

The impact of noise correlation on multivariate pattern classification in fMRI

Ruyuan Zhang (zhan1217@umn.edu), Kendrick Kay (kay@umn.edu)

Center for Magnetic Resonance Research, Department of Radiology

University of Minnesota, Minneapolis, MN 55455 USA

Abstract

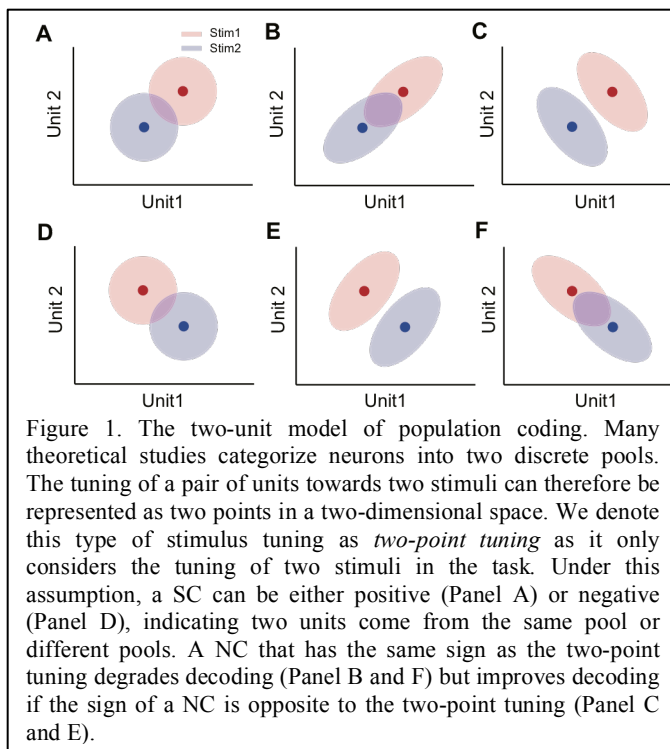
Previous theoretical work has shown that noise correlations (NC) that have the same sign as signal correlations (SC) limit the amount of information encoded in neural population responses. A recent fMRI study replicated this finding in functional magnetic resonance imaging (fMRI) data (Bejjanki et al., 2017). Here, we aim to gain further understanding of how NCs between voxels alter the accuracy of multivariate pattern classification (MVPC), a popular analysis method in fMRI research. In a simulated fMRI orientation experiment, voxel responses were simulated using an encoding model and then classified using linear discriminant analysis. We evaluated two forms of NC: one proportional (i.e., tuning-compatible NC) to the SC between voxels and the other one independent (i.e., tuning-independent NC) of the SC between voxels. Surprisingly, our results show that both the tuning-compatible NCs and the tuning-independent NCs improve MVPC accuracy. We show that these results stem from two major factors: (1) classifiers can “greedily” select voxels that have informative covariance and (2) the SC defined using tuning for two orientations is different from the SC defined using full orientation tuning curves. Taken together, our results provide a theoretical foundation for understanding the effect of NC on MVPC in future experimental studies.

Keywords: noise correlation, multivariate pattern analysis

Noise correlation and neural coding

It has been known that activity of a single neuron depends on other neurons and sometimes two neurons exhibit correlated activity. In computational neuroscience, the correlation between the tuning curves of two units is called *signal correlation* (SC) and the trial-by-trial response correlation is called *noise correlation* (NC). The SC has been extensively studied but the signatures of the NC in neural coding have not been well characterized. Recent multi-unit recording studies and theoretical work have examined to what extent NC will improve or impair population codes. Existing evidence has indicated that the effect of NC depends on many factors, including its behavioral relevance (Haefner et al., 2013), the form of noise, the sign of NC, and the relationship between NC and SC (Averbeck et al., 2006).

