An Information-Geometric Formulation of Pattern Separation and Evaluation of Existing Indices

Harvey Wang [1] Selena Singh [2] Thomas Trappenberg [1*] Abraham Nunes [1,3]

[1]Faculty of Computer Science, Dalhousie University, NS, Canada [2]Department of Psychology, Neuroscience and Behaviour, McMaster University, ON, Canada [3]Department of Psychiatry, Dalhousie University, NS, Canada

Outline

The hippocampal **dentate gyrus (DG)** performs **pattern separation (PS)** to aid in rapid and accurate encoding of associative memories

- Proper definition and measurement of PS are essential to understand its connection to conditions accompanied with cognitive decline, e.g. schizophrenia and Alzheimer's disease
- A common way to quantify PS is to compare pre-DG spike trains of individual neurons to corresponding post-DG spike trains

We describe PS as a computation that separates probability distributions that generate spike trains

- This gives a mathematically rigorous formalism that we used to expose open areas for further development and unify existing definitions

We show that existing approaches may not capture changes in temporal synchrony in neural patterns, which may carry neural information in hippocampal regions such as the DG [1, 2]

Our results encourage a search for new PS indices that are sensitive to coincident firing, correlated firing, and synchronization

Pattern Separation Described by a Statistical Manifold

We **formulate a neural network (e.g. the DG) as a statistical manifold $\mathcal{P}$**

- Each point on the manifold is a probability distribution $p(x;\xi)$ that generates spike trains x
- Each distribution is defined (parameterized) by the coordinates on the manifold ξ

The coordinates of such manifold includes intrinsic and extrinsic parameters:

- *Intrinsic:* network and cellular topology, geometry, and biophysics
- *Extrinsic:* inputs to the network from other areas (e.g. entorhinal cortex inputs to DG)

A manifold $\mathcal{P}$ with coordinates ξ describes PS iff small coordinate changes $\mathrm{d}\xi$ result in distant probability distributions $p(x;\xi)$ and $p(x;\xi+\mathrm{d}\xi)$

- $\mathcal{P}$ represents a neural pattern separator (e.g. the DG)
- ξ are functions of the inputs to the neural system (e.g. projections from the entorhinal cortex to the DG)
- Patterns x represent the input-driven output of the neural system (e.g. neural activity of the DG under a given stimulus)
- Distant distributions $p(x;\xi)$ and $p(x;\xi+\mathrm{d}\xi)$ generate highly dissimilar neural patterns x

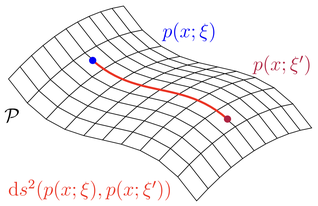

Figure 1. **Conceptualization of pattern separation on a 2-dimensional manifold $\mathcal{P}$.** The parameters ξ and $\xi' = \xi + \mathrm{d}\xi$ realize probability distributions $p(x;\xi)$ and $p(x;\xi')$ in the manifold, where the distance between distributions $d(p(x;\xi), p(x;\xi'))$ can be computed.

Distance between $p(x;\xi)$ and $p(x;\xi+\mathrm{d}\xi)$, in N-dimensions, is computed by summing up the changes in each individual coordinates $\mathrm{d}\xi_i, \mathrm{d}\xi_j \in \mathrm{d}\xi$

$$\mathrm{d}s^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} g(\xi_i,\xi_j)\,\mathrm{d}\xi_i\mathrm{d}\xi_j, \tag{1}$$

where g_{ij} is the Fisher information, which captures the magnitude and direction of change in the probability distribution with respect to each coordinate, between $\xi_i, \xi_j \in \xi$

$$g(\xi_i,\xi_j) = \left\langle \frac{\partial \log p(x;\xi)}{\partial \xi_i}\frac{\partial \log p(x;\xi)}{\partial \xi_j} \right\rangle_x. \tag{2}$$

There Exists No Single Definition of Pattern Separation

The common definition of pattern separation [3, 4, 5] can be formalized as follows:

A function performs pattern separation iff it maps similar neural patterns to dissimilar ones. More specifically, for some dissimilarity or distance index d, the mapping $x \mapsto f(x)$ performs pattern separation iff

$$d(x_1,x_2) < d\left(f(x_1), f(x_2)\right).$$

Clearly, the above gives **not a single, but a multitude of definitions**, depending on the form of the index d. We compare the most common definitions of PS [1], as well as novel indices based on information theory [6].

