

Noise correlations facilitate faster learning

Matthew R Nassar (matthew_nassar@brown.edu)

Neuroscience Department, 164 Angell Street
Providence, RI, 02906, USA

Abstract:

One major challenge for AI is that, while deep neural networks are capable of achieving human level performance on a wide variety of tasks, they typically require a greater number of learning trials than would be required by a human. This issue has stimulated an interest in the inductive biases that humans and other animals employ to constrain learning in complex natural environments. While the neural mechanisms used to implement inductive biases could be informative for both improving AI and providing a better mechanistic understanding of learning, these neural underpinnings remain elusive. Here I explore the possibility that stimulus-independent pairwise correlations between neurons, or so-called noise correlations, might reflect inductive biases used to constrain learning to specific task-relevant dimensions. I test this idea with a neural network model of a two-alternative forced-choice perceptual discrimination task in which the correlation among similarly tuned units can be manipulated independently of the overall population signal-to-noise ratio. Higher noise correlations among similarly tuned units led to faster learning through weight adjustments that favored homogenous weights assigned to neurons within a functionally similar pool. Such noise correlations emerge naturally with Hebbian learning. These results suggest that noise correlations may serve to reduce the dimensionality of learning thereby making it more rapid and robust.

Keywords: noise correlations; inductive biases; learning

Introduction

The brain represents information using distributed population codes in which particular feature values are encoded by large numbers of neurons. A theoretical advantage of distributed population codes is that a pooled readout across many neurons can effectively reduce the consequences of stimulus-independent variability (noise) in the firing of individual neurons. However, the degree to which this benefit can be employed in practice is limited by noise correlations, or the degree to which stimulus-independent variability is shared, particularly across the subset of neurons that encode a particular stimulus feature (Averbeck, Latham, & Pouget, 2006; Cohen & Kohn, 2011). In particular, positive noise correlations between neurons that share the same stimulus tuning reduce the amount of decodable information in the neural

population. Nonetheless, this type of noise correlation is reliably observed, particularly between pairs of neurons that provide evidence for the same choice or perceptual categorization (Cohen & Newsome, 2008), raising the question of why noise might be distributed in this task and tuning specific manner.

One possible explanation might be that such correlations serve a purpose for learning, namely to reduce the effective dimensionality of learning. Such an explanation would be consistent with computational analyses of Hebbian learning rules (Oja, 1982), which can both facilitate faster and more robust learning, and in turn may induce noise correlations. Perceptual learning studies support the notion that learning to readout available sensory information might be as large a challenge for the brain as representing the information in the first place; indeed effective readout of visual motion information can take years of training in monkeys (Law & Gold, 2008; 2009).

Here I explore this possibility using a simplified neural network model of a two alternative forced choice perceptual discrimination task in which the correlation among similarly tuned neurons can be manipulated independently of the overall population signal-to-noise ratio. Within this framework, noise correlations speed learning by forcing learned weights to be similar across pools of similarly tuned neurons, thereby ensuring learning occurs over the most task relevant dimension. Noise correlations can also be learned in the basic network architecture through Hebbian mechanisms. These results provide a first proof of concept for the notion that noise correlations could serve to control the dimensions over which learning occurs.

Methods

All simulations and analyses were performed using simplified and statistically tractable two-layer neural network. The input layer consisted of two pools of 100 neurons that were each “tuned” to one of two stimuli. On each trial normalized firing rates for the neural population were drawn from a multivariate normal



distribution that was specified by a vector of stimulus-dependent mean firing rates (signal: +1 for preferred stimulus, -1 for non-preferred stimulus) and a covariance matrix. All elements of the covariance matrix corresponding to covariance between units that were “tuned” to different stimuli were set to zero. The key manipulation was to systematically vary the magnitude of diagonal covariance components (eg. noise in the firing of individual units) and the “in pool” covariance elements (eg. shared noise across similarly tuned neurons) while maintaining a fixed level of variance in the summed population response for each pool:

