A framework for linking computations and rhythm-based timing patterns in neural firing, such as phase precession in hippocampal place cells

E. Paxon Frady (epaxon@berkeley.edu)

Pentti Kanerva (pkanerva@csli.stanford.edu)

Friedrich T. Sommer (fsommer@berkeley.edu)
Redwood Center, UC Berkeley, 575A Evans Hall, MC 3198
Berkeley, CA 94720 USA

Abstract

A major challenge in neuroscience is to develop models that bridge between observed neural firing patterns and computational functions. Here, we demonstrate the utility of Vector Symbolic Architecture (VSA) models in building a theory framework for neuroscience. Specifically, we present a VSA model expressing computations by operations on high-dimensional vectors of complex numbers, Fourier Holographic Reduced Representations (FHRR). We have developed a novel model of synaptic integration to implement FHRR operations with spiking neurons that express periodic population firing, where the timing of a spike relative to an internal oscillation represents the phase of a complex number. We illustrate how algorithms defined on a computational level, such as associative memory or spatial navigation, can be implemented by spiking neurons that exhibit similar firing patterns as observed in neural recordings. Thus, FHRR VSAs can establish a link between concrete computations and properties of neural firing such as oscillations and phase precession in hippocampus and cortex.

Keywords: spiking neurons, phase-precession,

Introduction

The goal of a theory framework for understanding brain recordings is the inference of principles of computation and brain function from complex, high-dimensional data. One approach has been to postulate a computation, build a "normative" network model that performs exactly this computation, and compare model predictions to experimental observations. This approach has led to the identification of specific types of computations in early sensory processing, e.g. contrast invariant feature detection (Troyer et al., 1998) and efficient coding (Olshausen Field, 1996; Rehn Sommer, 2007). However, most areas of cortex and hippocampus likely subserve multiple computational operations, leading to a combinatorial explosion of normative models.

We aim to develop a theory framework for neural computation that is modular and general enough to describe computations in different brain regions, e.g. hippocampus and visual cortex. To describe putative computations that underlie behavior, we will begin with models that have been proposed in the connectionist literature to understand cognitive functions (Plate, 1995; Kanerva, 1996; Plate, 2003; Gayler, 2004), often referred to as vector symbolic architectures (VSA). In these models, symbols or concepts are represented by pseudo-randomized high-dimensional vectors and the computations underlying cognitive reasoning are modeled by algebraic operations in the vector space (Kanerva, 2009). Different VSA models have been proposed using different vector spaces —binary, real or complex— and different operations. VSA models have in common that they include a few elementary operations that, when appropriately composed, can solve even challenging cognitive problems (Gayler, 2004). Typically, these operations include a binding operation to form data structures, such as key-value pairs, a superposition operation to represent sets of items, and a "scrambling" operation, such as a permutation. VSA models are also able to describe challenging problems of sensory-level processing, such as invariant object recognition (Eliasmith et al., 2012).

The idea that the brain uses pseudo-random activity for indexing structured computations seems far fetched at first. However, it is consistent with observations of apparent randomness in neural recordings, and some theories for hippocampus and cortex (Teylor and DiScenna 1986; Hawkins et al., 2017). Further, our recent work (Frady et al., 2018) has demonstrated how the structured representation of time series data formed in different types of VSA models is naturally implemented by recurrent networks with random weights, akin to those considered in reservoir computing (Lukoševičius Jaeger, 2009). Also, the feasibility of implementing cognitive computations with (a particular) VSA by rate codes in spiking neurons has been shown (Eliasmith et al., 2012).

Here, we use VSA as a flexible and modular two-level theory framework to describe neural computation. The first level uses a generic VSA formalism to define the putative computation underlying a behavior under study. The second level of the framework describes how the computational elements are implemented by spiking activity patterns and the structure of neural circuits. We demonstrate the utility of VSA models in a neuroscience theory framework by developing models for visual processing and memory (visual cortex), and navigation (hippocampus). Specifically, we will describe advances using a VSA model based on the algebra of vectors of complex numbers, Fourier Holographic Reduced Representations (FHRR: Plate, 2003).

