A motor cortical model of brain-machine interface learning, fast and slow

An extraordinary property of mammalian motor systems is their capacity to flexibly and quickly adapt to novel environments. A remarkable demonstration of this is primates' ability to acquire proficient control of a brain-machine interface (BMI). What are the algorithms underlying this learning process? An important clue is offered by the finding that primates can learn to use certain motor cortical (M1) BMI decoders with just a few hundred trials of practice (Sadtler et al. '14, *Nature*), while other decoders require many thousands (Oby et al. '19, *PNAS*). Here, we sought to build a mechanistic model of the changes underlying these two timescales of learning. We argue that the short timescales of learning are consistent with a "re-aiming" learning strategy whereby upstream M1 inputs are modified within a task-relevant low-dimensional space, while the local M1 circuit is left unchanged. The low dimensionality of the search space can allow for efficient optimization, but also imposes constraints on the BMI decoders that can be learned. By explicitly modelling "re-aiming" solutions, we show that these constraints are consistent with the results of Sadtler et al. We additionally find that the population activity generated by these solutions maintains its distribution after learning, as is observed empirically (Golub et al. '18, *Nat Neuro*). The model also makes a novel experimental prediction: an asymmetry in the set of achievable BMI velocities, such that BMI reaching ability should depend strongly on the direction of the reach. This asymmetry follows from the nonlinear dynamics inherent in neural circuits. Upon re-analysis of the data of Sadtler et al., we indeed find that this prediction holds. Finally, we show that a similar mechanism is consistent with longer timescales of learning, whereby learning comprises optimizing upstream M1 inputs within a higher-dimensional space encompassing the full space of natural motor commands.

Additional Detail: We consider a simple rate-based recurrent network model of an M1 circuit,

$$\tau\dot{\mathbf{x}} = -\mathbf{x} + \mathbf{W}^{rec}\mathbf{r} + \mathbf{W}^{in}\mathbf{u}, \quad \mathbf{r} = \phi(\mathbf{x}) \tag{1}$$

where $\mathbf{r} \in \mathbb{R}^N$ denotes the firing rates of the N neurons in the circuit and $\mathbf{u} \in \mathbb{R}^M$ denotes the activity of a set of M upstream neurons driving it. The task faced by the subject is to control a cursor on a screen, the velocity of which, $\mathbf{v}(t) = \mathbf{D}\mathbf{r}(t) + \mathbf{b}$, depends linearly on the firing rates of a subset of 100 neurons in the circuit ($\mathbf{D}$ has only 100 non-zero columns). What components of the circuit are modified during learning to solve this task?

A classical answer to this question is that, through synaptic plasticity, the connectivity of the local circuit, $\mathbf{W}^{rec}$ and $\mathbf{W}^{in}$, is optimized for the task (Legenstein et al. '10, Wärnberg & Kumar '19). But note that, because the BMI decoder $\mathbf{D}$ is not explicitly known by the subject, computing gradients of these parameters with respect to task performance is impossible, implying a gradient-free optimization algorithm must be used. Such algorithms are known to perform poorly in high dimensions (Werfel et al. '04, *NeurIPS*). This is inconsistent with the short timescales of learning observed in Sadtler et al. when one considers that, in the primate brain, the relevant motor cortical circuit likely has $N, M > 10^5$, implying that learning in this case would occur in a space with dimension greater than 10^{10}.

Motivated by these arguments, we propose an alternative hypothesis: learning optimizes the upstream inputs, $\mathbf{u}$, within a low-dimensional space. Reflecting on the experimental task used by Sadtler et al. (and by most BMI learning studies), this is a very natural strategy: prior to controlling the BMI, the primates were subjected to a "calibration task" in which they passively viewed 2D cursor reaches to targets on a circle; these 2D stimuli would have evoked 2D sensory inputs, $\mathbf{u}$. We postulate that, during subsequent BMI control, learning proceeds by optimizing inputs within this same low-dimensional space of inputs – a learning strategy we call *re-aiming*.

We model this strategy by letting the upstream inputs depend on the observed cursor velocities $\boldsymbol{\theta} \in \mathbb{R}^2$ during each trial of the calibration task,

$$\mathbf{u} = \phi(\mathbf{P}\boldsymbol{\theta}), \quad \mathbf{P} \in \mathbb{R}^{M\times 2} \tag{2}$$

This equation implies that the inputs $\mathbf{u}$ live within a two-dimensional manifold embedded in M-dimensional space. We can then ask: given a BMI decoder, what inputs *within this manifold* can drive the M1 population to produce the BMI velocities required by the task? Mathematically, this corresponds to optimizing the variable $\boldsymbol{\theta}$ in equation 2 w.r.t. task performance; that is, optimizing the inputs within this 2D manifold.

We first turn to the question of what BMI decoders admit such *re-aiming* solutions. A critical control-theoretic insight is that constraining the space of inputs to the circuit correspondingly constrains the set of dimensions reachable by the population activity. Thus, *re-aiming* will only work for BMI decoders "aligned" with these

dimensions. Indeed, in the case of linear dynamics ($\phi(\mathbf{x}) = \mathbf{x}$), we can rigorously prove that good *re-aiming* solutions exist only when the BMI decoder satisfies a precise criterion of alignment with the network dynamics. In nonlinear networks of rectified-linear neurons, these analytical results are qualitatively replicated in simulations.

