A Theory of Sensorimotor Learning for Brain-Machine Interface Control

Jorge Aurelio Menéndez

14th March, 2021

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London

Gatsby Computational Neuroscience Unit

Centre for Computation, Mathematics and Physics in the Life Sciences and

Experimental Biology

University College London

I, Jorge Aurelio Menéndez confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Jorge Aurelio Menéndez

14th March, 2021

Abstract

A remarkable demonstration of the flexibility of mammalian motor systems is primates' ability to learn to control brain-machine interfaces (BMI's). This constitutes a completely novel and artificial form of motor behavior, yet primates are capable of learning to control BMI's under a wide range of conditions. BMI's with carefully calibrated decoders, for example, can be learned with only minutes to hours of practice. With a few weeks of practice, even BMI's with random decoders can be learned. What are the biological substrates of this learning process? This thesis proposes a simple theory of the computational principles underlying BMI learning. Through comprehensive numerical and formal analysis, we demonstrate that this theory can provide a unifying explanation for various disparate phenomena observed during BMI learning in three different BMI learning tasks. By explicitly modeling the underlying neural circuitry, the theory reveals an interpretation of these phenoema in terms of the biological non-linear dynamics of neural circuits.

Impact Statement

The mammalian brain is an astonishingly flexible and adaptive organ. An extraordinary demonstration of this is offered by brain-machine interfaces, in which an external device is artificially controlled by activity in a neural population. Using such a device constitutes a completely novel and artificial form of "behavior", yet primates are able to learn to use brain-machine interfaces under a wide range of conditions. This thesis proposes a theory of the computations underlying this remarkable learning process.

From a practical perspective, brain-machine interfaces have several real-world uses. First and foremost, they hold great promise for restoring movement to individuals with movement disorders or injuries to the limbs or spinal cord. But they also have potential for amplifying and augmenting the way healthy individuals interact with digital devices. Many of the insights gleaned from the theory developed in this work may prove useful for improving the design of brain-machine interfaces with these goals in mind. The theory moreover provides a general framework for modeling how users' brains interact with brain-machine interfaces during learning.

From a more academic perspective, this work puts forward a re-interpretation of various pheneomena observed during BMI learning, providing a unifying explanation for these previously unrelated observations. Moreover, by explicitly modeling the underlying neural circuitry, the theory permits an understanding of these phenomena in terms of biological principles. This relies on the simulation and analysis of a class of non-linear dynamical systems called recurrent neural networks. In the elaboration of our theory, we thus develop a number of novel analytical and numerical tools for analyzing this class of models, which could prove useful for future theoretical research.

Acknowledgments

First and foremost, I have to acknowledge my incredible colleagues at the Gatsby Computational Neuroscience Unit and Sainsbury Wellcome Centre. It is this group of people that is most responsible for my development as a person and as a scientist over the course of my doctoral studies. In particular, special thanks go to my office-mates Lea Duncker, Michael Arbel, Kirsty McNaught, and Ilyes Khemakhem, both for putting up with endless linear algebra questions and for our numerous conversations about machine learning, neuroscience, life, and nothing (in reverse order), which helped inspire and push my research forward.

Of course, none of this research would have evolved into a thesis without the constant presence and advice of my supervisor Peter Latham. His sharp eye and sharp wit were there week in and week out throughout my PhD to make sure my research made any sense. But while I may owe the depth of my scientific training to Peter's supervision, to a large degree I owe its breadth to the special leadership of Maneesh Sahani and Peter Dayan. Through research talks, journal clubs, and, of course, teas, the Gatsby houses a special research environment that allowed me to organically grow into the scientist that I am now.

Going beyond Gatsby, I have to first thank my (now) wife Alex for the constant support over the past five years. This thesis would not have been possible without her loving support and her special ability to take my mind off work in the blink of an eye. I am also infinitely indebted to my parents, who have supported and inspired me in my academic pursuits ever since I can remember. It's hard to argue that my dad's perceptive rationality as the complement to my mom's mystic intuitions for human nature were not primary drivers in leading me to study psychology and neuroscience.

