A Computational Model of mPFC and HPC for Context Structure of Episodic Memory

Zhao Haiteng, Louis Tao

May 31, 2021

Abstract

Hippocampus and mPFC are important neural structures for context processing in episodic memory. The processing is combined with both positive and negative phase differences of theta oscillation between two structures. This paper attempts to give a possible theoretical explanation for the function of mPFC and hippocampus in the context structure problem of episodic memory, as well as the theta oscillation phenomenon, by a computational model. We define the context structure problem as minimum average cut problem of network. We propose a computational model which is based on the the biological structure and relationship of mPFC and hippocampus. Then we show how this model is related to the minimum cut problem, which is a possible theoretical explanation for the relationship between biological structure and its function. We also analyzed the dynamic characteristics of the model, which contains two kinds of oscillations and shows both positive and negative phase differences. This further explained the positive and negative phase differences of theta oscillation of hippocampus and mPFC.

1 Background

1.1 Psychological Theory of Memory and Context

Tulving defines episodic memory as the declarative memory with respect to a particular space over a specified period of time [Tulving, 2004]. An important feature is the connection between concepts. In an environment with abundant stimulation, subject forms connection between a large number of concepts. This memory structure can be considered as a network. However, the structure of episodic memory is not just the connection between projects; it contains more complex multi-level structural and abstract features. For example, experiences and items that are related in time and space tend to belong to the same whole in memory, which can be used abstractly in encoding and retrieval. This concept is often referred to as context. The computational model with context can well solve the structural problems of memory representation and

the feasibility of memory extraction [Howard and Kahana, 2002]. The items in the context are associated with the context, so that the episodic memory is organized, which reduces the complexity of the memory structure and provides a feasible calculation theory for memory retrieval [Raaijmakers and Shiffrin, 1981].

Although the concept of episodic is often involved in psychological research, its definition is not the same in different studies. Some researchers simply think of it as the coding of the surrounding environment which is homogeneous with the coding of the project Dickinson [1983]; some researchers define it as the controlled variables in psychological experiment [Donahoe et al., 2013]; some researchers associate it with cognitive control [Kee et al.]. Some researchers defined context as a combination of stimulus which are spatially and temporally adjacent [Balkenius and Lundagfird, 2000]. Here, based on the definition of space-time proximity, we further propose that the context is the structural feature in a network formed by a large number of stimulus items (such as a network based on spatial and temporal proximity). Theory about complex networks studied these structures in depth. In this study, we use the minimum average cut of the network as a computational criterion of context.

1.2 Neural Basis Related to Context of Episodic Memory

1.2.1 Hippocampus and Medial Prefrontal Cortex

Hippocampus plays a vital role in episodic memory. Damage to the hippocampus can lead to severe impairment of episodic memory [Carr et al., 2011]. The hippocampal indexing theory proposes that the basic mechanism of hippocampus in episodic memory is to sparsely encode the patterns of activity in the neocortex and form links between codes, thereby forming a network of connections between sparse representations in the hippocampus, enabling the storage of episodic memory [Teyler and Rudy, 2010].

Besides the encoding of concrete stimulus, the hippocampus is also involved in global context representation and general idea [Moscovitch et al., 2016]. The mechanism of such functions is not clear yet, but some study proposed that ventral medial prefrontal cortex works together with hippocampus, especially anterior hippocampus cortex, to achieve such functions [Benoit et al., 2014].

A variety of of evidence suggests that the prefrontal lobes play an important role in the formation, control, coding, and extraction of contexts. One of the important theoretical evidence is the relative slow change in neural activity in the prefrontal cortex [Postle, 2006], consistent with the fact that the context structure is a high-level feature. Different parts of the prefrontal lobes have different roles in context-related activities. Unlike the dorsolateral prefrontal lobes which is involved in information maintenance [Braver et al., 2001], the medial prefrontal cortex (mPFC) is considered more involved in the control of context [Eichenbaum, 2017]. But the specific mechanism of the control is

left unknown. Our study attempt to provide a theoretical explanation for this control mechanism, which is consistent with the physiological structure of the medial prefrontal cortex and hippocampus, and the theta oscillatory between them.

1.2.2 Physiological Structure of Hippocampus and mPFC

There are three projection pathways between the hippocampus and the medial prefrontal lobe. The first is a unidirectional junction from the hippocampal CA1 to the medial prefrontal cortex directly; the second is through thalamic nucleus reuniens (Re); the third is through other cortical areas, which form general connection indirectly.

Among the three pathways, the pathway through the thalamic Re is important for the control of context-related memory extraction, and is also an important pathway for theta synchronization between the mPFC and the hippocampus [Eichenbaum, 2017]. The destruction of this pathway can lead to the interference of the learned context [Wei and Südhof] when the subject learns in new context. This shows its important role of control in the process of context structure of episodic memory.