In FHRR, base symbols are represented by high-dimensional vectors of complex numbers with unit magnitude. We have developed a novel model of synaptic integration to implement FHRR with spiking integrate-and-fire neurons that express periodic population firing. In this model, the timing of spikes represents the relative phases of complex numbers. This makes FHRR useful for understanding phase coding in hippocampus and visual cortex. We demonstrate how the model of synaptic integration can be used to build a stable spiking complex Hopfield network (Jankowski et al. 1996) that can store arbitrary spiking phase patterns as attractor states, generates internal oscillations with no external clock, and that can recall information as a content-addressable memory with pattern completion. Finally, we show how applying sparsity to the FHRR vectors can produce a representation of two-dimensional space at the computational level, while also matching several observed phenomenon in the hippocampus at the mechanistic level, such as place fields and phase-precession.

Results

A circuit model for memorizing sensory inputs

Here we present a theory and model for spike-timing representations that follows the algebra of Fourier Holographic Reduced Representations (FHRR; Plate (5)). In this framework, symbols are represented by high-dimensional vectors (hypervectors) of complex numbers with random phase and unit magnitude. The transmission of complex numbers from one layer to another, or from time-step to time-step, can be related to the timing of action-potentials relative to an ongoing circuit oscillation or rhythm. This mechanism relates to many observations in the nervous system, for instance, in the hippocampal system it is well known that spiking is synchronized with the theta-rhythm and information can be encoded by the spiking relative to the phase of the oscillation (2).

The neurons in this model express periodic population firing, where the timing of a spike relative to an internal oscillation represents the phase of a complex number. The proposed synaptic dynamics transforms a presynaptic spike into an excitation/inhibition-balanced postsynaptic current oscillation. The timing of presynaptic spike and synaptic delay determines the phase of the postsynaptic current oscillation. The oscillatory currents from many synapses will sum in the neuron, in essence implementing a dot product between complex vectors. The effect of spikes depends now on their relative timing and stored synaptic patterns — it can be small in the decoherent case (Fig. 1A), or large in the coherent case (Fig. 1B).

We demonstrate the theory and mechanisms in a working model of a spiking complex Hopfield attractor network that can store and retrieve arbitrary phase patterns (Fig. 1C) as well as images encoded as FHRR hypervector symbols (Fig. 1D, E). The complex Hopfield network simply extends from the traditional Hopfield network (4). The outer-product learning rule is replaced with the conjugate outer-product learning rule. The neural non-linearity is changed to normalization to unit magnitude for complex numbers (3).

At the computational-level, the model we present is a network of three layers, an encoding layer, the Hopfield layer, and a decoding layer. The encoding layer, $\mathbf{x}(T) \in \mathbb{R}^D$, represents the inputs and is used to initialize the Hopfield network through the encoding matrix $\mathbf{X} \in \mathbb{C}^{N \times D}$, where D is the dimensionality of the input and N is the number of neurons in the Hopfield layer. The Hopfield layer, $\mathbf{h}(T) \in \mathbb{C}^N$, iterates through time with recurrent matrix $\mathbf{W} \in \mathbb{C}^{N \times N}$ and the complex magnitude $|\mathbf{h}_i(T)|$ renormalized to unity each time step by the non-linear function, Θ. The decode layer $\mathbf{y}(T) \in \mathbb{R}^D$ resolves the state of the complex Hopfield network with decode matrix $\mathbf{Y} \in \mathbb{C}^{N \times D}$. This gives a network described by these update equations:

$$\begin{aligned} \mathbf{h}(T) &= \Theta\left(\mathbf{W}\mathbf{h}(T-1) + \mathbf{X}\mathbf{x}(T)\right) \\ \mathbf{y}(T) &= \Re\left(\mathbf{Y}^\top \mathbf{h}(T)\right) \end{aligned} \tag{1}$$

Any phase pattern (vector of complex numbers with unit magnitude) can be stored in the complex Hopfield network using the conjugate outer-product rule, $\mathbf{W} = \mathbf{S}\mathbf{S}^{*\top}$, where $\mathbf{S} \in \mathbb{C}^{N \times M}$ is the matrix of phase patterns or of FHRR vector-symbols, and M is the number of patterns to store. Other arbitrary input patterns, such as RGB images, given by matrix $\mathbf{P} \in \mathbb{R}^{M \times D}$ can be stored in the network by associating each image pattern with a FHRR vector-symbol in the encoding and decoding matrices (1). However, the input patterns to be stored may have correlations. This can be resolved by storing the *decorrelated* patterns in the encoding matrix. Using singular value decomposition, $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$, the encoding and decoding matrices are:

$$\begin{aligned} \mathbf{X} &= \mathbf{S}\mathbf{U}\mathbf{V} \\ \mathbf{Y} &= \mathbf{S}^*\mathbf{P} \end{aligned} \tag{2}$$