Finally, I find it impossible not to briefly mention past mentors who, whether they know it or not, are to a significant degree indirectly responsible for the work behind this thesis. My first scientific mentors Jonathan Flombaum, Gi Yeul Bae, and Colin Wilson were the ones that put me on track to pursue a PhD after my undergrad. It took only a few months of working with them in the summer after my first year of university to immediately realize that a career in science was right for me. A special thanks also goes to Paul Smolensky, whose thought-provoking courses lead me to shift from doing research in cognitive psychology to research in the neural mechanisms underlying the computations I was previously studying only behaviorally. In large part, what I learned from these courses drove me to pursue a PhD in theoretical neuroscience. But if I really think back to the earliest indicators that I would write a thesis about biology with 65 equations in it, it would have to be the lessons I learned in my high school mathematics class with Mr. Omara. The IB internal assessment I was assigned in that class was the first time I discovered that modeling life with mathematics could be useful. And it has been a privilege to rediscover this fact over and over again since.

Contents

List of Figures

"No hay consuelo más hábil que el pensamiento de que hemos elegido nuestras desdichas; esa teleología individual nos revela un orden secreto y prodigiosamente nos confunde con la divinidad."

Jorge Luis Borges, Deutsches Requiem (1946)

Introduction

At the core of any nervous system is its capacity to generate and control movement. Without this critical ability, any other information processing would remain superfluous in the quest for survival [**?**]. It is the production of coordinated and purposeful movement that endows prey the capacity to avoid predators and predators the capacity to hunt prey. It thus goes without saying that one of the critical functions of vertebrate and invertebrate nervous systems is to generate goal-directed movement.

What is remarkable about mammalian nervous systems in particular is their ability to do so across an astonishingly wide variety of contexts. Through learning, mammals are able to adapt their movements to previously unexperienced environments or perturbations [Shadmehr and Mussa-Ivaldi, 1994, Li et al., 2001, Paz et al., 2005, Lopes et al., 2017, Kawai et al., 2015, Mathis et al., 2017, Perich et al., 2018, Sun et al., 2020]. Critically, they are able to do so flexibly and rapidly: learning a new motor behavior often requires relatively little practice, and newly learned motor behaviors rarely interfere with older ones. These two properties are crucial for intelligent and adaptive behavior. Moreover, they remain elusive in modern artificial intelligence systems, which still today struggle with generalizing their behavior to novel domains and with continual learning [Hassabis et al., 2017]. Understanding the algorithms and biological mechanisms by which mammalian brains support such motor learning is thus a central aim of neuroscience.

A typical neuroscientific approach to sudying motor learning is to record neural activity while an animal learns to perform a physical motor behavior [Li et al., 2001, Ungerleider et al., 2002, Yin et al., 2009, Chen et al., 2015, Cichon and Gan, 2015, Hayashi-Takagi et al., 2015, Vyas et al., 2018, Perich et al., 2018, Sun et al., 2020]. The idea is that, by characterizing the changes in a neural population over learning, we can hopefully glean some insight into the algorithms driving the behavioral improvements accompanying them. A major challenge to this approach, however, is that the functional roles of the various circuits underlying motor control are not well understood [Lopes et al., 2017]. Even the direct contributions of corticospinal tract neurons remain unresolved [Lemon, 2008], as they incorporate a myriad of downstream processes from the richly recurrent circuitry in the spinal cord [Brownstone et al., 2015] to the non-linear dynamics of muscle contraction [Todorov, 2000]. Consequently, while the teleology of the changes in behavior during motor learning may often be transparent, that of the changes in the underlying neural activity and circuitry typically remain unknown.

This makes it difficult to derive concrete conclusions about how these changes support learning and

what the computational principles behind them are. Are changes in activity driving improvements in the actual execution of the motor behavior? Are they instead driving modifications to a preparatory plan? Are they updates to an internal forward model? Or, on the other hand, do they simply reflect new memories of past errors, and have no direct role in driving the changes in behavior? While clever experimental paradigms have been developed to tease these possibilities apart, it is often impossible to do so since multiple such components are typically engaged by any given motor learning task.

If we perfectly understood the downstream consequences of the neural population in question, these ambiguities could potentially be resolved. But this is further comlpicated by the massively modular and distributed architecture of mammalian motor systems [Dum and Strick, 2002, Shadmehr and Krakauer, 2008, Lemon, 2008, Cisek and Kalaska, 2010], which obscures the biological susbstrates of learning-related changes in a given population and its downstream consequences. Local changes in activity could arise via modifications in either local or distal circuitry, and teasing these two possibilities apart is challenging without a precise understanding the roles of the relevant population(s). In sum, these considerations suggest that, to study the neural substrates of motor learning, we must first understand the long and winding road from neural activity to movement.