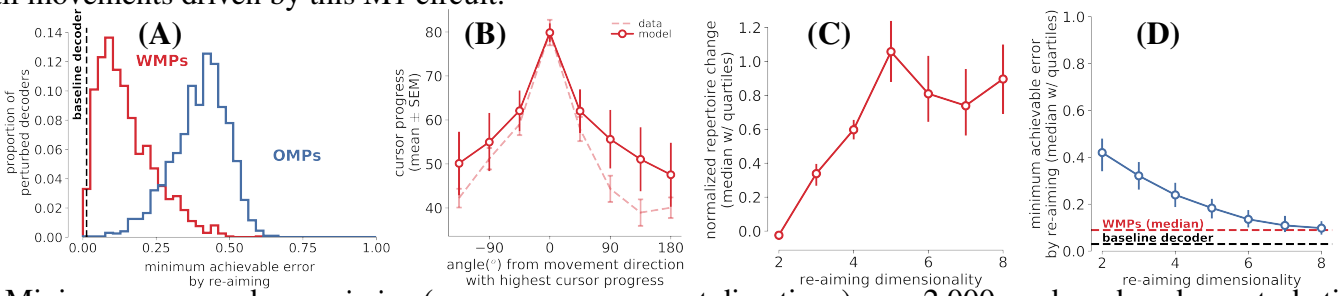
We further find that this criterion is consistent with the results of Sadtler et al. In their experiment, they fit a *baseline decoder* to the calibration task population activity and then considered two classes of decoder perturbations: *within-manifold perturbations* (WMPs) that preserved the decoder's "alignment" with these data, and *outside-manifold perturbations* (OMPs) that did not. After monkeys practiced using the baseline decoder, the authors observed that WMPs but not OMPs could be learned on short timescales. In agreement with these findings, we find that WMPs but not OMPs admit good *re-aiming* solutions (fig A), suggesting that short timescale learning may indeed involve something like *re-aiming*.

In nonlinear networks, this model of *re-aiming* also makes a novel experimental prediction: the dynamical asymmetry imposed by the non-negative firing rates $\mathbf{r}(t)$ induces an asymmetry in the set of BMI velocities $\mathbf{v}(t)$ achievable via *re-aiming* (fig B). Thus, if short timescale learning involves *re-aiming*, then performance improvements on these timescales should depend on movement direction. Re-analyzing the data of Sadtler et al., we find that this prediction holds (fig B), although the falloff in performance is steeper than the model prediction. Our model also predicts that the best movement directions should be predictable from the calibration task data.

To better understand the neural implications of *re-aiming*, we compared the population activity generated by the *re-aiming* solutions for the baseline decoder and for the within-manifold perturbations. This comparison allowed us to make predictions about how the population activity should change after learning a WMP by *re-aiming*. As observed previously for WMPs (Golub et al. '18), we find that the distributions of population activity underlying control of these decoders highly overlap (fig C, leftmost point), a phenomenon termed "neural reassociation". In other words, the distribution of population activity is highly conserved after learning by *re-aiming*.

Importantly, we find that this only occurs in our model when the inputs are constrained to two dimensions. If we allow inputs to be optimized over three or four dimensions rather than two (by augmenting $\mathbf{P}$ to be $M \times K$), we find that the resulting higher-dimensional *re-aiming* solutions for WMPs evoke novel activity patterns differing substantially from those evoked by the *re-aiming* solutions for the baseline decoder (fig C). This suggests that "neural reassociation" does not directly follow from assuming the motor cortical circuit remains unchanged – it is critical that the inputs be constrained to a low-dimensional space as well.

Note that generating novel activity patterns is exactly what is required to control OMPs (Oby et al.). We might then ask: do high-dimensional *re-aiming* solutions exist for these decoder perturbations? We find that, in random networks, the answer is yes, with high probability (fig D). This suggests that the reason these decoders cannot be learned on short timescales is not because *re-aiming* is not viable, but because *re-aiming* in this case requires optimization in a higher dimensional space. We speculate that this higher dimensional space constitutes the set of all upstream inputs used during natural motor control, with dimensionality roughly matching that of the space of of all movements driven by this M1 circuit.



(A) Minimum error under *re-aiming* (avg across movement directions) over 2,000 random decoder perturbations. Error normalized to that achieved under no input. **(B)** Dependence of "cursor progress" (as defined by Golub et al. '18) on movement direction for WMPs. Model results are re-scaled to match the peaks. **(C)** Difference in population activity generated by baseline decoder and WMP *re-aiming* solutions, quantified using the metric of Golub et al. '18. **(D)** Optimal error for OMPs as a function of input dimensionality used for *re-aiming* (controlling for input norm). Horizontal lines show optimal error under 2D *re-aiming*. All RNN simulations used $N{=}M{=}2000$, $\mathbf{W}^{rec}, \mathbf{W}^{in}, \mathbf{P}$ random Gaussian w/ normalized variance, $\tau{=}200$ms, and rectified linear nonlinearity $\phi(\cdot)$.