The neural activity of Re contains only little amount of information and does not represent the specific context [Ito et al., 2015]. It is considered to be a hub for excitatory or inhibitory regulation.

1.2.3 Theta Oscillation Synchronization between Hippocampus and mPFC

A large number of studies have found that encoding and retrieval of episodic memory involving context is related to theta (4-12 Hz) oscillation synchronization between hippocampus and mPFC (Eichenbaum, H., 2017).

Early studies found that theta oscillation in the hippocampus was shifted 30 ms earlier than mPFC [Siapas et al., 2005], but subsequent studies have found that the phase of the mPFC is shifted earlier for successful retrieval [Hallock et al., 2016]. Inconsistent studies means that there may be complex relationship between mPFC and hippocampus. Recent studies found both kinds of phase differences at the same time [Place et al., 2016] [Hallock et al., 2016]. These studies attempt to distinguish two kinds of phase differences by different cognitive phases, but the results of different studies are not consistent. For example, [Place et al., 2016] found that when the subject enters the environment, the hippocampus theta oscillation precedes mPFC; but in the task stage, the phase of the mPFC precedes hippocampus. However, studies such as [Torfi et al., 2010] get opposite result. How to explain the coexistence of the two kinds of phase differences is still unknown.

2 Related Theory

This section is organized as follows: in section 2.1, we briefly explain the relationship between Hopfield network and the minimal cut problem.In section 2.2, we will briefly describe the conversion from minimum average cut problem to the variable weight minimum cut problem.

2.1 Hopfield Network and Network Minimum Cut

Hopfield network is a nonlinear dynamical model with dense network structure similar to biological nervous system. The energy function of Hopfield network is Lyapunov function of the dynamical system when certain conditions hold, which guarantees the convergency. Thus many problems can be solved by the Hopfield network if they can be converted to the energy function of a Hopfield network.

[Bruck and Sanz, 2010] studied the relationship between Hopfield network energy function and the network minimum cut problem. They proved that the energy function of the discrete symmetric Hopfield network with -1/1 state corresponds to the value of cut of a undirected network with the same weight of the Hopfield network. They further proved that when the bias and weight of a discrete symmetric Hopfield network with -1/1 state are certain function of weight of a directed network (which is referred to as the original network), the energy function of the Hopfield network corresponds to the value of cut of the original network. The theories above connected Hopfield network and minimum cut problem.

We extended the results above to biased Hopfield network with 0/1 state.

Definition 1 Let G' denote a directed network, which is referred to as original network. Let W' be the weight matrix of G', and note the weight from node j to node i by w'_{ij}. We assume that the network is non-self-connected, which means that $w'_{ii} = 0$. Let G denote a discrete Hopfield network. Denote the state of neuron i by $g(u_i) \in \{0, 1\}$, and denote the weight matrix by W, while the weight from neuron j to neuron i is denoted by w_{ij}. When not confusing, a single subscript will also be used to distinguish different weights, i.e. w'_i or w_i.

Theorem 1 Let neurons of G is the image of identical mapping of nodes of G', and the weight of G satisfies $w_{ij} = \frac{1}{2}(w'_{ij} + w'_{ji})$, and the bias satisfies $I_j = -\frac{1}{2}(\sum_i w'_{ij})$. Then the energy function E of G satisfies $E = -\frac{1}{2}W'^{+-}$, where W'^{+-} is defined as $W'^{+-} = \sum_{g(u_i)=1, g(u_j)=0} w'_{ji}$. Note that W'^{+-} is the cut from the nodes whose corresponding neuron's state is 1 to the nodes of 0.

Proof

$$
\begin{aligned}
E &= \frac{1}{2}(\sum_{i,j} g(u_i) w_{ij} g(u_j)) + \sum_i g(u_i) I_i \\
&= \frac{1}{2}(\sum_{i,j} g(u_i) \frac{1}{2}(w'_{ij} + w'_{ji}) g(u_j)) - \frac{1}{2} \sum_i g(u_i) \sum_j w'_{ij} \\
&= \frac{1}{2}(\sum_{g(u_j)=1, g(u_i)=1} w'_{ji} - \sum_j \sum_{g(u_i)=1} w'_{ji}) \\
&= -\frac{1}{2} \sum_{g(u_j)=0} \sum_{g(u_i)=1} w'_{ji} \qquad (1)
\end{aligned}
$$

We further extend this result to continuous Hopfield network with the same weight mapping whose $g(u_i) \in (0, 1)$ is sigmoid function. This extension is based on the discussion of [Hopfield, 1984]. They argued that for a sigmoid function $g(u)$ with sufficiently large gain, the continues Hopfield network nearly has the same characterization and properties as a discrete Hopfield network. In fact, when the the gain of g(u) become large enough, the stable state will be close to 0 or 1, and the integral term $\int_0^{u_i} g^{-1}(v)dv$ can be ignored so that the energy function will be close to the directed Hopfeild network.