$$\sigma_{pool}^2 = n\sigma_{unit}^2 + n(n-1)Cov(\text{within pool})$$

Where σ_{pool}^2 is the variance on the sum of normalized firing rates from neurons within a given pool, n is the number of units in the pool and the within pool covariance ($Cov(\text{within pool})$) specifies the covariance of pairs of units belonging to the same pool. The signal to noise ratio for the population response was fixed to one for all simulations presented here. Given this constraint, the fraction of the total population noise that was shared across neurons was manipulated as follows:

$$\sigma_{unit}^2 = \frac{\sigma_{pool}^2}{n + n(n-1)\phi}$$

$$Cov(\text{within pool}) = \phi\sigma_{pool}^2$$

Where ϕ reflects the fraction of noise that is correlated across units, and was set to values ranging from 0 to 0.2 for this set of simulations. The output layer contained one unit for each pool in the input layer, and was fully connected to the input units in a feedforward manner. Output units were activated on a given trial according to weighted function of their inputs:

$$F_j = \sum_i F_i w_{i,j}$$

Actions were selected as a softmax function of output firing rates:

$$p(A_j) = \frac{e^{\beta F_j}}{\sum_k e^{\beta F_k}}$$

where β is an inverse temperature, which was held constant for all simulations. Learning was implemented through reinforcement learning:

$$\Delta w_{i,j} = \alpha \delta F_i$$

Where F_i is the normalized firing rate of the i th neuron, δ is the reward prediction error experienced on a given trial [+0.5 for correct trials and -0.5 for error trials], and α is a learning rate (held constant at 0.0001 for all simulations). The network was trained to correctly identify two stimuli (each of which was preferred by a single pool of input neurons) over 100 trials (the last 20 trials of which were considered testing). Simulations were repeated 1000 times for each level of ϕ and performance measures were averaged across all repetitions.

