

Neural coordination on encoding information from multiple simultaneously presented stimuli

Yunran Chen

Abstract

We are interested in how populations of neurons coordinate to encode information from multiple simultaneously presented stimuli. We conjecture that information about each presented stimulus may be aggregated within distinct subpopulations of neurons. We aim to provide some insights to this conjecture by principle component analysis (PCA) and factor model. From PCA, we find some evidence suggesting distinct subpopulations of neurons may exist depending on the preference of neurons to constituents of simultaneously presented stimuli. This evidence motivates a two-factor random factor Model selected by BIC, which provides a decent estimation for sample covariance matrix. And from the factor loadings matrix, we find the latent factors are interpretable, which may relate to the preference of a single neuron on constituents of the simultaneously presented stimuli.

1 Introduction

We are capable to respond when multiple objects are presented simultaneously in a natural scene. However, the underlying mechanism of how neurons encode information when multiple stimuli are simultaneously presented is still poorly understood. Recent research provides evidence suggesting that a single neuron may encode dual-stimuli information by fluctuating between two constituent signals (Caruso et al., 2018; Glynn et al., 2019). Here we are curious about how such fluctuation patterns co-vary across neurons. Specifically, is the coordination positive, suggesting all the neurons tend to encode the same signal? Is the coordination negative, where distinct subpopulations of neurons take complementary turns to encode different signals from trials to trials? Or is the coordination weak, where each neuron just makes an independent choice encoding either stimulus A or B from trials to trials? A plausible conjecture is that information about each presented stimulus may be aggregated within distinct subpopulations of neurons. We aim to provide some insights to this conjecture by exploratory analysis, which can serve as a building block on the future research aiming at developing a model capturing the coordinating encoding patterns among neurons when multiple stimuli are simultaneously presented.

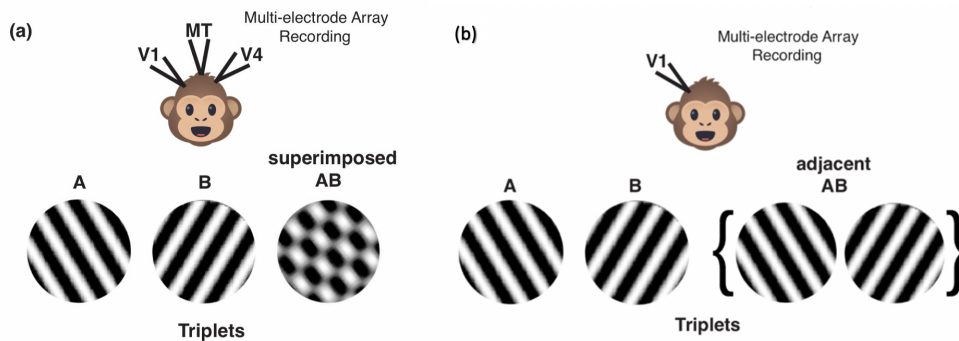


Figure 1: Experiment design for multiple cells recordings when presented two types of condition triplets.

Our analysis is based on a subset of the dataset provided by Ruff and Cohen (2016). In the experiment, a monkey was trained to make saccades towards single stimulus (gratings) or two different types of dual stimuli (superimposed or adjacent gratings) as shown in Figure 1. A sensor is plugged in V1 brain area (visual cortical area) of the monkey, recording the total spike counts of multiple neurons simultaneously. Here we denote single stimulus conditions as A and B respectively, and dual stimuli condition as AB. Under each condition, a spike count matrix is provided with each row representing repeated trials under the same condition, each column representing 97 recorded neurons. 21, 22, 24 repeated trials were conducted under $\{A, B, AB(\text{superimposed})\}$ conditions, and 16, 10, 17 repeated trials were conducted under $\{A, B, AB(\text{adjacent})\}$ conditions.