Figure 1 shows the min cut calculated by continuous Hopfield network, which is very close to the ground truth.

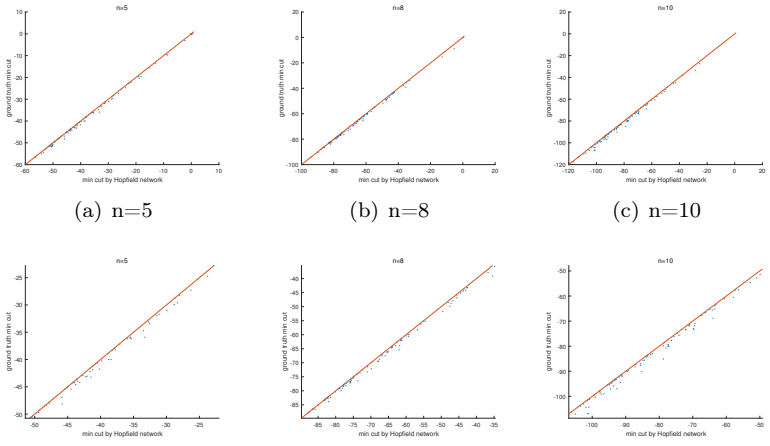
2.2 The Minimum Average Cut and the Varied Weight Minimum Cut

Definition 2 Define the minimum average cut of the network G_1 as $H(G_1) = \min_{c \in C} \frac{1}{|c|}(\sum_{i \in c} w_i)$, where c is the subscript set of a cut of G_1, C is the set of c of G_1, and $||$ indicates the number of elements of the set. We also denote $cH(G_1) = \arg\min_{c \in C} \frac{1}{|c|}(\sum_{i \in c} w_i)$.

This problem attempts to find the cut with the smallest average weight on the network.

Theorem 2 Define the varied weight minimum cut of the network G_1 as $F(\lambda, G_1) = \min_{c \in C} \sum_{i \in c}(w_i - \lambda)$, and denote $cF(\lambda, G_1) = \arg\min_{c \in C} \sum_{i \in c}(w_i - \lambda)$. Then $F(\lambda, G1)$ is strict decreasing function of λ, and $F(\lambda, G_1) = 0$ if and only if $\lambda = H(G_1)$, and these are also equal to $cF(\lambda, G_1) = cH(G_1)$.

Proof This is a direct application of Dinkelbach's Theorem [Schaible, 1976].



(a) n=5 (b) n=8 (c) n=10

(d) n=5,part (e) n=8,part (f) n=10,part

Figure 1: The minimum cut calculated by continuous Hopfield network and ground truth, which is very close to the diagonal line. Each point is the min cut corresponding to an original network with random weights. n is the number of node of the network. The second row are the local of the first row. Note that the min cut calculated by continuous Hopfield network won't surpass the ground truth.

Therefore, we convert minimum average cut problem into the common minimum cut problem by introduces the parameter λ, and give the necessary and sufficient conditions for the equivalence of solutions of two problems. Therefore, if $F(\lambda, G_1)$ and $cF(\lambda, G_1)$ can be solved easily, $H(G_1)$ and $cH(G_1)$ can be solved by searching the space of λ.

Combining Sections 2.1 and 2.2, it's easy to see the feasibility of solving the minimum average cut problem by continuous Hopfield network. In the next section, we will use this association to propose a model of hippocampus and mPFC in context problem.

3 Model

In Section 3.1, we define the problem of finding the context structure as the minimum average cut problem and explain the rationality of this definition. In Section 3.2, we propose our model from the perspective of biological structure and explain its computational ability of context structure problem.

3.1 Context Structure and Minimum Average Cut

Context structure can be considered as substructure problems of network of items and their relationship. When the edges of the network represent similarity, substructure problem seeks for subnetworks that have the smallest similarity, which reflects the inherent structure of the network. The network cut is widely used in structural partitioning problems, such as image segmentation [Wu and Leahy, 1993], because it reflects the object well.

However, as [Shi and Malik, 2000] pointed out, the minimum cut objective function prefers substructures that are too small or too large, because such segmentation usually can reduce the number of edges in the cut, and thus can more easily get smaller cuts. Therefore, a series of intuitive revision methods have been proposed [Soundararajan and Sarkar, 2003]. These revisions use features such as the number of nodes in subgraph, or the connectivity between subgraph and full network to normalize the cut, which can avoid the bias of minimum cuts. However, these graph segmentation methods are mostly used for grid networks [Özgür Şimşek et al., 2005], rather than complete graph. The number of edges in cut of complete graph grows faster than cut of grid network, so revisions such as the number of subgraph nodes perform not well.

So we propose that average cut can be a suitable objective function for finding substructures. Formally, we define the context structural partitioning problem as the minimum average cut problem, which can be written as $cH(G')$, where G' denotes the relationship network of items.