In a typical perceptual decision making task, previous theoretical work usually categorizes neurons into two pools whose tuning favor two stimuli respectively (Shadlen et al., 1996). Under this assumption, any pair of neurons reside within the same pool or different pools. Accordingly a NC



can either have the same or opposite sign of the SC. For example, a positive NC between two neurons having a positive SC (i.e., two neurons in the same pool) is considered as a tuning-compatible NC. The tuning-compatible NC has been found to be ubiquitous in the brain (Gawne & Richmond, 1993; van Kan et al., 1985) and detrimental to information encoding (Averbeck et al., 2006). This view is elaborated in Figure 1.

In a different line of research, fMRI provides simultaneous measurement of many voxel responses in the brain, and a popular approach to fMRI is multivariate pattern classification (MVPC). Despite the tremendous success of MVPC in fMRI research, little is known with respect to how NCs alter its performance.

In this paper, we aim to perform a series of theoretical analyses to understand how NCs between voxels in fMRI data will impact the performance of MVPC. We focus on two forms of NC: one proportional to (i.e., tuning-compatible NC) and the other independent of SC (i.e., tuning-independent NC). Using an orientation encoding model, we simulate the trial-by-trial responses of many voxels and use the responses to classify two orientations.

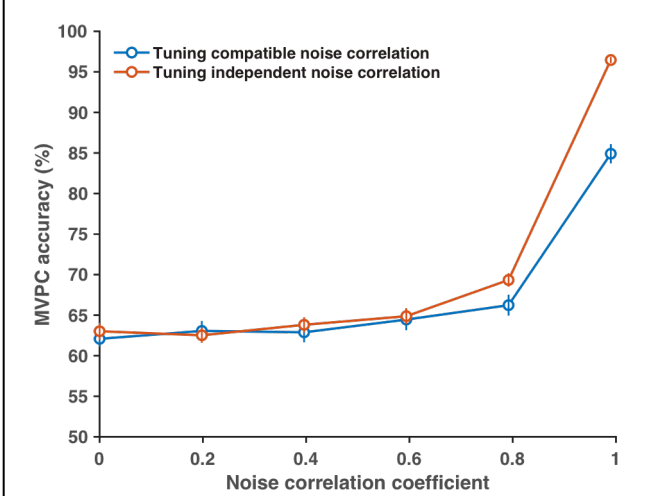


Figure 2. The impact of tuning-compatible and tuning-independent NC on the performance of MVPC. Increasing both forms of NC improves MVPC performance. Error represents the SEM across ten independent simulations.

Simulating noise correlation in fMRI data

The orientation encoding model. All simulations in this paper were performed under a hypothetical experiment: classifying orientations 45° and 135° based on multi-voxel responses. This is reminiscent of early brain decoding investigations (Kamitani & Tong, 2005).

The orientation encoding model assumes 8 neural orientation channels, whose tuning functions are specified as positively half-wave rectified cosine functions raised to the fifth power:

$$g_k(s) = \left[\cos\left(\frac{\pi}{90}(s - \varphi_k)\right) \right]_k^5, \quad (1)$$

where k indicates the k -th channel defined by the preferred orientation φ_k . Preferred orientations of 8 channels are equally spaced between $[0, 180^\circ)$. We further assume that a single voxel response $f_i(s)$ is the linear combination of all orientation channels:

$$f_i(s) = \sum_{k=1}^8 w_{ki} g_k(s), \quad (2)$$

where w_{ki} is the connection weight between the k -th orientation channel to the i -th voxel. \mathbf{W} is given by:

$$w_{ki} \sim \text{uniform}(0, 3), \quad (3)$$

Thus the mean of a group of voxel response can be represented by $\mathbf{f}(s) = [f_i(s)]$. However, empirically measured

voxel responses are inevitably corrupted by noise. Thus trial-by-trial responses of a group of voxels should be:

$$\mathbf{b} = \mathbf{f}(s) + \mathbf{e}, \quad (4)$$

where \mathbf{b} is estimated beta weights in normal general linear model analysis in fMRI and \mathbf{e} represents the multivariate normal distribution:

$$\mathbf{e} \sim N(0, \Sigma), \quad (5)$$

where Σ is the covariance matrix.

