Optimal Population Vector Spatial Correlation Predicts Place-like Firing fields

February 2023

1 Introduction

We postulate that all optimal representations of (allocentric) spatial position share a fundamental set of properties

1. Uniqueness of representations
    - All spatial positions should be uniquely encoded
2. Robustness
    - Only large perturbations can force confusion between distant locations
    - Energy conservation!
3. Ordering
    - Representations of locations that are close in physical space, are similar in representation
4. Invariance
    - There are no "special" locations in space, so there should be no "special" representations, and the magnitude of the representation is equal everywhere.

We can get all of these properties if we demand that the population vector correlation has a rotationally symmetric, peaked structure. One of the simplest ways of achieving this, is by demanding that the population vector correlation approximate the following Gaussian correlation function at every spatial point

$$C(\mathbf{r}, \mathbf{r}') = e^{-\frac{1}{2\sigma_c^2} d(\mathbf{r},\mathbf{r}')^2}, \tag{1}$$

where σ_c is the population correlation tuning width, and d is a distance function. In this work, we will only consider the Euclidean distance $d(\mathbf{r}, \mathbf{r}') = \sqrt{|\mathbf{r} - \mathbf{r}|^2}$.

Intuitively, if the (un-normalized) population vector spatial correlation adheres perfectly to (1), all spatial locations are uniquely encoded (as the correlation is only maximal at one location), thus meeting property 1. Furthermore, points that are close in space are represented using similar population vectors, while distant locations become uncorrelated (depending on the tuning width σ_c), satisfying condition 3.

Notably, the choice in (1) also leads to robust neural codes, in the sense that spatially distant points are encoded by dissimilar population vectors. As such, small perturbations will only change

the population vector into similar states, and the representation will remain in the neighborhood of the correct representation. Thus, the choice in (1) is one way of achieving a robust representation, and meeting requirement 3.

Finally, point 4 needs some more thought, but could either be enforced through regularization, or is indirectly imposed by using an unnormalized correlation measure!

A Variational Model of Place Cells

Consider a firing rate model of an ensemble of N neurons. Collectively, the firing rate functions of these neurons set up a population vector $\mathbf{p} \in \mathcal{P}$, which in the noise-free case, can be viewed as a functions of the true position in a given enviroment Ω. We consider only flat, two-dimensional environments, so $\Omega \subset \mathbb{R}^2$.

For a given firing rate function p_i, $i = 1, 2, 3, ..., N$, we impose biologically plausible constraints in the form of non-negativity and boundedness, i.e.

$$0 \leq p_i \leq M, \quad i = 1, 2, 3, ..., N,$$

where $M \in \mathbb{R}$ is some upper bound on the largest attainable firing rate.

With the properties of our ensemble established, we turn to formulating an optimization objective that encourages a population vector correlation structure similar to that given in (1). However, as our ensemble is finite, we seek the rate functions $\{p_i\}_{i=1}^N$ that most closely approximates the Gaussian correlation structure at every point in the domain. Since these are unknown functions, we consider the variational objective

$$S = \int_\Omega \left(C(\mathbf{r}, \mathbf{r}') - \mathbf{p}(\mathbf{r}) \cdot \mathbf{p}(\mathbf{r}')\right)^2 d\mathbf{r} d\mathbf{r}' \tag{2}$$

where $\mathbf{p}(\mathbf{r}) \cdot \mathbf{p}(\mathbf{r}') = \sum_i p_i(\mathbf{r}) p_i(\mathbf{r}')$ is an (un-normalized) measure of the population correlation at distinct coordinates r and r'.

Neural Network Parametrization

As a first investigation of what optimal position representations $\{p_i\}_{i=1}^N$ may emerge from extremizing (2), we iteratively minimize the Monte Carlo approximation

$$S \approx \sum_{j \geq i;}^{B} (C(\mathbf{r}_i, \mathbf{r}_j) - \mathbf{p}(\mathbf{r}_i) \cdot \mathbf{p}(\mathbf{r}_j)))^2,$$

using stochastic gradient descent, with B being the minibatch size. We parametrize the firing field functions by a neural network. Specifically, we use a densely connected, feedforward network with all-ReLU activation functions, and 50 output units (full architecture: [FC512, ReLU, FC512, ReLU, FC50, ReLU]). Notably, the activation function guarantees that the learned representations are all non-negative.

The model was trained with the Adam optimizer with a learning rate of 0.001 for a total of 10000 training steps. For every training step, space was sampled in minibatches of size $B = 256$. The domain Ω was taken to be a square 2×2 region, sampled randomly and uniformly. The tuning width σ_c was taken to be 0.4.