Exploring the dependence structure of spike counts matrices and make comparison between the single-stimulus and dual-stimuli conditions may help to unveil coordinate patterns among populations of neurons. We first apply Principle Component Analysis to explore how the dependence structure among neurons under dual-stimuli conditions are different from those under single-stimulus condition. Specifically, we compare the loadings of neurons onto the first two principle components under dual-stimuli condition to those under single-stimulus condition. We find some evidence suggesting distinct subpopulations of neurons may exist depending on the preference of neurons to constituents of simultaneously presented stimuli. This evidence motivates a Factor Analysis Model to assume latent factors governing the neurons' firing patterns when exposed to simultaneously presented stimuli. Such latent factors may relate to the preference of a single neuron on constituents of the simultaneously presented stimuli.

It is worth noting that this dataset contains two different types of dual-stimuli trials: superimposed and adjacent gratings. The adjacent AB signal represents the case where dual stimuli are presented simultaneously. In contrast, the superimposed AB signal may be recognized as a new signal which is uncorrelated to either A or B. Comparison of our analysis under this two different conditions validates this conjecture.

2 Exploratory Data Analysis

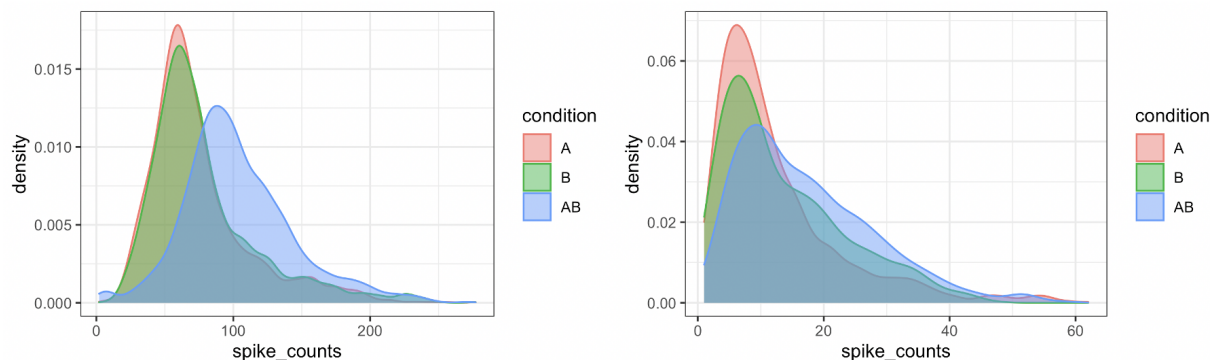


Figure 2: Density plots of spike counts among different conditions for superimposed triplet (left) and adjacent triplet (right)

Figure 2 presents the density of spike counts under superimposed-type triplet and adjacent-type triplets respectively. For the superimposed-type triplet (left), the medians of the spike counts are around 65, 68, 98 respectively for $\{A, B, AB\}$. For the adjacent-type triplet, the medians of spike counts are around 9, 11, 15 respectively for $\{A, B, AB\}$. We can see spike counts are more disperse under dual-stimuli trials compare to that under single-stimulus trials. And for superimposed stimuli, neurons exhibits higher spike count compared to that under single-stimulus trials, suggesting neurons may preceive the superimposed dual-stimuli as a new signal. In contrast, for adjacent dual-stimuli trials, a center shift is not obvious, suggesting neurons may preserve the information of the constituents of dual stimuli.

Figure 3 presents the density of pairwise spike counts correlation under two different types of triplets. Under single-stimulus and superimposed dual-stimuli conditions, the center of the curve is positive evidently. In contrast, the pairwise correlation under adjacent dual-stimuli condition is centered around 0 with a slight