Voxel-wise noise correlations. In computational neuroscience, the signal correlation between two units refers to the correlation of their tuning curves:

$$r_{ij}^{\text{signal}} = \text{corr}(f_i(\mathbf{S}), f_j(\mathbf{S})), \quad (6)$$

where \mathbf{S} indicates all possible orientations between $[0, \pi)$ and $f_i(\mathbf{S})$ is the tuning curve of the i -th voxel.

On the other hand, the noise correlation refers to the trial-by-trial covariation between two units and the NC in theory is independent of the SC. Empirical recording studies found that the NC is usually weakly proportional to the SC.

As such, two forms of NC are proposed in this model. In the first case, the NC is directly related to the SC:

$$r_{ij}^{\text{tuning}} = r_{ij}^{\text{signal}}, \quad (7)$$

We denote this type of NC as r_{ij}^{tuning} because it is directly related to two voxels' tuning.

In the second case, we randomly shuffle the voxel index such that the NC is independent of the SC while keeping the overall amount of noise in the population constant:

$$r_{ij}^{\text{random}} = r_{ij}^{\text{signal}}, \quad (8)$$

where \mathbf{z} is a vector of voxel indices with a random order.

Given the correlation matrix, the covariance matrix can be formulated as:

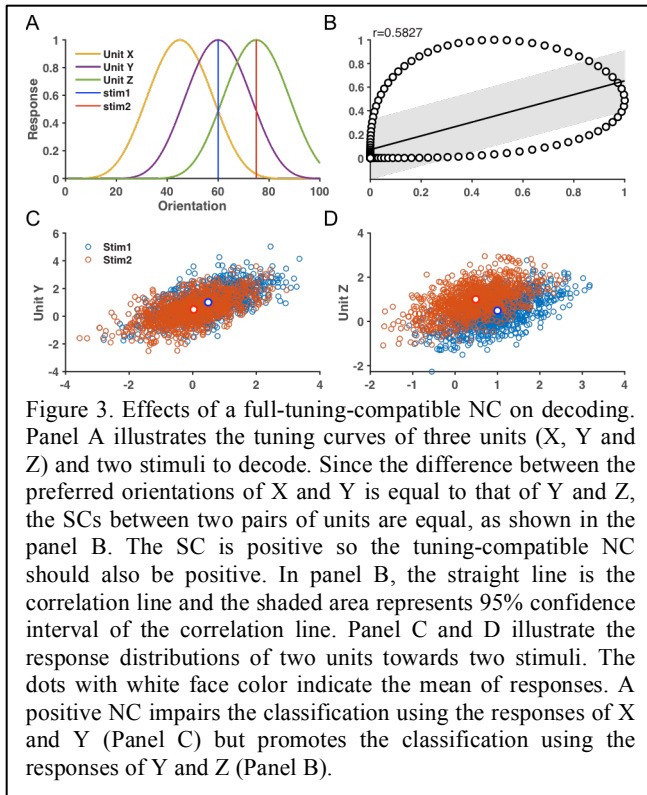
$$\Sigma_{ij} = \begin{cases} \tau_i^2, & i = j \\ c * \tau_i \tau_j r_{ij}, & i \neq j \end{cases}, \quad (9)$$

where τ_i^2 is the trial-by-trial response variance of the i -th voxel and the NC between voxels i and j is r_{ij} , which could be either r_{ij}^{tuning} or r_{ij}^{random} . c is the NC coefficient that controls the strength of the NC. The key question we focus here is how the c value alters MVPC accuracy.

In addition, we assume all τ s follow a normal distribution:

$$\tau \sim N(u, v), \quad (10)$$

Here, μ and v represent the mean and the variability of noise level across voxels respectively. We set $u = 15$, $v = 6$.



