Brief note on the impact of redundancy on stable decoding

M. Rule

January 16th 2013

Neural activity is redundant: many states in motor cortex can generate similar movements. When we record from motor cortex, we capture only a small fraction of the total neurons. Redundancy makes it possible to observe the overall state of motor cortex from limited observations, but might also impair the generalization performance of a linear decoder.

Consider two neurons, A and B, that combine linearly to produce movement $C{=}\alpha_1 A{+}\alpha_2 B$. (Perhaps both neurons drive the same targets in spinal cord.) An animal could use any linear combination of activations of units A and B to perform behavior C, so long as the sum $\alpha_1{+}\alpha_2$ is constant. What if there is an unobserved variable γ that sets whether neuron A or B is used more (Fig. 1)?

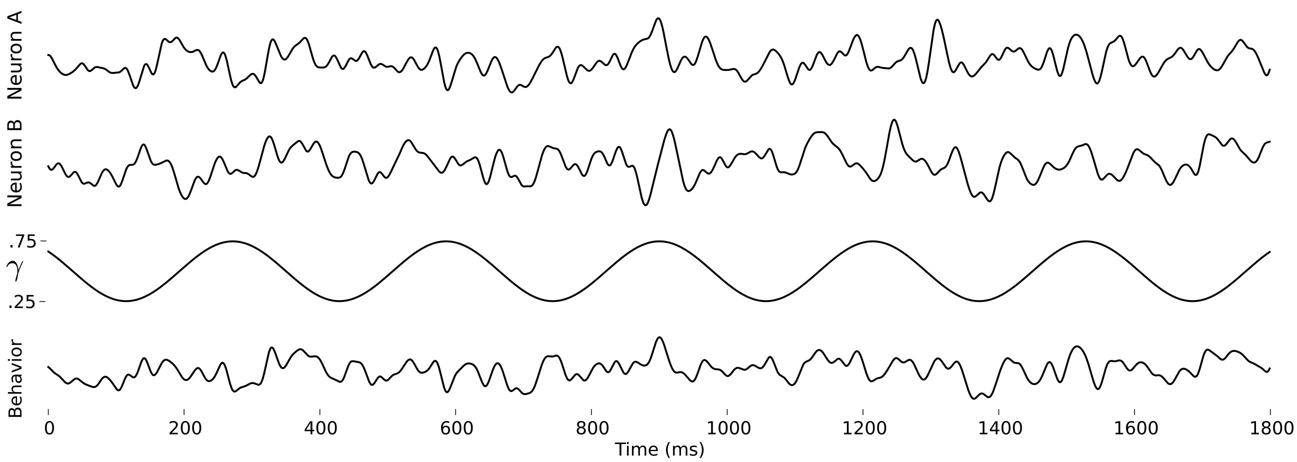

Figure 1: (simulated hypothetical scenario) Neural signals A and B combine linearly according to weight γ to form behavioral output $C = \gamma A + (1 - \gamma)B$. Parameter γ modulates sinusoidally between 0.25 and 0.75.

Let's say we record only from neuron A. Building a linear decoder $\hat{C} = \alpha A$ leads to an over-fit (and erroneous) estimate of the contribution of A to behavior: $\hat{\alpha} = (A \cdot C)/(A \cdot A) \approx 0.996$. When predicting behavior from A, the reconstruction error varies depending on the unobserved slow variable γ (figure 2). This error resembles transient noise, or perhaps an independent source of neuronal variability. But, the activation of A and B always drives behavior in a predictable way. Hidden sources of variability, and under-sampling of the neural population, leads to apparent instability when there is none (Fig. 2).

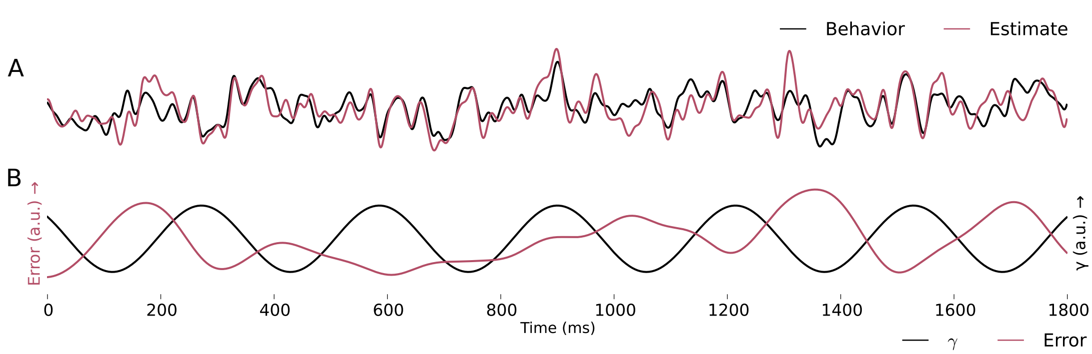

Figure 2: (simulated hypothetical scenario) **(A)** Reconstructed behavior using only unit A leads to unstable decoding accuracy. **(B)** The the smoothed (Gaussian kernel $\sigma = 60$ ms) absolute reconstruction error varies with this hidden parameter γ, which sets A's contribution to the motor output.