At the mechanistic level, these equations are implemented by spiking integrate-and-fire neurons. The excitatory synapses have a tunable magnitude and delay, which are set by the magnitude and phase of the complex numbers in the matrices $\mathbf{W}$, $\mathbf{X}$ and $\mathbf{Y}$. A population of inhibitory neurons provides the inhibition to create a balanced oscillatory current. The oscillatory current is self-generated, and no external clock is needed, but the frequency of the oscillation is arbitrarily set to 5 Hz. This frequency relates the phases of the complex numbers to delays. The neurons have a refractory period half the cycle time, which prevents the neurons from firing more than once per cycle. This acts as the neural non-linearity that renormalizes the complex numbers to unit magnitude. The Hopfield dynamics drives the system into a stable oscillation that can be used to store and recall information.

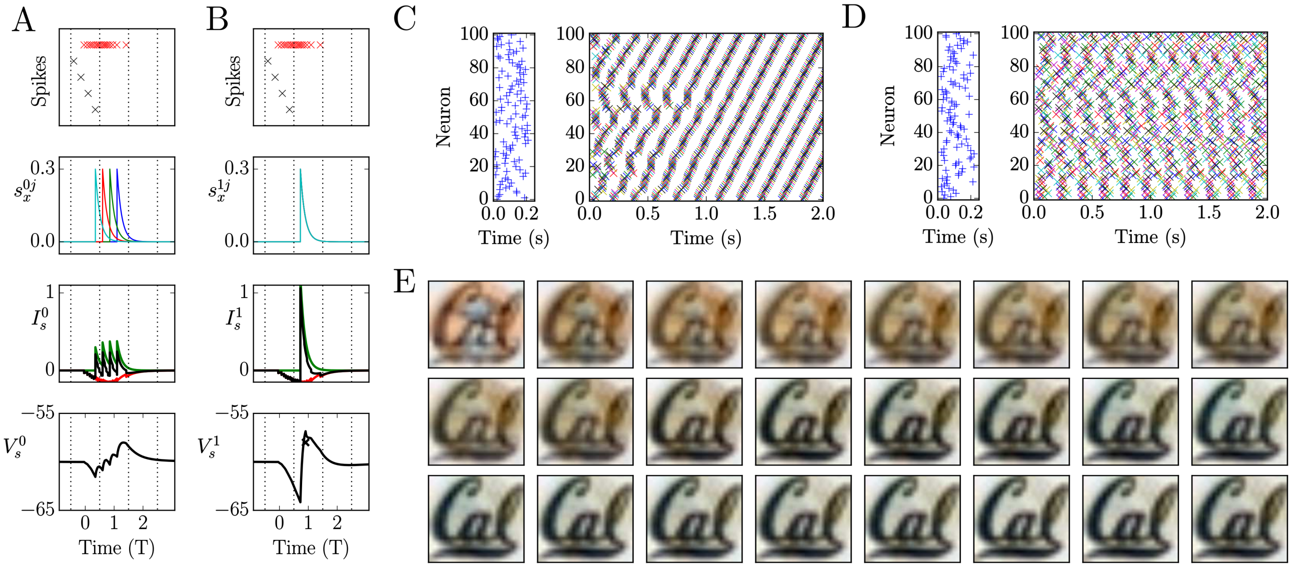

Figure 1: Memory network with phase-precession coding. A, B. The effect of a presynaptic spike train (Row 1; red x's: inhibitory neuron spikes) depends on synapse phase patterns (Row 2) impinging on a target neuron. The current into a postsynaptic neuron (Row 3) can be rather small (A, decoherent) or quite large (B, coherent). C. Several linear phase patterns are stored in a spiking Hopfield attractor, and the network is initialized randomly. After a short time it converges to a stored attractor state. D. E. Several images are stored in the spiking Hopfield attractor, and the network is initialized to a mixture of 3 of the images. After a few iterations, the network converges to the most strongly initialized stored pattern. Spiking patterns (D) and readout of spiking patterns (E).

A model of place representation in hippocampus

We next demonstrate how the computations in the FHRR model can be implemented in spiking neurons to produce neural response activity similar to CA1 principal neurons, exhibiting place fields and phase-precession. Some properties of the FHRR model described above can be modified to better match neuroscience mechanisms without losing the connection to the computational theory. In essence, we manipulate sparsity and the binding mechanism in the FHRR model to form unique neural codes of 2-D locations in different environments (Teylor and DiScenna 1986) as well as mimic observed physiology. This type of representation is interesting because sparse firing is a prominent feature of biological neurons and is naturally able to represent latent representations in sparse coding models (Olshausen Field, 1996; Rehn Sommer, 2007).