Simulations. Six levels of the NC coefficient were set from 0 to 0.99. For each NC coefficient level, we simulated the responses of 30 voxels for a total of 2000 trials (1000 trials for each orientation). A linear discriminant analysis was used to classify orientations based on all 2000 simulated trials. Six NC coefficient levels and two types of NC (r_{ij}^{tuning} and r_{ij}^{random}) constituted a total of twelve conditions. The simulated experiment was repeated 100 times in each condition to estimate the reliability of MVPC accuracy. Note that in each of 100 simulated experiments, the connection weight W was reset and fixed throughout. The NC coefficient but not the weight W varied such that endowed results at different NC coefficients were directly comparable.

Noise correlation improves MVPC accuracy

We examined how the two forms of NC alter MVPC accuracy. We noticed that increasing the amount of a tuning-compatible NC improves the MVPC accuracy (Figure 2). This result contradicts conventional thinking and a recent fMRI study showing that the tuning-compatible NC is usually detrimental to information decoding (Bejjanki et al., 2017). We also observed that a similar result occurred if we increased the amount of tuning-independent NC. This result demonstrates that even randomly injecting some covariance in fMRI data will improve the performance of pattern classification.

Why do voxel-wise NCs improve MVPC accuracy, regardless of whether the NCs are related to the voxel tuning or not? We identified two major factors contributing

to the results and will elaborate them in next two sections.

The distinction between the tuning for two stimuli and the tuning curves for all stimuli

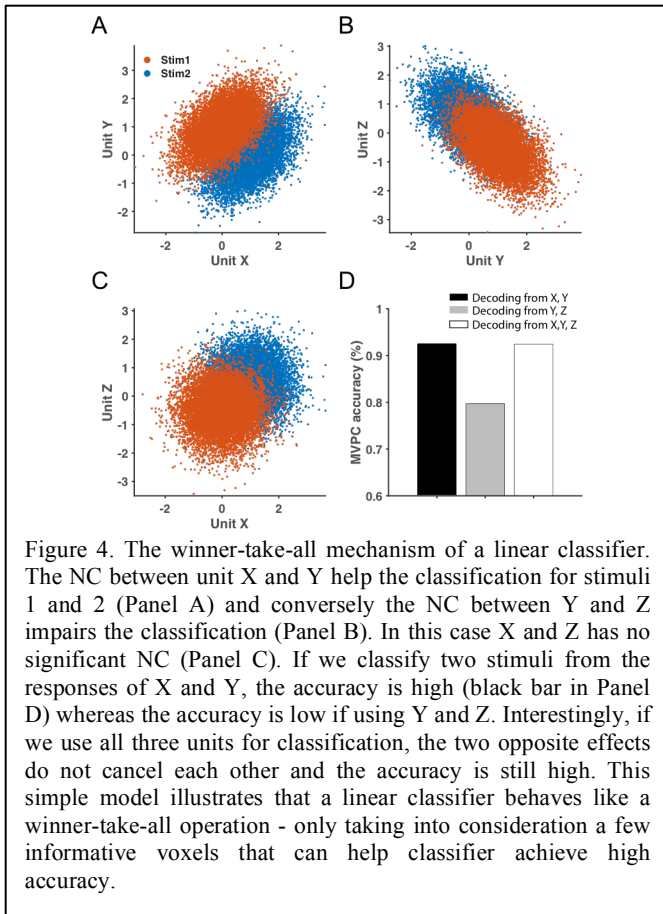
Why have previous studies demonstrated that a tuning-compatible NC is detrimental for information decoding whereas here we show it is beneficial? We call attention to the difference between the neuronal responses to two stimuli and the full response tuning curves across all possible stimuli. In general, a tuning-compatible NC means having the same sign of a SC. However, a SC can be two distinct cases, as we will show below, and therefore a tuning-compatible NC should be also considered under two circumstances..

Many previous studies categorize neurons into two pools that selectively respond to two stimuli. Any pair of neurons can be either in the same pool (i.e., positive SCs) or different pools (i.e., negative SCs). We denote the SC in this case as *two-point SC* and the NC that has the same sign of this SC as *two-point-compatible NC*. Figure 1 depicts how a NC and a SC in the two-point case jointly determine classification performance. Note that under this circumstance, a two-point-compatible NC indeed worsens classification.