Existing Definitions of Pattern Separation are Incomplete

We use Eq. 3 to define two probability distributions representing the activity of two systems, $p(x;\eta_1,\eta_2,0)$ and $p(x;\eta_1,\eta_2,\mathrm{d}\theta)$. These systems have equal marginal rates in each neuron (quantified by η_1 and η_2), but different rates of coincidental firing. Specifically, the neurons in the first system fire independently ($\theta = 0$), and the rate of coincidental firing in the second system range from highly anti-correlated ($\mathrm{d}\theta < 0$) to highly correlated ($\mathrm{d}\theta > 0$).

The systems are set up such that the distance between two probability distributions are always *only* proportional to their difference in coincidental firing rate $\mathrm{d}\theta$ (Eq. 4).

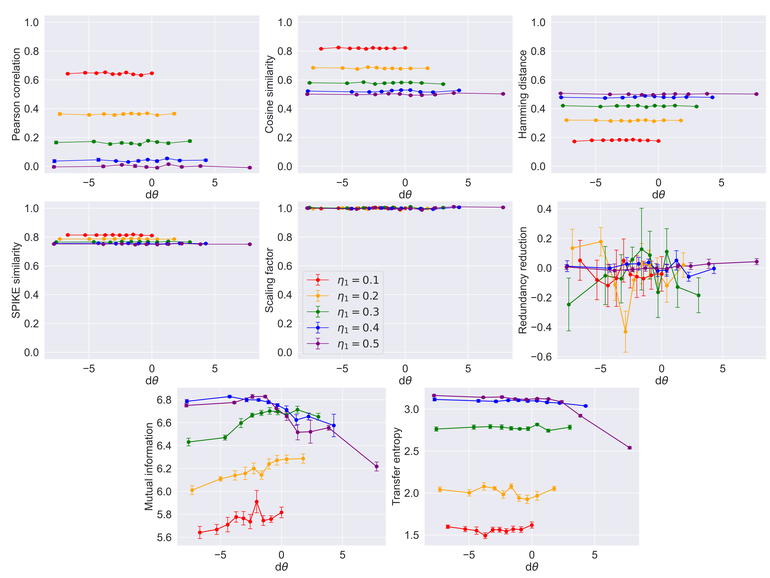

Figure 2. **Measurements of existing PS indices versus change in the rate of coincidental firing.** Indices were applied to spike trains generated from $p(x;\eta_1,\eta_2,0)$ and $p(x;\eta_1,\eta_2,\mathrm{d}\theta)$, over different values of η_1 and η_2 shown by different coloured plots (we set $\eta_1+\eta_2=1$), and averaged over 10 samples. The error bars show the standard error of the mean. SPIKE similarity was implemented based on [1, 7]. Mutual information, transfer entropy, and redundancy reduction were computed using the toolbox from [6].

Figure 2 shows measurements of existing PS indices between patterns with equal firing rates but differing spike timing. The flatness of the plots show that these indices are insensitive to PS that arise from the differences in the relative timing of the spikes and/or the correlation between neurons.

2-Neuron Model in a 3-Dimensional Manifold

For simplicity, we consider a 2-neuron system and its firing pattern x, given by the probability $p(x)$. At any given time, the neurons have some probabilities of:

(1) firing together, p_{11} (2) firing separately, p_{10} and p_{01} (3) *not* firing together, p_{00}

Since $p_{00}+p_{10}+p_{01}+p_{11}=1$, the manifold for this model is 3-dimensional.