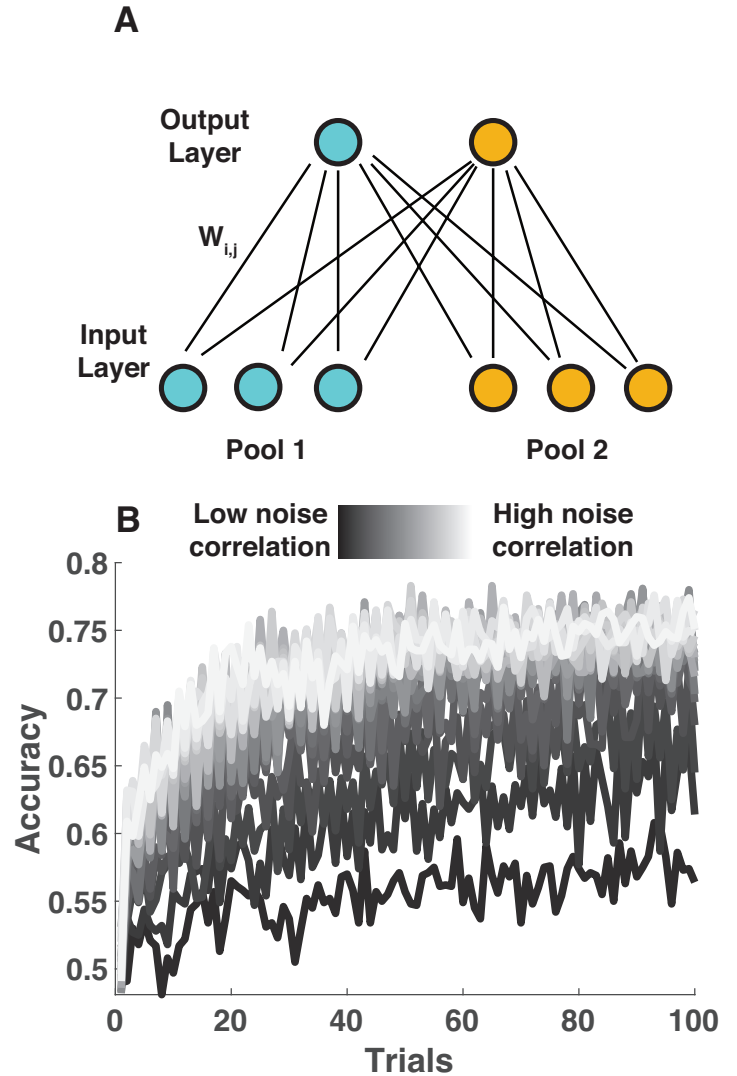


Figure 1: Noise correlations can speed learning. A) Schematic of the two layer neural network used in simulations. Network contained two pools of input units that responded preferentially to different stimulus categories. The primary manipulation was to control the degree to which response variability was correlated across neurons of a given input pool while maintaining a fixed signal-to-noise ratio at the population level (equal to 1 in this example). Output unit responses were a weighted sum of input layer firing. Weights were learned through reinforcement learning. B) Networks in which input layer variability was correlated across units within the same pool (lighter colors) tended to learn the task more rapidly.

Results

Two-layer feed-forward neural networks were trained and tested on a perceptual categorization task that required identification of a stimulus that was encoded by increased activity in one of two pools of units (Fig 1a). Firing rates of individual input units were variable on each trial, and the degree to which this stimulus-independent variability was shared across units in the same pool was manipulated while holding signal-to-noise ratio of the population response constant. All networks learned to perform the categorization task (Fig 1b), however the networks in which noise was more highly correlated across units in the same pool tended to do so more rapidly (Fig 1b; lighter colors). Mean performance across the last twenty trials revealed a positive relationship between noise correlations and performance up to noise correlations of 0.2 (Fig 2a; red curve). Models with higher “in pool” noise correlations approached the theoretically achievable performance level (Fig 2a; blue line) even after only 80 training trials. The faster learning in the high noise correlation conditions was made possible by constraining the degree to which weights associated with units in the same pool could diverge from one another (Fig 2b, compare light and dark weight profiles for pool 1/2 units [left/right]).

Additional simulations revealed that the same basic qualitative results were obtained when 1) learning was supervised rather through reinforcement, 2) more than two pools of units were included, 3) signal-to-noise ratio was varied [0.5 – 2], 4) the size of the pools were adjusted [10-1000 units]. Furthermore, extending the network to include a fully connected intermediate layer that learned connectivity weights to the input layer through a Hebbian mechanism was capable of reproducing beneficial noise correlations.