However, the SC in our study is defined as the correlation between the full tuning curves of two units, not just the selective responses for two stimuli. We denote this type of SC as *full-tuning SC* and the NC that has the same sign of this SC as *full-tuning-compatible NC*. Note that these two types of NC are dissociable and a full-tuning-compatible NC, unlike a two-point-compatible NC, can either help or hamper MVPC. We use a three-unit model to depict this (Figure 3). Units X and Y have very similar tuning curves (i.e., a positive full-tuning SC, Figure 3A-B) and their full-tuning-compatible NC therefore should also be positive. The same holds for units Y and Z. If we attempt to classify two stimuli based on responses of X and Y, a full-tuning-compatible NC is in line with the two-point-compatible NC and thus worsens the MVPC accuracy (Figure 3C). However, if we attempt to classify two stimuli based on responses of Y and Z, the response geometry is flipped and a positive NC turns to be beneficial in this case (Figure 3D). Thus, given the same full-tuning-compatible NC, it can be either consistent with or opposite to a two-point-compatible NC and thus either improves or impairs classification.

Previous observations (Bejjanki et al.) that a tuning-compatible NC is detrimental only holds true with respect to the two-point-compatible NC; while this case is tractable and easy to understand, it is overly simplified since neurons usually continuously span a whole feature space. The full tuning-compatible NC is a more realistic characterization for the case where discrete stimuli are classified from the responses of a population with diverse tuning.

A “winner-take-all” mechanism of MVPC



The second factor that contributes to the observed results is the winner-take-all operation of a linear discriminant. Again, we run a simple three-unit model to illustrate this (Figure 4). In the model, units X and Y have a covariance beneficial to classification, whereas the covariance of units Y and Z is detrimental. If the classification is based on all three units, two opposite effects do not “cancel” each other. Rather, the performance of classification remains high, indicating that the classifier behaves according to a “MAX” or “winner-take-all” rule. In other words, the covariance structures of a small group of “good” voxels will determine the overall accuracy of a linear classifier that can characterize and make use of noise covariance structure. Thus, as long as a NC creates such structure in a subset of voxels, an appropriately designed linear classifier can utilize these voxels to achieve better performance.

Conclusion

We simulated a fMRI experiment and classified two orientations based on simulated responses of a population voxels. We found that adding noise correlations to the population, no matter compatible with signal correlations between voxels or not, improves classification accuracy. This result at first glance contradicts with previous work showing that tuning compatible NC is detrimental for information decoding. We show that these results mainly

stem from two major factors: (1) the distinct definitions of tuning for two stimuli and the tuning curves for all stimuli; (2) the greedy selection of beneficial covariance by a linear discriminant. We contend that NCs in a large population are almost guaranteed to be helpful since as long as there exists at least one pair of units for which NC is beneficial, a linear discriminant can exploit these units to improve classification accuracy. Note that these two factors not only manifest in the simulation here but also are broadly true when considering other neuroscience data (i.e., multi-unit neural recordings). Our analyses thus provide insight into the nature of MVPC, as well as general principles of neural coding.

References

- Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nat Rev Neurosci*, 7(5), 358-366.
- Bejjanki, V. R., da Silveira, R. A., Cohen, J. D., & Turk-Browne, N. B. (2017). Noise correlations in the human brain and their impact on pattern classification. *PLoS Comput Biol*, 13(8), e1005674.
- Gawne, T. J., & Richmond, B. J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? *The Journal of Neuroscience*, 13(7), 2758-2771.
- Haefner, R. M., Gerwinn, S., Macke, J. H., & Bethge, M. (2013). Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat Neurosci*, 16(2), 235-242.
- Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nat Neurosci*, 8(5), 679-685.
- Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(4), 1486-1510.
- van Kan, P. L., Scobey, R. P., & Gabor, A. J. (1985). Response covariance in cat visual cortex. *Exp Brain Res*, 60(3), 559-563.