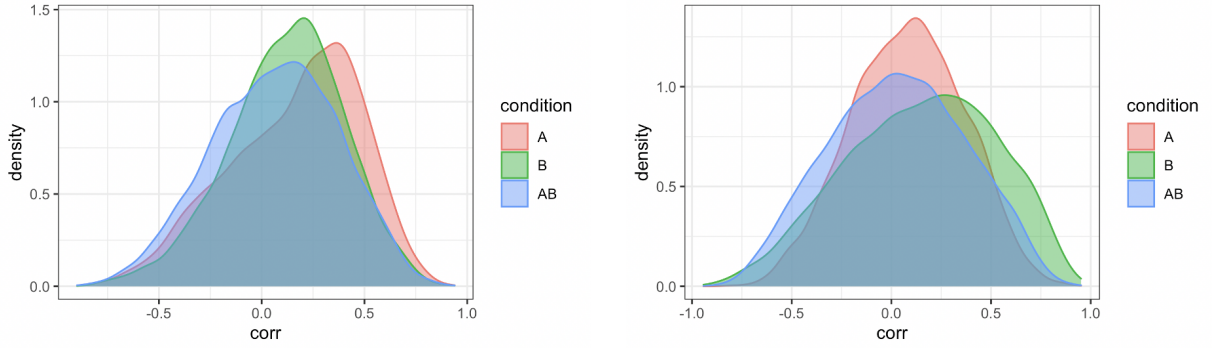


Figure 3: Density plots of pairwise correlation of correlation matrices among different conditions for superimposed triplet (left) and adjacent triplet (right)

positively shift. This findings provide evidence favoring an existence of distinct subpopulations of neurons for preserving different constituents of simultaneously presented stimuli opposed to positive coordination and negative coordination mentioned in Section 1.

3 Methods

3.1 Principle Component Analysis

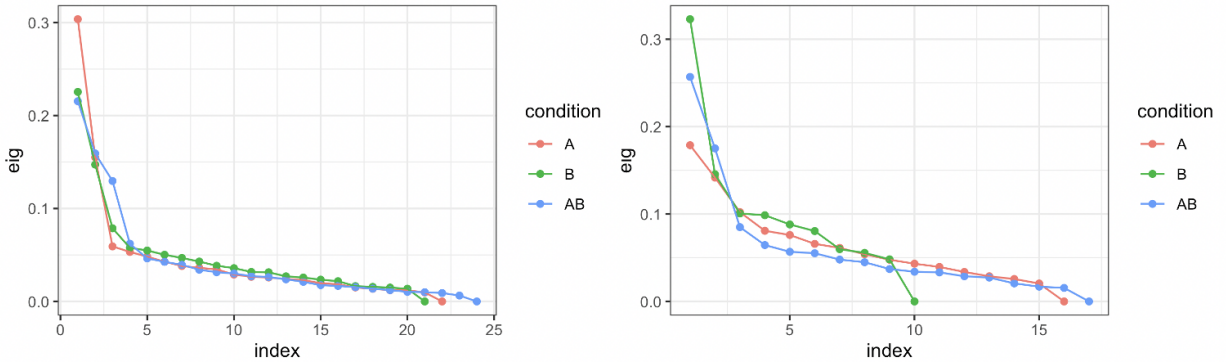


Figure 4: Scatter plot of eigen values in percentage $\lambda_k / \sum_i \lambda_i$ for the superimposed-type triplet (left) and the adjacent-type triplet (right) respectively

The multi-electrode sensor provides spike count recordings for 97 neurons, but only a small number of repeated trials are available under each condition (ranging from 10 to 24). Such a high dimension covariance matrix is common in nature science since the budget for experiments is limited. To visualize the high dimension covariance for some insights on the dependence structure, we consider the Principle Component Analysis (PCA) for a dimension reduction. Principle Component Analysis contributes to projecting data onto the lower-dimensional space spanned by the first few principle components while preserving variation of data to a great extent. Furthermore, the loadings of variables onto the principle axis can provide us insights on the relations among variables. Let Y be a $n \times p$ matrix with each row representing i.i.d observations (trials) under a specific condition. We apply an eigen decomposition to the sample covariance matrix (centered) $S = Y^T C Y = E^T E = V \Lambda V^T$, where C , E , Λ , and V represent centering matrix, columnwise-centered data matrix, diagonal matrix with eigen values ranked in a descending order and corresponding eigen vectors