Representations in the FHRR model typically are vectors of unit complex numbers with random phases as the basic set of symbols, but this translates to every neuron firing once each cycle of the oscillation. Our work (Frady et al. 2018) showed that rate modulation can be applied to the basic set of vector-symbols without detriment to the capacity. By including rate modulation, we can model both place-fields and phase-precession within the activity of the neurons, as observed in neuroscience (Buzsaki, 2006).

Sparsity enables the modeling of both place-fields and phase-precession within the activity of the neurons as follows. If x and y are continuous values representing the position of a mouse in a 2D arena, one can form unique *location keys* using the powers of some base vectors $\mathbf{X}$ and $\mathbf{Y}$. The base vectors can be designed to capture the rate of phase-precession and the size of place fields, as well as preserve the sparsity of the neural code. In FHRR, circular convolution implements the binding operation: $\circledast$. The joint key uniquely representing a particular location in a particular environment can then be formed by: $\mathbf{K} \circledast \mathbf{X}^{\circledast x} \circledast \mathbf{Y}^{\circledast y}$, where K is a sparse random vector representing an index for the environment/context. [1]

The network of spiking neurons with the synaptic properties described above can be used to represent the FHRR vector $\mathbf{K} \circledast \mathbf{X}^{\circledast x} \circledast \mathbf{Y}^{\circledast y}$. We demonstrate how such a network encodes the position of a mouse exploring a 2-D box on a particular trajectory (Fig. 5A). The resulting spike responses of the network are shown in (Fig. 5B), in which both place-field responses and phase-precession are present (Fig. 5C). Another population of readout neurons with the same synaptic dynamics described above can decode the

[1] The expression $\mathbf{X}^{\circledast x}$ describes circular convolution exponentiation, i.e. $\mathbf{X}^{\circledast 3} = \mathbf{X} \circledast \mathbf{X} \circledast \mathbf{X}$. This is well-defined for continuous values in the exponent.

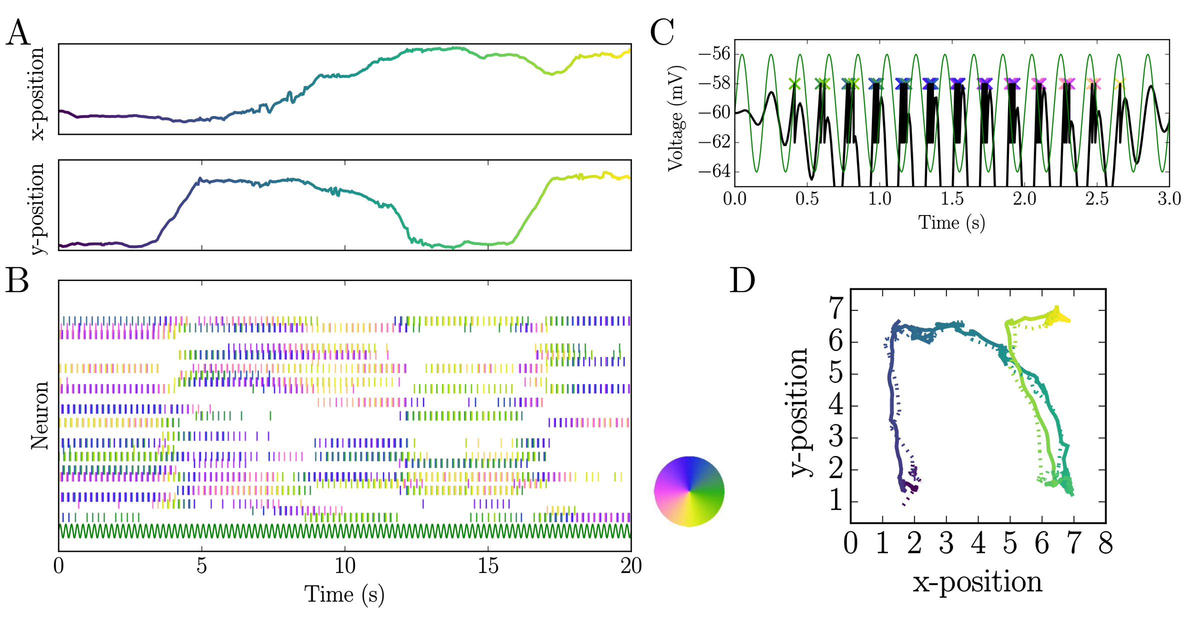

Figure 2: Place representations using phase precession neural coding model. A. The position of a mouse over time. B. Model spiking activity encoding position. The theta rhythm is depicted as the green trace at the bottom. Spikes are colored based on phase relative to theta oscillation (inset). C. Spikes of individual neuron phase-precess while mouse moves through a place field. The rate modulations and phase-precession reproduce characteristics of CA1 place cells. D. Position decoded from spiking activity (dashed line) compared to real position (solid line).

neural activity (Fig. 5D).