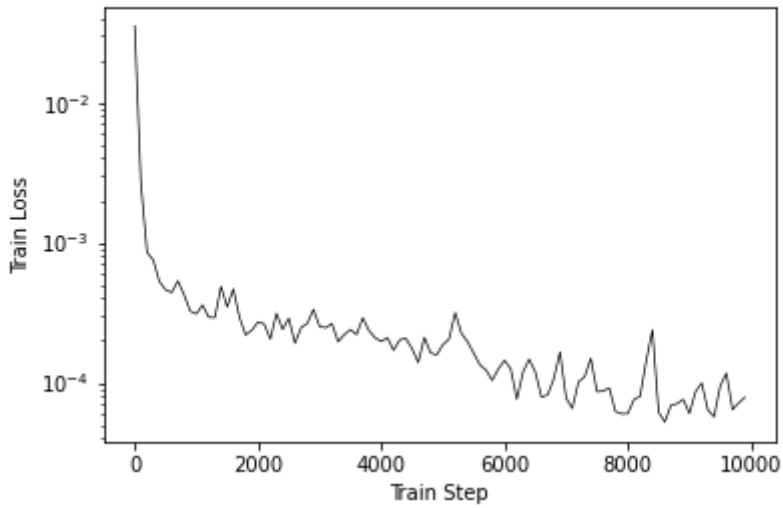

Figure 1: Training loss history for the variational Monte Carlo neural network objective.

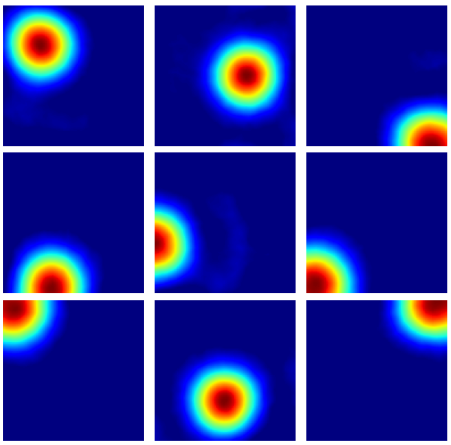

Figure 2: Example of the learned spatial representations. Shown are spatial ratemaps of 9 select output units of the network trained to minimize the Monte Carlo variational correlation objective.

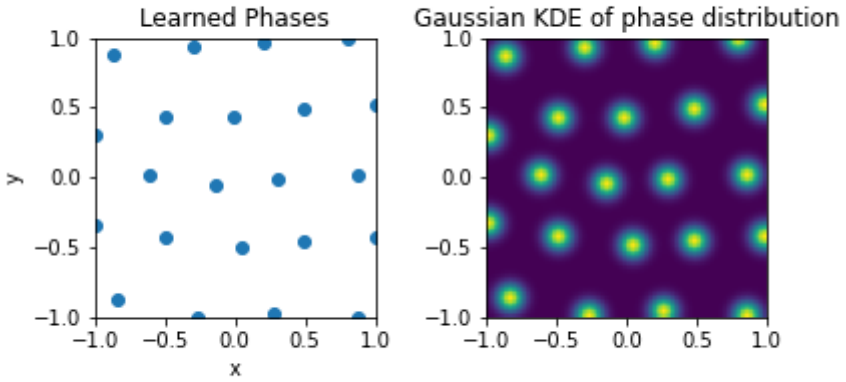

Figure 3: Phase distribution of the trained, place-like ensemble. Left: locations corresponding to the maximal rate of every output unit in the network trained to minimize the Monte Carlo variational correlation objective. Right: kernel density estimate of the phase distribution, highlighting that phases appear to fall on the vertices of a hexagonal grid.

Results

2 The First Variation

$$S = \int C^2(\mathbf{r}, \mathbf{r}') + C(\mathbf{r}, \mathbf{r}')\mathbf{p}(\mathbf{r}) \cdot \mathbf{p}(\mathbf{r}') + (\mathbf{p}(\mathbf{r}) \cdot \mathbf{p}(\mathbf{r}'))^2 d\mathbf{r}d\mathbf{r}'$$

$$S = \int C^2(\mathbf{r}, \mathbf{r}')d\mathbf{r}d\mathbf{r}' + \int C(\mathbf{r}, \mathbf{r}') \sum_i p_i(\mathbf{r})p_i(\mathbf{r}')d\mathbf{r}d\mathbf{r}' + \int \left(\sum_i p_i(\mathbf{r})p_i(\mathbf{r}') \right)^2 d\mathbf{r}d\mathbf{r}'$$