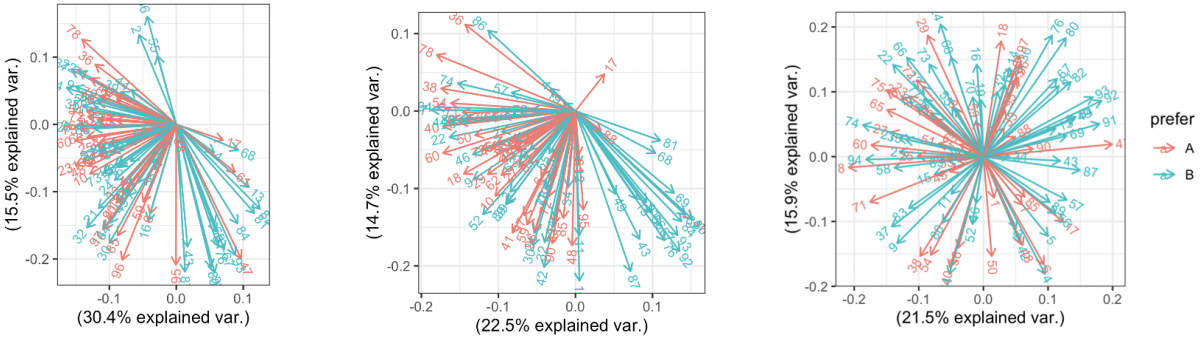


Figure 5: Loadings of neurons on first two principle axis under superimposed triplet condition $\{A, B, AB\}$

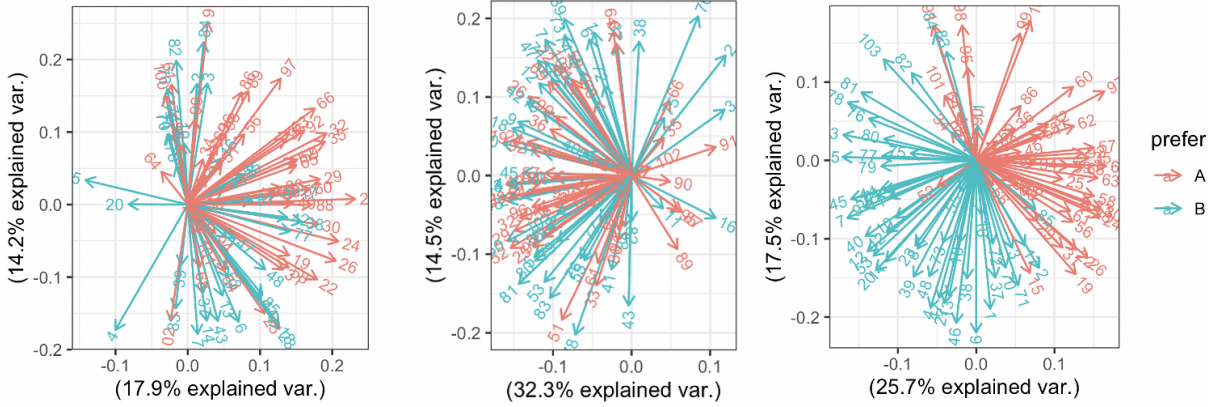


Figure 6: Loadings of neurons on first two principle axis under adjacent triplet condition $\{A, B, AB\}$

respectively. Let $F = EV$, where the k th column of F is $f_k = Ev_k$, the k th principle component, and v_k depicts the loadings of variables onto the k th principle axis.

Figure 4 depicts the percentage of variation explained by each principle component, denoted as $\lambda_k / \sum_i \lambda_i$ for the superimposed-type triplet (left) and the adjacent-type triplet (right) respectively. For the superimposed-type triplets, the elbow points appear in the third eigen values under single-stimulus conditions; in contrast, the elbow value appears in the fourth eigen value under the dual-stimuli condition. Compared to other two conditions, the first principle component under single stimulus A explains greater percentage of the total variance, which around 30% opposed to 20% for other two conditions. For the adjacent-type triplets, the elbow points appear the same for three conditions, which is the third largest eigen values. Interestingly, we observe the percentage of variance explained by the first principle component under the dual-stimuli condition lies between those under the single-stimulus condition.