Discussion

The two-level theory framework described here can serve to link computation with neural mechanisms in a modular and flexible manner. This allows theorists to postulate computations in a generic VSA formalism and create models that can be compared to neuroscience data. The novel model of spike-timing and synaptic integration highlights a new way to understand and utilize neural oscillations and phase-precession for computation. This will serve to enhance our understanding of oscillatory neural computation observed in the nervous system. Further, our framework can serve as a new way to design algorithms for distributed computation, such as on emerging devices from neuromorphic computing (Davies et al. 2017).

Acknowledgements

This work was supported by the Intel Corporation (ISRA on Neuromorphic Architectures for Mainstream Computing), NSF award 1718991.

References

H. Aoki and Y. Kosugi. An Image Storage System Using Complex-Valued Associative Memories. *IEEE*, pages 626–629, 2000. ISSN 10514651.

G. Buzsáki and A. Draguhn. Neuronal oscillations in cortical networks. *Science*, 304(June):1926–1929, 2004. ISSN 1095-9203. 10.1126/science.1099745. URL `http://www.ncbi.nlm.nih.gov/pubmed/15218136`.

A. Hirose. Dynamics of Fully Complex-Valued Neural Networks. *Electronics Letters*, 28(16):1492–1494, 1992. ISSN 00135194. 10.1049/el:19920948.

J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982. ISSN 0027-8424. 10.1073/pnas.79.8.2554.

T. A. Plate. *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*. Stanford: CSLI Publications, 2003. ISBN 978-1575864303.

Buzsáki, G. (2006). *Rhythms of the Brain.* Oxford University Press.

Davies, M., Srinivasa, N., Lin, T. H., Chinya, G., Cao, Y., Choday, S. H., ... Liao, Y. (2018). Loihi: A neuromorphic manycore processor with on-chip learning. IEEE Micro, 38(1), 82-99.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., Rasmussen, D. (2012). A large-scale model of the functioning brain. Science, 338(6111), 1202-1205.

Frady, E.P., Kleyko, D., Sommer, F.T. (2018). A theory of sequence indexing and working memory in recurrent neural networks. Neural Computation, 30(6).

Gayler, R. W. (2004). Vector symbolic architectures answer

Jackendoff's challenges for cognitive neuroscience. arXiv preprint cs/0412059.
Jankowski, S., Lozowski, A., Zurada, J. M. (1996). Complex-valued multistate neural associative memory. IEEE Transactions on Neural Networks, 7(6), 1491-1496.
Kanerva, P. (1996). Binary spatter-coding of ordered K-tuples. Artificial Neural Networks—ICANN 96, 869-873.
Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. Cognitive Computation, 1(2), 139-159.
Knoblauch, A., Sommer, F.T. (2016). Structural plasticity, effectual connectivity and memory in cortex. Frontiers in Neuroanatomy
Knoblauch, A., Palm, G., Sommer, F. T. (2010). Memory capacities for synaptic and structural plasticity. Neural Computation, 22, 289-341
Olshausen, B. A., Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381, 607–9.
Palm, G., Sommer, F.T. (1992). Information and pattern capacities in neural associative memories with feedback for sparse memory patterns. In: Neural Network Dynamics, Springer New York. Eds.: J.G.Taylor, E.R.Caianello, R.M.J. Cotterill, J.W.Clark, 3-18
Plate, T. A. (1995). Holographic reduced representations. IEEE Transactions on Neural networks, 6(3), 623-641.
Plate, T. A. (2003). Holographic Reduced Representation: Distributed Representation for Cognitive Structures. Stanford: CSLI Publications, 2003. ISBN 978-1575864303.
Rehn, M., Sommer, F. T. (2007). A network that uses few active neurones to code visual input predicts the diverse shapes of cortical receptive fields. J. Comp. Neurosci., 22, 135–146.