A unique way to address this challenge is to replace this serpentine road with a simpler and more transparent alternative. Indeed, circumventing this pathway is a necessary treatment for patients with tetraplegia, for whom alternative routes to movement must be developed to allow them to move. With this goal in mind, biomedical engineers and clinicians have developed brain-machine interfaces (BMIs) that transform internally generated neural activity (recorded through an implant) into movement of an external effector. Critically, this novel mapping from neural activity to movement, termed the BMI decoder, must be learned by the user. This provides an opportunity to study motor learning under complete knowledge and control of how the activity of the recorded population translates to movement [Golub et al., 2016]. Moreover, by making these mappings simple, we can invert them and make explicit statements about how neural activity must be reshaped over learning in order to support BMI movements [Athalye et al., 2017, Golub et al., 2018, Hennig et al., 2018]. This in turn allows us to make concrete hypotheses about how such learning might be implemented. This is exactly the approach we will exploit in the research presented in this doctoral thesis.

While motor control with a BMI is not exactly the same as natural motor control, there are many reasons to expect that similar mechanisms are engaged by both types of motor behaviors. Substantial evidence suggests that BMI control employs many of the same computations associated with natural motor control, such as sensory prediction errors [Vendrell-Llopis et al., 2019], state estimation with forward models [Golub et al., 2015], online error correction using sensory feedback [Stavisky et al., 2017], and motor planning [Vyas et al., 2018]. Additionally, the basal ganglia – a structure classically associated with motor learning in mammals and other phylogenetic classes – seem to play a direct causal role

in BMI learning [Koralek et al., 2012, Neely et al., 2018]. At the behavioral level, BMI learning occurs on a similar timescales as manual motor adaptation [Shadmehr and Mussa-Ivaldi, 1994, Li et al., 2001, Paz et al., 2005, Mathis et al., 2017, Perich et al., 2018, Sun et al., 2020] when the BMI decoder is carefully calibrated [Sadtler et al., 2014], and knowledge acquired during BMI learning can carry over to manual movements [Vyas et al., 2018]. These observations strongly suggest that the processes underlying BMI learning at the very least overlap with those driving natural motor learning.

Most importantly, many of the characteristic properties of mammalian motor learning are conserved in BMI learning:

 ▷ *adaptability*: human and non-human primates are able to learn to use completely novel decoders [Oby et al., 2019], even randomly generated ones [Ganguly and Carmena, 2009]

 ▷ *efficiency*: when the BMI decoder is carefully calibrated, they can do so on relatively fast timescales, with only minutes to hours of practice [Jarosiewicz et al., 2008, Chase et al., 2012, Gilja et al., 2012, Sadtler et al., 2014, Sakellaridi et al., 2019]

 ▷ *flexibility*: they can learn to control a new decoder without forgetting an old one [Ganguly and Carmena, 2009] and without forgetting how to control their natural limbs. They can also seamlessly switch back and forth between different BMI decoders [Ganguly and Carmena, 2009] and between BMI and manual control [Ganguly et al., 2011].

A comprehensive understanding of BMI learning thus has the potential to provide insight into the biological machinery giving rise to each of these aspects of natural motor learning.

This doctoral thesis presents the development of a theory of the biological structures underlying BMI learning. We will in particular focus on properties 1 and 2 above, by developing a theory that can capture both the efficiency with which carefully calibrated decoders can be learned and the adaptability with which non-human primates can learn arbitrary unstructured decoders when they are given enough practice to do so. We will apply this theory to model BMI learning across three different BMI tasks, providing a unifying explanation for BMI learning phenomena that had been previously treated completely separately in the literature. We will explicitly focus on BMIs that record from motor cortex, although many of the ideas we propose may be relevant to BMI's in other brain areas as well.

The thesis follows a somewhat untraditionally asymmetric organization:

1. In chapter 1, we motivate the theory from first principles and present it in its simplest form. We also demonstrates its ability to account for various phenomena observed during BMI learning in the context of three different tasks. In the remaining chapters, we focus on only one of these three tasks, so the first chapter comprises more than half of the thesis.

2. In chapter 2, we attempt to provide some mathematical grounding to the arguments made in chapter 1, through formal analysis of a simple case. This reveals certain connections between our

theory of BMI learning to classical concepts in control theory.

3. Finally, in chapter 3 we address one of the most substantial oversimplifications of the simpler models analyzed in chapters 1 and 2. We show that many of the intuitions and results from the simpler models extend to a more complex setting, and discuss connections to additional experimental literature.