We consider visualizing the loadings of variables on first two principle axis as shown in Figure 5 and Figure 6. To further interpret the principle components, we introduce a variable depicting the preference between stimulus A and B for a single neuron. If a neuron exhibits a larger spike count under stimulus A compared to that under stimulus B, we call the neuron prefers stimulus A, and vice versa. The third panel in Figure 6 (adjacent dual-stimuli condition) suggests neurons with the same preference tend to cluster together in the space spanned by the first two principle axis. This clustering behavior suggests neurons with the same preference may have similar fluctuation patterns between constituent stimulus A and B. In contrast, there is no obvious clustering indicated by neurons' preference for other conditions. For the single-stimulus conditions, we observe the most of loadings of neurons gathering together within a subspace, suggest a positive correlation among most of neurons, which is consistent to exploratory analysis shown in Figure 3.

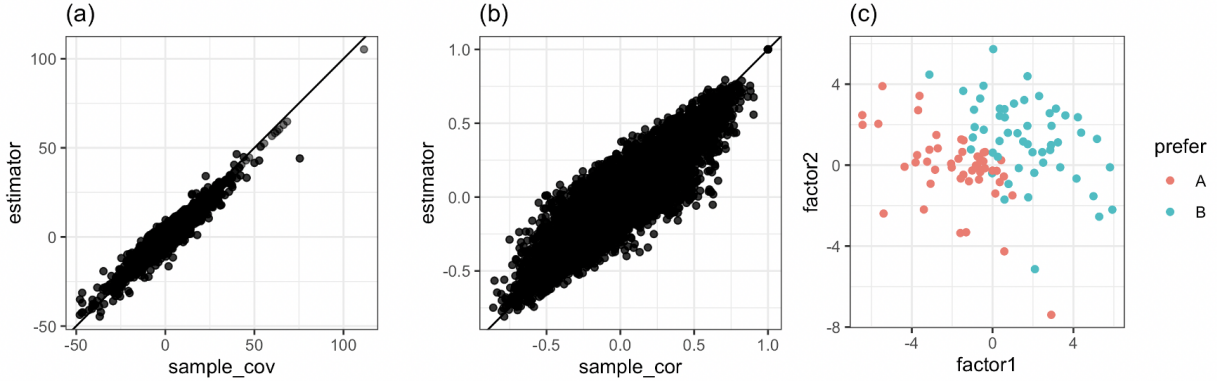


Figure 7: Factor analysis on spike count matrix under adjacent dual-stimuli condition. (a) estimated covariance matrix versus sample covariance matrix. (b) estimated correlation matrix versus sample correlation matrix. (c) factor loadings colored by the preference of each single neuron

3.2 Factor Analysis

Based on the findings by PCA, we may expect a lower-dimensional dependence structure embedding in the covariance matrix of the spike count matrix under adjacent dual-stimuli condition. We consider a Factor Model to describe the variation in terms of unobserved latent factors. We consider Gaussian q-factor model described as follows:

$$Y = 1\mu^T + ZA^T + E\Psi^{1/2},$$

with corresponding representation of the covariance matrix as

$$\text{Var}(y_i) = AA^T + \Psi,$$

where $\mu \in \mathbb{R}^p$, $A \in \mathbb{R}^{p \times q}$, $\Psi = \text{diag}(\psi_1, \dots, \psi_p)$, $Z \sim N_{n \times q}(0, I \otimes I)$, $E \sim N_{n \times p}(0, I \otimes I)$, Z and E are independent.

We estimate MLEs for (μ, A, Ψ) based on EM algorithm, and select the number of factors based on BIC. We built 2-factor random factor model selected by BIC, the resulting estimated covariance and correlation matrix are presented in Figure 7. Figure 7(a)(b) show most points are lying along the $y = x$ line, indicating a good performance of proposed estimator. Figure 7(c) presents the factor loading matrix colored by the preference of each single neuron. We can observe two distinct clusters, suggesting the latent factors may relate to the preference of each single neuron, which is consistent from the inference we obtained from PCA (Figure 6).