The first term, l_1 is unchanged by a perturbation $p_i \to p_i + \varepsilon\eta_i$. The second term, denoted l_2 does change:

$$l_2 \to \int C(\mathbf{r}, \mathbf{r}') \sum_i p_i(\mathbf{r})p_i(\mathbf{r}') + \varepsilon(p_i(\mathbf{r})\eta_i(\mathbf{r}') + p_i(\mathbf{r}')\eta_i(\mathbf{r})) + \varepsilon^2\eta_i(\mathbf{r})\eta_i(\mathbf{r}')d\mathbf{r}d\mathbf{r}'.$$

Taking the ordinary derivative, and evaluating at $\varepsilon = 0$, we find

$$\left.\frac{dl_2}{d\varepsilon}\right|_{\varepsilon=0} = \int C(\mathbf{r}, \mathbf{r}') \sum_i p_i(\mathbf{r})\eta_i(\mathbf{r}') + p_i(\mathbf{r}')\eta_i(\mathbf{r})d\mathbf{r}d\mathbf{r}'.$$

The variation of the last term, l_3 becomes

$$dl_3 \to \int \left(\sum_i (p_i(\mathbf{r})p_i(\mathbf{r}') + \varepsilon(p_i(\mathbf{r})\eta_i(\mathbf{r}') + p_i(\mathbf{r}')\eta_i(\mathbf{r})) + \varepsilon^2\eta_i(\mathbf{r})\eta_i(\mathbf{r}') \right)^2 d\mathbf{r}d\mathbf{r}',$$

and so

$$\left.\frac{dl_3}{d\varepsilon}\right|_{\varepsilon=0} = 2\int\left(\sum_i p_i(\mathbf{r})p_i(\mathbf{r}')\right)\left(\sum_i p_i(\mathbf{r})\eta_i(\mathbf{r}') + p_i(\mathbf{r}')\eta_i(\mathbf{r})\right)d\mathbf{r}d\mathbf{r}'.$$

and so

$$\left.\frac{dS[\mathbf{p}+\varepsilon\boldsymbol{\eta}]}{d\varepsilon}\right|_{\varepsilon=0} = 0 \implies \int\left(\sum_i p_i(\mathbf{r})\eta_i(\mathbf{r}') + p_i(\mathbf{r}')\eta_i(\mathbf{r})\right)\left(C(\mathbf{r},\mathbf{r}') - 2\sum_i p_i(\mathbf{r})p_i(\mathbf{r}')\right)d\mathbf{r}d\mathbf{r}' = 0.$$

If we denote the factor on the right hand $u(\mathbf{r},\mathbf{r}')$, we can write

$$\int\left(\sum_i p_i(\mathbf{r})\eta_i(\mathbf{r}') + p_i(\mathbf{r}')\eta_i(\mathbf{r})\right)u(\mathbf{r},\mathbf{r}')d\mathbf{r}d\mathbf{r}' = 0,$$

and by noticing that $u(\mathbf{r},\mathbf{r}') = u(\mathbf{r}',\mathbf{r})$, i.e. u is symmetric in its arguments, we can change the order of integration and switch labels on one term!

$$0 = \sum_i \int u(\mathbf{r},\mathbf{r}')p_i(\mathbf{r}')\eta_i(\mathbf{r})d\mathbf{r}d\mathbf{r}' + \int u(\mathbf{r}',\mathbf{r})p_i(\mathbf{r}')\eta_i(\mathbf{r})d\mathbf{r}'d\mathbf{r}.$$

$$0 = \sum_i \int u(\mathbf{r},\mathbf{r}')p_i(\mathbf{r}')\eta(\mathbf{r})d\mathbf{r}d\mathbf{r}'$$

.

$$0 = \sum_i \int\int u(\mathbf{r},\mathbf{r}')p_i(\mathbf{r}')d\mathbf{r}'\eta(\mathbf{r})d\mathbf{r},$$

, which, according to the fundamental lemma of VC, means that

$$\int u(\mathbf{r},\mathbf{r}')p_i(\mathbf{r})'d\mathbf{r}' = \int\left(C(\mathbf{r},\mathbf{r}') - 2\sum_i p_i(\mathbf{r})p_i(\mathbf{r}')\right)p_i(\mathbf{r})d\mathbf{r}' = 0.$$

Ouch.