3.2 Model with Biological Basis

We will present our model from the perspective of biological structure first. As shown in Figure 2, our model contains the hippocampus and the mPFC controlling signal through the Re of the thalamus. The hippocampus model contains excitatory and inhibitory neurons. Due to the dense interconnection of hippocampal, we consider it as a fully connected network [Dubreuil and Brunel, 2013], and we use the ordinary differential equations with the form of Hopfield network to describe the excitatory neurons [Natschläger, 1997] [Hopfield, 1984].

Let u_i denote the average membrane potential of the neuron i. $g(u) = \frac{1}{1+e^{-u}}$ is the sigmoid function, which describes the nonlinear relationship between average membrane potential and average action potential. Thus the input received from other excitatory neurons by neuron i is $\sum_{j \neq i} w_{ij} g(u_j)$. We introduce linear decay current $-a_i u_i$, and constant bias current I_i for each neuron. We also introduce a self-feedback term $e_i g(u_i)$ for each neuron i , where e_i is the coefficient.

We consider all inhibitory neurons as a population and assume that they have homogenous connection with all excitatory neurons, which functions as global

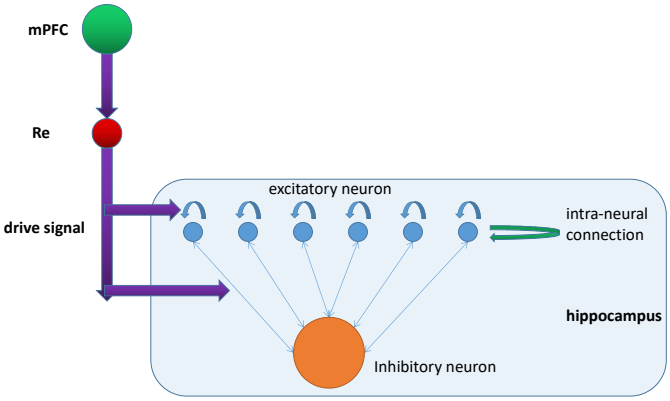


Figure 2: The Architecture of our model. The hippocampus consist of excitatory neural network and a global inhibitory neural population. the control signal of mPFC acts on hippocampus through Re, which drives excitatory neurons, and amplifies the self-feedback current and inhibitory current.

normalization for all the excitatory neurons (Filip, P., Hopfield, JJ. (2013)). We denote the weight from excitatory neurons to inhibitory neurons population by w_{ei}, and denote the weight from inhibitory neurons to excitatory neurons by w_{ie}. So they receive the input $w_{ei}\sum_j g(u_j)$ from global excitatory neurons. We ignore the dynamics of inhibitory neurons and assume that the output is linearly related to the input, i.e. $e_t w_{ei}\sum_j g(u_j)$, where e_t is the coefficient. Thus the inhibitory input to the excitatory neurons is $w_{ie}e_t w_{ei}\sum_j g(u_j)$. We use coefficient e to note $w_{ie}e_t w_{ei}$, so this term is $e\sum_j g(u_j)$.

Here, we emphasize the rationality of the linear function between the input and output of inhibitory population, because the activity of population is usually almost linear when input is small. For example, for divisive normalization models $out = R_{max}\frac{in^n}{in^n+\sigma}$ Olsen et al. [2010], we have $out \approx \frac{R_{max}}{\sigma}in$ when $n = 1$ and $in \ll \sigma$.

We consider the control signals from the Re to hippocampus as an drive signal. Let the signal be denoted by λ. Let the drive signal be an input to the excitatory neuron through the weight w_λ. And we assume it be the linear amplifying coefficient to the self-feedback current of the excitatory neurons and the inhibitory current of the inhibitory neurons to the excitatory neurons.

In conclusion, our model can be expressed as

$$\tau \frac{du_i}{dt} = -a_i u_i + \sum_{j \neq i} w_{ij} g(u_j) + \lambda e_i g(u_i) + I_i + w_\lambda \lambda - \lambda e \sum_j g(u_j) \tag{2}$$

where τ is the time constant.

We will next explain the association between this model and the context structure problem. From theorem 2, $cH(G')$ on G' can be done by searching λ satisfying $F(\lambda, G') = 0$. Then $cF(\lambda, G')$ is the solution of $cH(G')$.

