J. Neurosci.*, vol. 16, no. 6, pp. 2112–2126, 1996.

[18] H. S. Seung and D. D. Lee, "The manifold ways of perception," *Science*, vol. 290, no. 5500, pp. 2268–2269, 2000.

[19] S. M. Stringer, E. T. Rolls, T. P. Trappenberg, and I. E. T. de Araujo, "Self-organizing continuous attractor networks and motor function," *Neural Netw.*, vol. 16, no. 2, pp. 161–182, 2003.

[20] D. A. Robinson, "Integrating with neurons," *Annu. Rev. Neurosci.*, vol. 12, pp. 33–45, Mar. 1989.

[21] A. A. Koulakov, S. Raghavachari, A. Kepecs, and J. E. Lisman, "Model for a robust neural integrator," *Nature Neurosci.*, vol. 5, no. 8, pp. 775–782, 2002.

[22] S. M. Stringer, T. P. Trappenberg, E. T. Rolls, and I. E. T. de Araujo, "Self-organizing continuous attractor networks and path integration: One-dimensional models of head direction cells," *Netw., Comput. Neural Syst.*, vol. 13, no. 2, pp. 217–242, 2002.

[23] A. Samsonovich and B. L. McNaughton, "Path integration and cognitive mapping in a continuous attractor neural network model," *J. Neurosci.*, vol. 17, no. 15, pp. 5900–5920, 1997.

[24] S. Treue, K. Hol, and H.-J. Rauber, "Seeing multiple directions of motion—Physiology and psychophysics," *Nature Neurosci.*, vol. 3, no. 3, pp. 270–276, 2000.

[25] A. Pouget, K. Zhang, S. Deneve, and P. E. Latham, "Statistically efficient estimation using population coding," *Neural Comput.*, vol. 10, no. 2, pp. 373–401, 1998.