1. **Abstract**
2. **Introduction**

Individual neurons in the brain provide limited and noisy information, as such- information is encoded across populations of neurons in the brain (cite). The nature of how this population coding occurs, however, is not fully understood. A primary limitation of neural coding is neuronal noise. Even if an identical stimulus is presented multiple times, the neural response to the stimuli will deviate slightly for each presentation (cite). As a result, the interpretation of neural responses is probabilistic (i.e. the stimulus information encoded in the neural response to said stimulus must be decoded by downstream brain regions in the presence of noise). This calls to question the accuracy of how information can be represented in the brain which depends, in large part, on whether noise is correlated between neurons. As it turns out, neuronal noise is indeed correlated (cite) in a manner which typically limits the information encoded in neural population activity. These correlations can be modulated in the brain, for example attention to a stimulus can reduce noise correlation and improve neural coding. It has been shown that proprioception, the sense of the position of the body in space, can be improved during active movement relative to passive movement of the body, which could be caused by a reduction in noise correlations. In this project, **we address the role of noise correlation in the population coding of active versus passive whole-arm movements in non-human primates.** **We further evaluate whether linear decoding is reasonable proxy for neural representation of whole-arm movement** by comparing the performance of linear and nonlinear models. Previous research has efficacy of linear models in decoding visual categories from neural activity in the inferior temporal cortex (cite). We expect that arm movements will be similarly representable using linear models.

1. **Methods**

**3.1 Data**

We will be using electrophysiological data recorded from non-human primates while they use a manipulandum to reach for targets presented on a screen. There were XX reaching trials per condition. There were 2 conditions, active and passive. In the active condition, the subject reaches for the target on their own, and in the passive condition, the subject’s arm is moved using the manipulandum. Data were recording using 100-electrode arrays into the arm representation of area 2 of S1 somatosensory cortex in these monkeys. Signals were sampled at 30 kHz. Neural spikes were detected online using a threshold set at -5 \* signal RMS. …more info

The dataset also includes the position of the handle using encoders on the manipulandum joints, and the interaction forces between the monkey’s hand and the handle.

**3.2 Preprocessing (offline)**

Neural spike data filtering, noise rejection etc. (For detail methods, see: https://elifesciences.org/articles/48198#s4) . The data was split into tree subsets of randomly selected trials: (1) A training set (XX % of the trials) which was used to train the classification algorithm, (2) a cross-validation set (XX% of the trials) which was used to refine the model via regularization, and (3) a test set (XX% of the data) which was used evaluate the performance of the model on completely unseen data. We used a 5-fold stratified cross validation split.

**3.3 Features**

We used neural spike data from individual neurons to predict the target location of a given arm-movement. We standardized the neural data by removing the mean off all features and scaling to unit variance. This is especially important, as we will be implementing a Support Vector Machine (SVM) classifier. The objective function of the RBF kernel of SVM assume that all features are centered around 0 and have similar variance (cite).

**3.4 Classification**

We use SVM classifiers to predict target location based (from a set of XX discrete targets) using neuronal recordings. SVM is a non-probabilistic binary linear classifier, however, using the kernel trick, SVMs can map inputs into high-dimensional feature spaces to perform no-linear classification (cite). We implement both linear and nonlinear (RBF) kernel functions separately. For both SVMs, we use a search over XX regularization coefficients between XX and XX during 5-fold stratified cross-validation.

**3.5 Noise Correlation Reduction**

In order to eliminate noise correlation from the data, we shuffled trials corresponding to the same target for each neuron.

**3.6 Evaluation**

We compared the classification accuracy before and after trial-shuffling to evaluate the role of noise correlation in the neural representation of whole-arm movements. This comparison was made for the linear model only, and irrespective of movement type (active vs passive).First, we compared the bootstrapped 95% confidence intervals of the stratified 5-fold cross validation accuracies with and without trial-shuffling. We also applied McNemar’s test to a 2x2 contingency table containing the frequencies of correct (and incorrect) classifications before and after noise shuffling (see Figure X). This allows us to compare whether the row and column marginal frequencies are equal.

|  |  |  |
| --- | --- | --- |
|  | **Correct classification (shuffled)** | **Incorrect classification (shuffled)** |
| **Correct classification (unshuffled)** | # Yes/Yes | # Yes/No |
| **Incorrect classification (unshuffled)** | # No/Yes | # No/No |

Figure X: Example of a contingency table comparing the frequency of correct and incorrect classifications for data with and without noise correlation. Each cell reports the frequency value given a specific set of conditions. For example, the top left cell reports the number of trials where both the unshuffled and shuffled conditions result in a correct classification.