4 Results

Results from exploratory data analysis and principle component analysis both provide evidence favoring the conjecture that neurons perceive superimposed dual-stimuli as a new signal but encode adjacent dual-stimuli based on constituents of stimuli. Results from principle component analysis suggest neurons with the same preference may have similar fluctuation patterns when exposed to multiple simultaneously presented stimuli (adjacent dual-stimuli in the experiment setting). The firing behavior of most neurons tend to be positively related when exposed to a single stimulus (single stimulus or superimposed dual-stimuli). Two-factor random factor model can well capture the covariance matrix (or the correlation matrix) under adjacent presented dual stimuli. And the factor loadings are related to the preference of a single neuron. Both analysis from PCA and factor model provide evidence favoring the existence of distinct subpopulations based on neurons' preference on constituents of stimuli.

5 Discussion

Our two-factor random factor model has a decent performance on approximating the covariance matrix and correlation matrix. However, we assume the discrete spike count data follows normal distribution approximately. Spike count data is usually modelled as a Poisson distribution, which approximates to a normal distribution only when the rate is sufficiently large. In our dataset, the mean of spike counts under adjacent dual-stimuli condition is 17, which may not be promising enough. We can apply a factor analysis on count data using Poisson model (Larsson, 2020; Wedel et al., 2003).

Results from principle component analysis and factor models provide evidence showing populations of neurons may coordinate to encode multiple simultaneously presented stimuli by assembling information from distinct constituents of stimuli within distinct subpopulations related to neurons' preference. The future research direction is to develop a constrained factor model based on the conjecture (Chen et al., 2020). Specifically, we are interested in imposing a block structure on the factor loading matrix A , which require A having at most k distinct rows. This decomposition of covariance matrix can be reparameterized in the following form:

$$S = PBP^T + \Psi, \tag{1}$$

where P is a $n \times k$ binary matrix indicating the group allocation for each neuron, B is a $k \times k$ non-negative definite matrix. Neurons within the same clusters share the same factor loadings and behave the same while neurons from distinct clusters behave differently. The block structure can directly validate the conjecture that the neurons may coordinate based on subpopulations, providing an attractive biological interpretation. In addition, block constraint on factor loading matrix can further embed the matrix into a lower dimensional space to simplify the estimator and ease the computation. Furthermore, we can consider incorporating the spike count matrix under constituent stimulus A and B into the modeling. Caruso et al. (2018) assume spike count under dual-stimuli condition Y^{AB} comes from Poisson mixtures $\alpha \text{Poi}(\lambda^A) + (1 - \alpha) \text{Poi}(\lambda^B)$, where assuming spike count under single-stimulus condition Y^A and Y^B come from $\text{Poi}(\lambda^A)$ and $\text{Poi}(\lambda^B)$ respectively. We may extend the model for a single neuron to multivariate analysis for multiple neurons using Gaussian Copula model, with the marginal distribution as Poisson mixtures, the covariance matrix specified in equation 1.

Code

The code and data are available in github: <https://github.com/YunranChen/832DataProject.git>

References

- Caruso, V. C., J. T. Mohl, C. Glynn, J. Lee, S. M. Willett, A. Zaman, A. F. Ebihara, R. Estrada, W. A. Freiwald, S. T. Tokdar, et al. (2018). Single neurons may encode simultaneous stimuli by switching between activity patterns. *Nature communications* 9(1), 1–16.
- Chen, E. Y., R. S. Tsay, and R. Chen (2020). Constrained factor models for high-dimensional matrix-variate time series. *Journal of the American Statistical Association* 115(530), 775–793.
- Glynn, C., S. T. Tokdar, A. Zaman, V. C. Caruso, J. T. Mohl, S. M. Willett, and J. M. Groh (2019). Analyzing second order stochasticity of neural spiking under stimuli-bundle exposure. *arXiv preprint arXiv:1911.04387*.
- Larsson, R. (2020). Discrete factor analysis using a dependent poisson model. *Computational Statistics*, 1–20.
- Ruff, D. A. and M. R. Cohen (2016). Attention increases spike count correlations between visual cortical areas. *Journal of Neuroscience* 36(28), 7523–7534.
- Wedel, M., U. Böckenholt, and W. A. Kamakura (2003). Factor models for multivariate count data. *Journal of Multivariate Analysis* 87(2), 356–369.