Define G'' as a network with the same nodes as G', whose weight matrix is $W' - \lambda E$, where E is an $n \cdot n$ matrix whose diagonal elements are 0 and the remaining elements are 1. Solving $F(\lambda, G')$ is identical to solving the minimum cut problem of G''. According to Theorem 1, this can be achieved by Hopfield network corresponding to G''. For weight matrix $W' - \lambda E$, the weight of the corresponding Hopfield network is

$$w''_{ij} = \frac{1}{2}(w'_{ij} - \lambda + w'_{ji} - \lambda) = w_{ij} - \lambda, i \neq j \tag{3}$$

$$w''_{ii} = 0 \tag{4}$$

$$I''_i = -\frac{1}{2} \sum_{j \neq i} (w'_{ji} - \lambda) = I_i + \lambda \frac{n-1}{2} \tag{5}$$

So the corresponding Hopfield network is

$$\begin{aligned} \tau \frac{du_i}{dt} &= -a_i u_i + \sum_{j \neq i} (w_{ij} - \lambda) g(u_j) + I_i + \lambda \frac{n-1}{2} \\ &= -a_i u_i + \sum_{j \neq i} w_{ij} g(u_j) + \lambda g(u_i) + I_i + \frac{n-1}{2} \lambda - \lambda \sum_j g(u_j) \end{aligned} \tag{6}$$

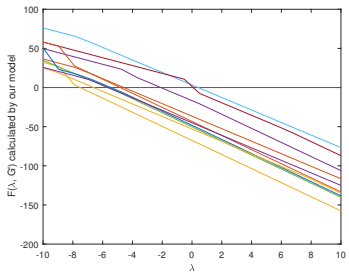
Comparing (2) and (6), it's obvious that when $e_i = e$ and $w_\lambda = e\frac{n-1}{2}$, the model in (2) is the Hopfield network for solving the minimum cut problem on G'', which can be performed by search of λ.

From the above, we modeled the biological structure and showed how it can solve the context structure problem.

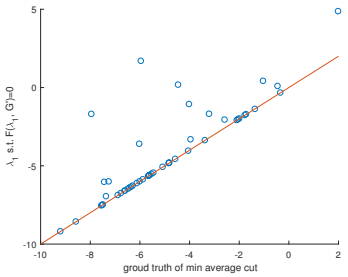
Figure 3 shows the validity of out model. The $F(\lambda, G')$ solved by or model is monotone decreasing function of λ, which follows theorem 2. And the critical $\lambda_1 | F(\lambda_1, G') = 0$ is close to the ground truth $H(G')$.

4 Dynamic Properties of Model

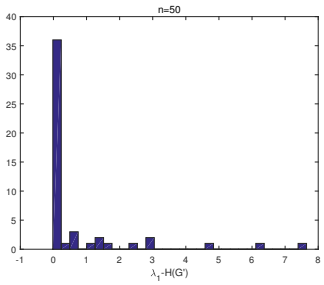
In this section, we will analyze the dynamic properties of the model and explain its relationship with the phase difference of theta oscillation between the



(a) the monotone decreasing function



(b) scatter of model's result and ground truth



(c) the histogram of error

Figure 3: (a) The relationship between $F(\lambda, G')$ calculated by our model and λ. Each line is function of a random original network. (b) Most λ_1 found by our model is close to the ground truth $H(G')$. (c) The histogram of error of $\lambda_1 - H(G')$. The total number of samples is 50.

hippocampus and the mPFC.

We first declare the good convergence properties of our model, which is the fast dynamics of our model. That is, for any given λ, our model will converge to a local minimum of varied weight minimum cut. Note that the Hopfield network can't be guaranteed to converge to the global minimum solution, which is the inner property of the minimum cut itself.

Theorem 3 For any given λ, our model is globally asymptotically stable, which means that, starting from any initial value, when $t \to +\infty$, the network state converges to the set of equilibrium points.

Proof Because the weight of Hopfield network is always symmetric, theorem 3 can be derived from the theorem that a continues Hopfield network with symmetric weight and monotonically increasing active function is globally asymptotically stable [Hopfield, 1984].

With theorem 3, we get guarantee that the network converges with any given λ value.

Because solving the minimum average cut problem requires searching in the possible interval $[\lambda_1, \lambda_2]$, a simple method is to let λ traverse the interval slowly enough. In this situation, the system becomes a fast-slow system. When λ changes in the slow time scale, the Hopfield network is a fast system and can be considered to be approximately at its equilibrium all the time. If we let λ traverse in the form of a piecewise continuous periodic function of time, and the frequency is close to theta frequency (4-8Hz), the assumption of the fast-slow system is reasonable, because the time constant of the Hopfield network is much smaller than theta period. Thus, we consider the theta oscillation of control signal from mPFC to hippocampus as traversing and searching.

In the following content, we will further analyze the slow dynamics with respect to λ, which are the properties of response of the hippocampus to theta oscillation, especially its phase relationship to the drive signal.

The form of λ can take any form of piecewise continuous function. But for the simplicity of analysis, we use a piecewise linear function, $\lambda = \lambda(t) = \max\{0, 2(t - (\frac{n+1}{2})T)\frac{\lambda_{max}}{T}\}$, where T is the period, $n = \max\{n_1 | n_1 \in Z, n_1 T \leq t\}$. And when we consider the interval when λ increases linearly, let's take the form of $\lambda = ct$ for simplicity. Note that the specific form of the function λ does not affect our analysis. For the interval where λ is continuous to time, the results below are always true.