We further investigate the role of noise correlation with respect to the type (passive or active) of movement and type (linear or nonlinear) of model. To do this, we performed a three-way ANOVA and subsequent Tukey’s Honestly Significant Difference (HSD) test on the stratified 5-fold cross validation accuracies obtained by constructing XX bootstrapped data samples for each condition (See figure XX)

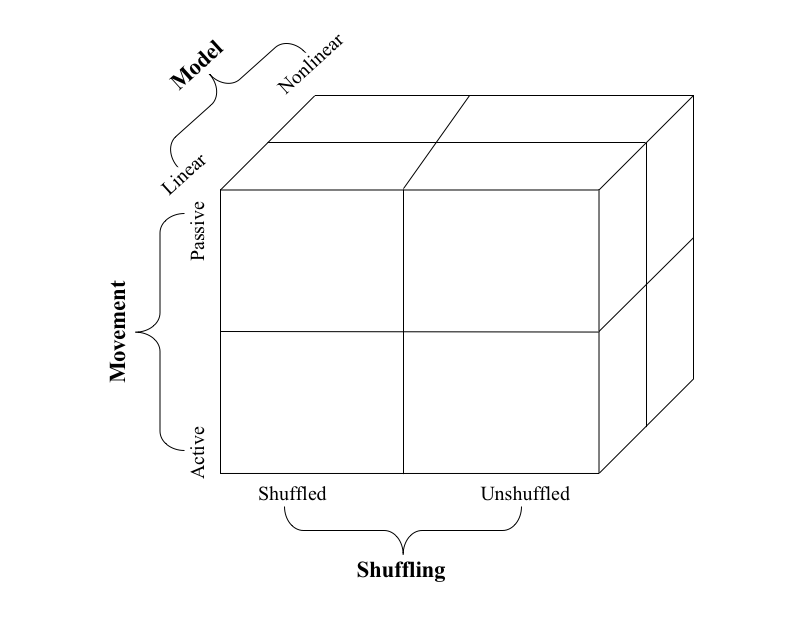


Figure X: Schematic of all comparisons to be evaluated. Each cell represents a group (e.g. Active, Linear, Unshuffled), with a total of 8 groups to compare. Note that the movement condition will be identified post-model selection based on the concatenation of both movement types (i.e. no separate classifier will be trained for active/passive movements only).

We will then eliminate the role of noise correlation by shuffling trials corresponding to the same target for each neuron. Note that several different arm movements can be associated with the same target. We can then evaluate the role of noise correlation in the neural representation of whole-arm movements by comparing the classifiers performance before and after elimination of noise correlation for both the active and passive conditions.

1. **Results**

**4.1 Noise correlation impedes neural representation of arm movement**

We assessed the role of noise correlation in computational modeling of neural representation of whole-arm movement in non-human primates. We found that, irrespective of movement and model type, classification accuracy was improved when noise correlation was eliminated via trial shuffling.

**4.2 Neural representation is improved for active movements**

**4.3 Active limb movements mitigate the impact of noise correlation**

We further assessed the role of movement-intent (i.e. was the movement active or passive) on the accuracy of our computational model. We found that there was an interaction between movement type and shuffling in that shuffling (i.e. removal of noise correlation had a greater effect on accuracy for passive movements.

**4.4 Linear decoding is a proxy for neural representation of arm movements**

We compared the accuracy of a non-linear (or linear TBD) SVM for classifying target location neural spike data. We found that, irrespective of movement type and noise correlation, the XX classifier achieved better accuracies when evaluated on unseen data, suggesting that the XX model is a better proxy of neural representation. However, the (other) classifier still achieved an accuracy of XX on unseen data, which indicates that it performs better than chance, which would be XX %. This implies that, while the XX model may be a better model for neural representation of arm movements, both XX and YY models are reasonable.

Figure X: ANOVA table

Figure X: Tukey’s HSD test

Figure X: Confusion matrices for each model

1. **Discussion**

**5.1 Noise correlation impedes neural representation of arm movement**

**5.2 Active limb movements mitigate the impact of noise correlation**

1. The role of efference copy (movement-producing signal generated by motor system) in active versus passive movements. In passive movements, we may not see the same attenuated processing of self-generated afferents that we should expect in the active condition.
2. Internal mechanisms to deal with noise correlation may be more robust in active condition

**5.3 Linear decoding is a proxy for neural representation of arm movements**

.