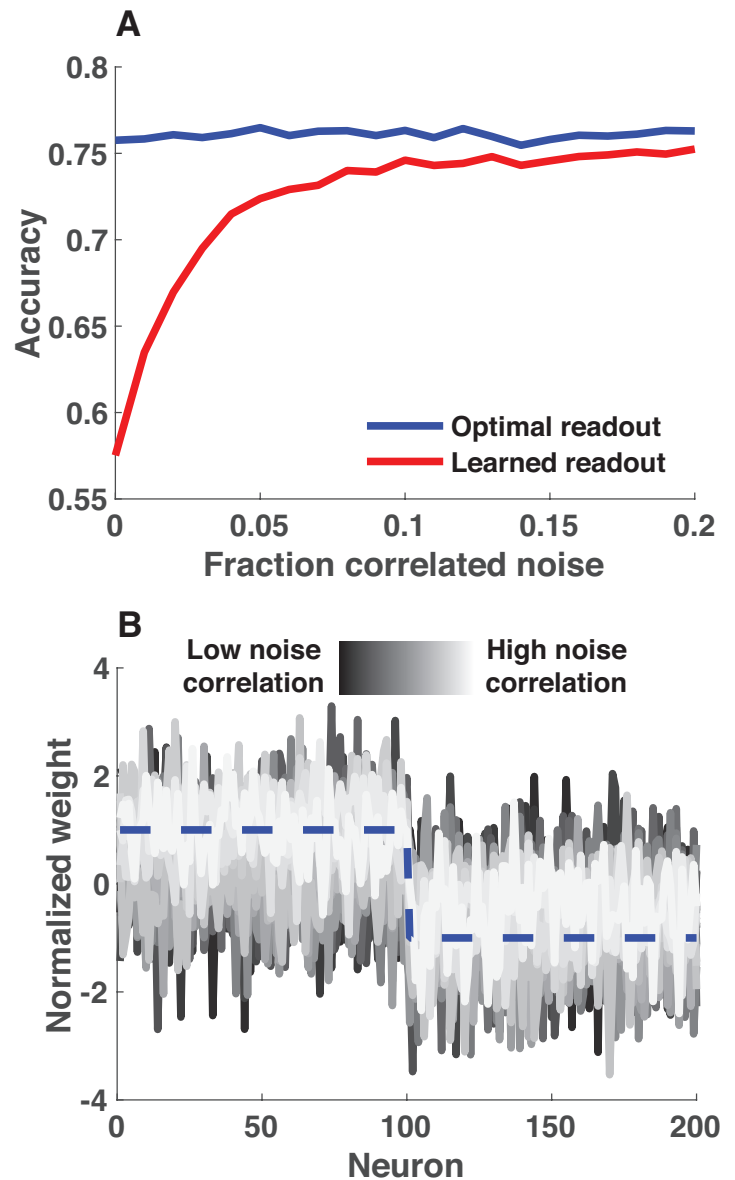


Figure 2) Noise correlations can improve learning by targeting learning to relevant dimensions A) Optimal readout of input layer in all networks yielded a similar level of performance (blue) but learned weights yielded better performance (red) for by models that had higher noise correlations (abscissa). B) learning was faster because correlated networks (lighter colors) tended to learn similar weights (ordinate) for all units within a given pool (abscissa), whereas less correlated networks (darker colors) tended to learn more variable weight profiles.

Discussion

Positive noise correlations between similarly tuned neurons are typically observed in paired recordings (Cohen & Newsome, 2008) despite their apparent deleterious effects on the quantity of decodable information in neural populations (Averbeck et al., 2006). One possible reason for this is that such noise correlations can speed learning by effectively reducing the dimensionality of the learning problem. In the very simplified set of simulations presented here, learning the appropriate weights for independently firing units requires estimating one parameter for each neuron in the population and in the presence of noise this leads to over-fitting (high variance on weight profile in 2b). In contrast, increased noise correlations force learning to ascribe similar weights to all neurons within the same pool, which, in the extreme as noise correlations approach 1, is analogous to estimating one parameter per pool. Thus, while the simulations presented here are by necessity an oversimplification of the processes implemented in the brain, they suggest that noise correlations might help constrain learning to meaningful dimensions. If this is the case, then a closer look at noise correlation profiles in the brain might shed light on the sorts of inductive biases that guide learning in humans and animals.

Acknowledgments

This work was supported by NIA K99/R00 Pathway to Independence Award: K99AG054732. The author would like to thank Michael Frank, Apoorva Bhandari, and Drew Linsley for helpful discussion.

References

- Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5), 358–366. <http://doi.org/10.1007/978-3-662-02728-8>
- Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Publishing Group*, 14(7), 811–819. <http://doi.org/10.1038/nn.2842>
- Cohen, M. R., & Newsome, W. T. (2008). Context-Dependent Changes in Functional Circuitry in Visual Area MT. *Neuron*, 60(1), 162–173. <http://doi.org/10.1016/j.neuron.2008.08.007>
- Law, C.-T., & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience*, 11(4), 505–513. <http://doi.org/10.1038/nn2070>
- Law, C.-T., & Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, 12(5), 655–663. <http://doi.org/10.1038/nn.2304>
- Oja, E. (1982). Simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 1–7.