Theorem 4 For $\lambda = ct$ and $g(u_i)$ of all neurons are close to 0 or 1, (1) if $|\{i|g(u_i) \approx 0\}| = |\{i|g(u_i) \approx 1\}| = \frac{n}{2}$ for n is even, and $|\{i|g(u_i) \approx 0\}| = \frac{n-1}{2}$,

$|\{i|g(u_i) \approx 1\}| = \frac{n+1}{2}$ or the inverse for n is odd, then $g(u)$ will be stable, which means that $g(u_i)$ close to 0 will approximate 0 or remain stable, and the same for $g(u_i)$ close to 1; (2) if the condition of (1) does not hold, then all the u_i and $g(u_i)$ will change by the same sign with $\frac{n}{2} - \sum_i g(u_i)$.

Proof Let A denote the diagonal matrix where $a_{ii} = a_i$. With discussion above, let's assume that the system is at equilibrium state u_e at time t, which means

$$-Au_e + (W - Ect)g(u_e) + I + \frac{n-1}{2}ct = 0 \tag{7}$$

After dt, the system is at new equilibrium state $u_e + du$, and we get

$$-A(u_e + du) + (W - Ec(t + dt))g(u_e + du) + I + \frac{n-1}{2}c(t + dt) = 0 \tag{8}$$

Let $J(u)$ denote the Jacobian matrix of $g(u)$ to u, which is a diagonal matrix with $J_{ii}(u) = g'(u_i)$. Then combining (7) and (8) we get

$$[-A + (W - Ect)J(u_e)]du = (Eg(u_e) - \frac{n-1}{2}\mathbf{1})cdt \tag{9}$$

where $\mathbf{1}$ denote the vector all the elements of which are 1. When all the $g(u_i)$ is close to 0 or 1, $g'(u_i) \approx 0$. So

$$du_i = -\frac{c}{a_i}(\sum_{j \neq i} g(u_j) - \frac{n-1}{2})dt \tag{10}$$

In condition (1), let's assume that n is odd first. if $|\{i|g(u_i) \approx 0\}| = |\{i|g(u_i) \approx 1\}| = \frac{n}{2}$, then $\sum_{j \neq i} g(u_j) - \frac{n-1}{2} > 0$ for $i|g(u_i) \approx 1$, and < 0 for $i|g(u_i) \approx 0$. So $du_i < 0$ for $i|g(u_i) \approx 0$, and $du_i > 0$ for $i|g(u_i) \approx 1$. Next, assume that n is even. Without loss of generality, assume that $|\{i|g(u_i) \approx 0\}| = \frac{n-1}{2}$, and $|\{i|g(u_i) \approx 1\}| = \frac{n+1}{2}$. Then $\sum_{j \neq i} g(u_j) - \frac{n-1}{2} = 0$ for $i|g(u_i) \approx 1$, and < 0 for $i|g(u_i) \approx 0$. So $du_i < 0$ for $i|g(u_i) \approx 0$, and $du_i = 0$ for $i|g(u_i) \approx 1$. Above all, $g(u)$ is stable in condition (1).

If condition (1) does not hold, then the sign of $\sum_{j \neq i} g(u_j) - \frac{n-1}{2}$ is the same with $\sum_i g(u_i) - \frac{n}{2}$, so the sign of du_i is the same with $\frac{n}{2} - \sum_i g(u_i)$.

Note that (1) and (2) shows that that state in (1) is like an attractor of the dynamics with respect to λ.

Considering the sparsity of spontaneous activity of biological neural network, we assume that the initial condition g(u) does not meet the condition of (1) of theorem 4, and $|\{i|g(u_i) \approx 0\}| > |\{i|g(u_i) \approx 1\}|$. Thus theorem 4 shows that the neurons in the network are divided into two classes. With the increase of λ,

they show different dynamic properties. Nearly half neurons' $g(u)$ increase at the beginning and increase to nearly 1 finally, and the other half of the neurons' $g(u)$ increase in the initial stage, but in the final stage they decrease to 0 again. We call them class I and class II neurons respectively. We can see that the wave crest of class I is at the end of time interval when λ increase, while the wave crest of class II must appear in the middle of this interval. We further analyze the phase of the wave crest of class II neurons.

To simplify the analysis, we consider the situation when $g'(u_i)$ of only one neuron is significantly not close to 0, and $g'(u) \approx 0$ for the remaining neurons. Note that this is equal to the situation that one neural's $g(u)$ is close to 0.5, while others are still close to 0 or 1. Because if some neurons have changed their $g(u)$ from nearly 0 to nearly 1, the moment when some neurons' $g(u)$ are close to 0.5 must exist, so the assumption only means that we don't consider the situation when more than one neurons' $g(u)$ are close to 0.5. In such situation, the influence between two neuron is difficult to analyze. But we should note that for large sigmoid gain, this condition almost always holds.