Chapter 1

Previous models of motor cortical BMI learning have postulated that synaptic plasticity within motor cortex underlies learning [Legenstein et al., 2010, Engelhard et al., 2019, Wärnberg and Kumar, 2019, Feulner and Clopath, 2021]. Indeed, models of the synaptic connectivity required for a recurrent network to solve a BMI reaching task [Wärnberg and Kumar, 2019] and the plasticity rules by which that connectivity might be learned [Feulner and Clopath, 2021] can account for differences in learning between certain classes of BMI decoders [Sadtler et al., 2014]. However, a fundamental limitation of synaptic plasticity is the curse of dimensionality: mammalian cortical circuits contain billions to trillions of synapses, meaning that learning via optimization of their weights would entail solving an extremely high-dimensional optimization problem. In the best of cases – when the objective function and its gradient are explicitly known – solving such problems typically requires vast amounts of training data. In the BMI learning case, this difficulty is only exacerbated. Because the subject's motor system has no explicit access to the BMI decoder, the relationship between internal neural activity and movement – and, by extension, task performance – is unknown. This means that gradients of task performance with respect to internal biological parameters must be estimated through trial and error [Héliot et al., 2010, Feulner and Clopath, 2021]. Moreover, this estimation problem is made even more difficult by the biological constraints of neurons and synapses, which preclude synaptic plasticity rules from back-propagating gradients through the neural circuitry [Crick, 1989, Bartunov et al., 2018, Lillicrap et al., 2020]. In sum, these considerations suggest that BMI learning by synaptic plasticity should be slow and highly limited (although see [Raman et al., 2019] for exceptions to parts of this argument).

This is inconsistent with the strikingly fast and flexible learning observed in many BMI experiments, in which non-human primates are able to achieve proficient control with only 10s to 100s of trials of practice [Fetz and Baker, 1973, Jarosiewicz et al., 2008, Chase et al., 2012, Gilja et al., 2012, Sadtler et al., 2014]. Moreover, the hypothesis that motor cortex undergoes substantial synaptic changes over learning is inconsistent with the observation that the statistical structure of motor cortical population activity remains remarkably conserved after learning. The repertoire of activity patterns employed for BMI control is unchanged after training on a new decoder for a few hours [Golub et al., 2018, Hennig et al., 2018], and motor cortical neuron tuning to manual reaches remains largely fixed after performing a BMI reaching task [Ganguly et al., 2011]. This preservation of motor cortical operation is also consistent with the fact that learning a BMI task does not seem to interfere with natural limb control

[Ganguly et al., 2011].

Together, these observations call into question the plausibility of synaptic plasticity in motor cortex as the mechanism underlying BMI learning, at the very least for the short timescales of learning observed in some of the studies cited above. These studies suggest instead that the brain might take a more parsimonious learning strategy, whereby (1) the dimensionality of the learning problem is kept low to enable fast learning, and (2) the motor cortical machinery for natural movements is kept intact.

A learning strategy that satisfies these two criteria is that of "re-aiming" [Jarosiewicz et al., 2008, Chase et al., 2010, Chase et al., 2012] or "intrinsic variable learning" [Hwang et al., 2013, Sakellaridi et al., 2019]. Under this strategy, the animal exploits the pre-existing motor cortical circuitry by learning an association between intended BMI movements and motor commands that would otherwise be used during natural motor behavior. For example, if the BMI decoder is such that motor cortical activity generated during a *leftward* arm reach would lead to a *rightward* BMI movement, then the animal would learn to employ the motor command usually reserved for *leftward* arm reaches in order to achieve this *rightward* BMI movement (fig. 1a). This strategy easily satisfies criteria 1 and 2 above. The dimensionality of the learning problem is kept low because both BMI movements – movements of a 2D [Ganguly and Carmena, 2009, Chase et al., 2012, Sadtler et al., 2014, Oby et al., 2019] or 3D [Taylor et al., 2002, Carmena et al., 2003, Jarosiewicz et al., 2008] effector – and natural motor commands [d'Avella and Bizzi, 1998, d'Avella et al., 2003, Ivanenko et al., 2003, Todorov, 2004, Kutch and Valero-Cuevas, 20 Kuppuswamy and Harris, 2014] are low-dimensional. The motor cortical circuit can be kept intact because the patterns of activity used for manual and BMI control are the same.