We use the neurons' marginal firing rates, η_1 and η_2, and the coincidental firing rate, θ, as the 3 coordinates of this manifold,

$$\eta_1 = \langle x_1\rangle = p_{10}+p_{11}, \quad \eta_2 = \langle x_2\rangle = p_{01}+p_{11}, \quad \theta = \log\frac{p_{11}p_{00}}{p_{10}p_{01}}. \tag{3}$$

These coordinates were used because (η_1,η_2) or θ are *orthogonal* [2], which means that their Fisher information vanishes: $g(\eta_1,\theta) = g(\eta_2,\theta) = 0$ (from Eq. 2).

This orthogonality allows us to generate patterns with (1) equal coincidental firing rates, (2) equal marginal firing rates, or (3) equal coincidental *and* marginal firing rates in a very controlled manner. In particular, for two distributions with equal marginals η_1 and η_2, their distance is simply proportional to their difference in the rates of coincidental firing (from Eq. 1):

$$\mathrm{d}s^2\big|_{\mathrm{d}\eta_1=\mathrm{d}\eta_2=0} = g(\theta,\theta)\,(\mathrm{d}\theta)^2. \tag{4}$$

Discussions

Figure 3 shows a potential factor that inhibits the ability for existing PS indices to capture neuron-neuron correlations within an ensemble. Common experimental protocols convert a pattern of spike times (from each neuron in a given system) into binary vectors, which are then concatenated end-to-end into a single vector. As a result, the spiking activity of a neuron is only compared to itself, as opposed to that of other neurons in the system. If a PS index is only sensitive to the rate of firing, and not to the relative timing of spikes, this would explain the flatness observed in Figure 2.

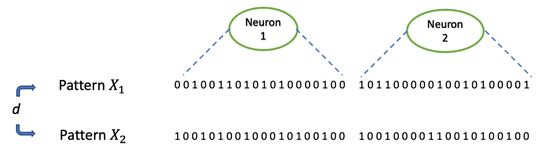

Figure 3. Binary representations of neural patterns X_1 and X_2 from two 2-neuron systems. Common experimental protocols divide the observation window into bins that are assigned the value 1 if the neuron is active at least once, and 0 otherwise. Vectors representing the activities of different neurons in the system are then concatenated into a single vector.

There is mounting evidence for the relevance of population structure and relative timing of spikes in conveying neural information. In particular, mossy fibre sprouting (MFS) in the DG, which has been linked to PS performance. MFS is characterized by the growth of granule cell axons into their own dendritic field, creating *de novo* recurrent excitatory circuits that may change pairwise neuron-neuron interactions and alter neural information. Therefore, we must incorporate into PS research methods that are sensitive to within-system correlational structures.

References

[1] A. D. Madar, L. A. Ewell, and M. V. Jones, "Temporal pattern separation in hippocampal neurons through multiplexed neural codes," *PLOS Computational Biology*, vol. 15, p. e1006932, Apr. 2019.

[2] H. Nakahara and S.-i. Amari, "Information-Geometric Measure for Neural Spikes," *Neural Computation*, vol. 14, pp. 2269–2316, Oct. 2002.

[3] H. Wigström, "Spatial propagation of associations in a cortex-like neural network model: Spatial Propagation of Associations," *Journal of Neuroscience Research*, vol. 3, no. 4, pp. 301–319, 1977.

[4] E. T. Rolls and R. P. Kesner, "A computational theory of hippocampal function, and empirical tests of the theory," *Progress in Neurobiology*, vol. 79, pp. 1–48, May 2006.

[5] M. A. Yassa and C. E. Stark, "Pattern separation in the hippocampus," *Trends in Neurosciences*, vol. 34, pp. 515–525, Oct. 2011.

[6] A. D. Bird, H. Cuntz, and P. Jedlicka, "Robust and consistent measures of pattern separation based on information theory and demonstrated in the dentate gyrus," preprint, Neuroscience, Nov. 2022.

[7] T. Kreuz, D. Chicharro, C. Houghton, R. G. Andrzejak, and F. Mormann, "Monitoring spike train synchrony," *J Neurophysiol*, p. 16, 2013.