Assume that $g'(u_i) \neq 0$, while $g'(u_j) \approx 0$ for $j \neq i$. Then only the i_{th} column of $(W - Ect)J(u_e)$ is not 0. Thus we can get that

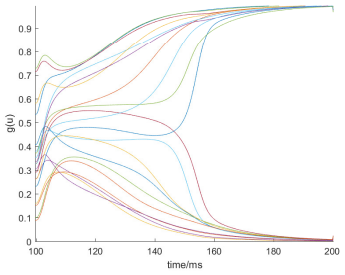
$$du_j = \begin{cases} -cdt(\sum_{k \neq i} g(u_k) - \frac{n-1}{2}) \frac{a_i + (w_{ji} - ct)g'(u_i)}{a_i a_j} & j \neq i \\ -cdt(\sum_{k \neq i} g(u_k) - \frac{n-1}{2}) \frac{1}{a_i} & j = i \end{cases} \tag{11}$$

Therefore, the condition that the neuron j shows a wave crest is $a_i + (w_{ji} - ct)g'(u_i))$ changes from positive to negative, which means $t > \frac{w_{ji} + a_i / g'(u_i)}{c}$.

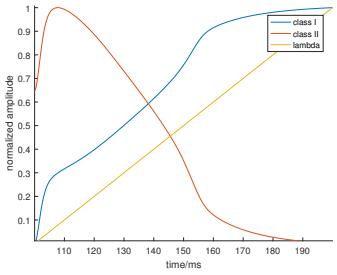
The analysis above only holds locally. Neurons whose $g(u)$ increases from 0 to more than 0.5 may decrease to 0 again, and neurons whose $g'(u)$ is temporarily reversed from positive to negative may reverse again, so the global properties of the dynamic properties of two classes of neurons cannot be predicted by the result simply. However, we can still obtain qualitative relationship between the phase difference between the wave wrest of two classes of neurons, which is negatively related to the weight and the decay coefficient of membrane voltage.

Figure 4 (a) shows the two classes of dynamics of neurons. The same with our analysis, all the $g(u_i)$ increase at the beginning, while half of neurons changes directions and finally be close to 0.

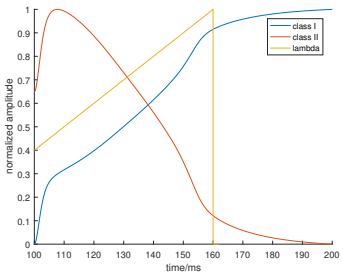
We will next briefly discuss the relationship between this biphasic dynamics and the mPFC-hippocampus phase difference. Since λ is the control signal from mPFC to hippocampus through the thalamic Re, the theta oscillation measured in the mPFC must precede the λ signal received by hippocampus $\delta\phi_1$. Denote



(a) the two classes of dynamics



(b) normalized wave of two classes of neurons and λ



(c) the phase relationship with delay between mPFC and hippocampus

Figure 4: (a)Two classes of dynamics of neurons. (b) Normalized wave of two classes of neurons with respect to λ. (c) The wave pattern where λ corresponds to the signal measured in mPFC, which precede the signal in hippocampus $\delta\phi_1 = 40ms$.

the phase difference of wave wrests of two classes of neurons by $\delta\phi$. Then the phase differences between mPFC and hippocampus will be $\delta\phi_1$ and $\delta\phi - \delta\phi_1$. which are positive and negative, respectively. Thus the property of out model can explain the bidirectional phase relationship between mPFC and hippocampus.

Figure 4 (b) and (c) shows the phase differences without and with delay. Note that with $\delta\phi_1 = 40ms$ [Eichenbaum, 2017], they show quite similar pattern with the theta oscillations.

5 Conclusion

Context structure is an important cognitive task of episodic memory. Hippocampus and mPFC are key neural structures for context processing in episodic memory, while hippocampus represent the items and their relationship and mPFC control the context processing. The control is through Re of thalamus. The processing of context is combined with both positive and negative phase differences of theta oscillation between mPFC and hippocampus. But the mechanism of the function and dynamics is not clear yet.

This paper proposed a theoretical explanation for the function of mPFC and hippocampus in the context structure problem of episodic memory, as well as the theta oscillation phenomenon, by a computational model. We define the context structure problem as mean minimum cut problem of network. We propose a computational model which is based on the the biological structure and relationship of mPFC and hippocampus. Then we show how this model is related to the minimum cut problem, which is a possible theoretical explanation for the relationship between biological structure and its function. We also analyzed the dynamic characteristics of the model, which contains two kinds of oscillations and shows both positive and negative phase differences. This further explained the positive and negative phase differences of theta oscillation of hippocampus and mPFC.

References

Christian Balkenius and Kungshuset Lundagfird. A computational model of context processing. 06 2000.