There is substantial evidence both for and against the re-aiming strategy. [Jarosiewicz et al., 2008, Chase et al., 2010, Chase et al., 2012, Hwang et al., 2013, Sakellaridi et al., 2019] However, previous treatments have only considered the effects of re-aiming on individual neuron responses, relying on a feedforward tuning curve description of how re-aiming affects neural activity and behavior. Here, we re-visit the re-aiming hypothesis at a more mechanistic level, by formulating a theory of how such a learning strategy might shape population activity in a recurrent circuit model of primary motor cortex. Via simulation and analysis of this circuit, we derive predictions about how neural activity and behavior would change under a pure re-aiming learning strategy in the context of four commonly employed BMI tasks. This in turn will allow us to re-evaluate the extent to which various observations about BMI learning are, or are not, consistent with re-aiming.

Results

I.1.1 Re-aiming as optimization of low-dimensional inputs to motor cortex

We begin by modeling motor cortex as a recurrent neural network driven by an upstream population of neurons (fig. 1b),

$$\tau\frac{\mathrm{d}x_i}{\mathrm{d}t} = -x_i + \sum_{j=1}^{N} W_{ij}^{\mathrm{rec}} r_j + \sum_{j=1}^{M} W_{ij}^{\mathrm{in}} u_j \tag{1}$$

$$r_i = \phi(x_i) = \begin{cases} x_i & \text{if } x_i > 0 \\ 0 & \text{else} \end{cases}$$

where $r_1, r_2, \ldots, r_N$ and $u_1, u_2, \ldots, u_M$ denote the firing rates of the motor cortical and upstream neurons, respectively. A rectified linear activation function $\phi(\cdot)$ is used to ensure that firing rates are strictly non-negative. In the simulations presented here, we assume that firing rates are low at the start of each trial of BMI control, and thus set the initial conditions to 0, $x_i(0) = 0$. The weights W_{ij}^{rec} and W_{ij}^{in} represent the strengths of the synaptic connections between motor cortical neurons and from the upstream population, respectively. To avoid making any strong commitments as to the structure of recurrent connectivity within motor cortex, we use randomly connected networks throughout the results section of this chapter; simulations with other relevant connectivity structures are presented in the supplementary results section (fig. 9).

Inspired by recent models and theories of motor cortical function [Shenoy et al., 2013, Hennequin et al., 2014, Sussillo et al., 2015], we assume that movements are executed by driving motor cortex with a temporally simple upstream input. We assume that the rich intrinsic dynamics of the motor cortical circuit suffice to generate the spatiotemporal patterns of cortical activity necessary for the execution of a given motor behavior [Russo et al., 2018]. This leaves to the upstream inputs the role of simply selecting which patterns to generate at which time in order to execute that behavior. Upstream firing rates are therefore assumed to have little temporal structure, here fixed to be constant in time during the execution of the behavior (more complex input dynamics are considered in chapter 2).

Most importantly, we assume that the spatial structure of the upstream inputs is modulated by a "motor command" representing the desired behavior to be executed. In accordance with the fact that motor behaviors are generally low-dimensional [d'Avella and Bizzi, 1998, d'Avella et al., 2003, Ivanenko et al., 2003, Todorov, 2004, Kutch and Valero-Cuevas, 2012, Kuppuswamy and Harris, 2014], we assume that the space of all possible motor commands is low-dimensional as well. We formalize this by expressing the motor command as a K-dimensional vector $\boldsymbol{\theta}$, parameterized by $K \ll N$ motor variables $\theta_1, \theta_2, \ldots, \theta_K$. These are encoded in the firing rates of the upstream population according to a set of encoding weights
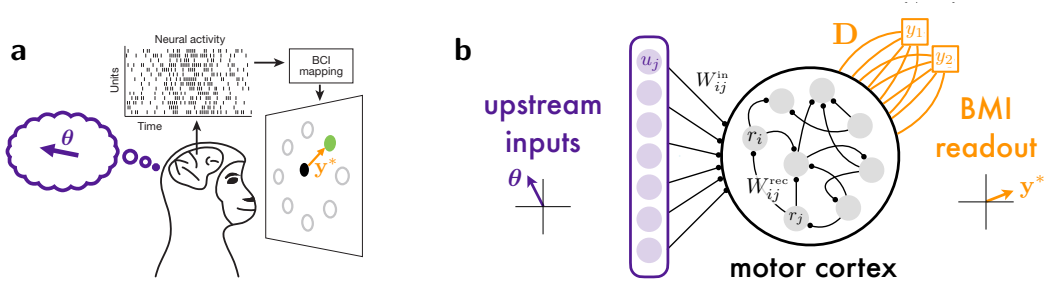
Figure 1: Modeling the re-aiming learning strategy.
(1a) Depiction of the re-aiming strategy for BMI learning. If activity evoked by imagining a leftward planar movement moves the BMI cursor right, then the animal learns to use this motor command for moving the cursor to the right. Critically, the space of imagined planar movements is low-dimensional.
(1b) Proposed model of re-aiming. Upstream inputs u_j to motor cortex are modified within a low-dimensional space, parameterized by a motor command vector $\boldsymbol{\theta}$. BMI readouts are a linear readout from motor cortical firing rates, through a decoding matrix $\mathbf{D}$. Re-aiming is formalized as identifying the motor command $\boldsymbol{\theta}$ that will achieve some target BMI readout $\mathbf{y}^*$ at some future point in time.