R. G. Benoit, K. K. Szpunar, and D. L. Schacter. Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *Proceedings of the National Academy of Sciences of the United States of America*, 111(46):16550, 2014.

T S Braver, D M Barch, B A Keys, C S Carter, J D Cohen, J A Kaye, J S Janowsky, S F Taylor, J A Yesavage, and M S Mumenthaler. Context processing in older adults: evidence for a theory relating cognitive control to neurobiology

in healthy aging. *Journal of Experimental Psychology General*, 130(4):746–63, 2001.

Jehoshua Bruck and Jorge Sanz. A study on neural networks. *International Journal of Intelligent Systems*, 3(1):59–75, 2010.

Margaret F Carr, Shantanu P Jadhav, and Loren M Frank. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, 14(2):147–53, 2011.

A Dickinson. *Conditioning and associative learning.* 1983.

J. W. Donahoe, D. C. Palmer, and J. E. Burgos. The s-r issue: its status in behavior analysis and in donahoe and palmer's learning and complex behavior. *Journal of the Experimental Analysis of Behavior*, 67(2):193–211, 2013.

Alexis Dubreuil and Nicolas Brunel. Memory capacity of networks with stochastic binary synapses. *Bmc Neuroscience*, 14(S1):P43, 2013.

Howard Eichenbaum. Prefrontal-hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, 18(9):547, 2017.

H. L. Hallock, A. Wang, and A. L. Griffin. Ventral midline thalamus is critical for hippocampal-prefrontal synchrony and spatial working memory. *Journal of Neuroscience*, 36(32):8372–8389, 2016.

J J Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc Natl Acad Sci U S A*, 81(10): 3088–3092, 1984.

M. W Howard and M. J Kahana. A distributed representation of temporal context. 2002.

H. T. Ito, S. J. Zhang, M. P. Witter, E. I. Moser, and M. B. Moser. A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation. *Nature*, 522 (7554):50, 2015.

Loh Kep Kee, Hadj Bouziane Fadila, Petrides Michael, Procyk Emmanuel, and Amiez Céline. Rostro-caudal organization of connectivity between cingulate motor areas and lateral frontal regions. *Frontiers in Neuroscience*.

Morris Moscovitch, Roberto Cabeza, Gordon Winocur, and Lynn Nadel. Episodic memory and beyond: The hippocampus and neocortex in transformation. *Annual Review of Psychology*, 67(1):105–134, 2016. doi: 10.1146/annurev-psych-113011-143733. URL `https://doi.org/10.1146/annurev-psych-113011-143733`. PMID: 26726963.

Thomas Natschläger. Networks of spiking neurons can emulate arbitrary hopfield nets in temporal coding. *Network Computation in Neural Systems*, 8(4): 355–371, 1997.

Shawn R. Olsen, Vikas Bh, and Rachel I. Wilso. Divisive normalization in olfactory population codes. *Neuron*, pages 287–299, 2010.

R Place, A Farovik, M Brockmann, and H Eichenbaum. Bidirectional prefrontal-hippocampal interactions support context-guided memory. *Nature Neuroscience*, 19(8):992–994, 2016.

B. R. Postle. Working memory as an emergent property of the mind and brain. *Neuroscience*, 139(1):23–38, 2006.

Jeroen G Raaijmakers and Richard M Shiffrin. Search of associative memory. *Psychological Review*, 88(2):93–134, 1981.

Siegfried Schaible. Fractional programming. ii, on dinkelbach's algorithm. *Management Science*, 22(8):868–873, 1976.

Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Trans.pattern Anal.mach.intell*, 22(8):888–905, 2000.

Athanassios G Siapas, Evgueniy V Lubenov, and Matthew A Wilson. Prefrontal phase locking to hippocampal theta oscillations. *Neuron*, 46(1):141–151, 2005.

P. Soundararajan and S. Sarkar. *An in-depth study of graph partitioning measures for perceptual organization.* 2003.

T. J. Teyler and J. W. Rudy. The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus*, 17(12):1158–1169, 2010.

Sigurdsson Torfi, Kimberly L Stark, Karayiorgou Maria, Joseph A Gogos, and Joshua A Gordon. Impaired hippocampal-prefrontal synchrony in a genetic mouse model of schizophrenia. *Nature*, 464(7289):763–7, 2010.

E Tulving. Episodic memory: from mind to brain. *Rev Neurol*, 53(4 Pt 2):S9, 2004.

Xu Wei and Thomas C. Südhof. [research article] a neural circuit for memory specificity and generalization.

Zhenyu Wu and Richard M. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Trans.pattern Anal.machine Intell*, 15(11):1101–1113, 1993.

Özgür Şimşek, Alicia P. Wolfe, and Andrew G. Barto. Identifying useful subgoals in reinforcement learning by local graph partitioning. In *International Conference on Machine Learning*, 2005.