U_{ij},

$$u_i(\boldsymbol{\theta}) = \phi\left(\sum_{j=1}^{K} U_{ij}\theta_j\right), \tag{2}$$

where the rectified linear activation function $\phi(\cdot)$ is again used to ensure firing rates are non-negative. For the sake of simplicity, we set the encoding weights U_{ij} randomly.

During BMI control, motor cortical firing rates $\mathbf{r}(t) = \begin{bmatrix} r_1(t) & \cdots & r_N(t) \end{bmatrix}$ are directly translated to behavior through the readout from a BMI decoder, here assumed to be linear,

$$\mathbf{y}(t) = \mathbf{D}\left(\mathbf{r}(t) - \boldsymbol{\mu}\right). \tag{3}$$

The constant vector $\boldsymbol{\mu}$ is included to center the strictly positive firing rates $\mathbf{r}(t)$ (see Methods). The BMI readout $\mathbf{y}(t)$ could determine, for example, the position [Serruya et al., 2002, Hochberg et al., 2006] or velocity [Velliste et al., 2008, Jarosiewicz et al., 2008, Ganguly and Carmena, 2009, Chase et al., 2012, Sadtler et al., 2014] of the BMI effector. A subject learning to control a novel BMI decoder must find a way to generate motor cortical activity patterns $\mathbf{r}(t)$ that will produce the target readouts $\mathbf{y}^*(t)$ demanded by the task at hand.

Various components of the motor cortical circuit could be optimized to do so, such as its local recurrent connectivity W_{ij}^{rec} [Wärnberg and Kumar, 2019, Feulner and Clopath, 2020] or input connectivity from the upstream population W_{ij}^{in} [Legenstein et al., 2010, Engelhard et al., 2019]. The re-aiming hypothesis, however, makes the strong commitment that the local motor cortical circuit structure remains intact, and that instead only the upstream motor commands $\boldsymbol{\theta}$ are optimized [Chase et al., 2010]. The same motor cortical machinery used for natural motor control is thus exploited for BMI control.

This BMI learning strategy has a number of advantages, the primary of which is that it reduces the dimensionality of the learning problem. Optimizing the recurrent weights of the motor cortical circuit, for example, would require optimizing N^2 synaptic weights, whereas optimizing the motor command $\boldsymbol{\theta}$

only requires optimizing $K \ll N$ motor variables. Moreover, not all K motor variables need to optimized – we will argue below that, in certain settings, subjects may be optimizing only a subset of $\widetilde{K} < K$ task-relevant motor variables. Such a reduction in the number of optimized parameters may be critical for efficient learning in the absence of gradient information.

However, while reducing the dimensionality of the learning problem may make it easier to find a solution, it also reduces the space of available solutions. A reduction in the number of motor variables that can be manipulated for BMI control constrains how motor cortical activity can be adjusted to perform the BMI task. This reduction in flexibility implies limitations on the BMI problems that can be learned via this low-dimensional learning strategy. Our goal here is to develop a formal theory of re-aiming that will allow us to better understand these limitations, and evaluate whether they are consistent with the limitations empirically observed in BMI learning experiments. What determines whether a BMI decoder can or cannot be learned by re-aiming? To what extent can motor cortical population activity be modified by re-aiming? How might this constrain behavior during BMI control? And, crucially, are these neural and behavioral signatures of re-aiming consistent with empirical observation? In the following, we address these questions in the context of four different BMI learning experiments.

To obtain general insights, we analyze a simplified model of re-aiming in which the motor command is optimized to produce a target readout $\mathbf{y}^*$ at some fixed future endpoint time t_{end},

$$\hat{\boldsymbol{\theta}}(\mathbf{y}^*) = \underset{\theta_1, \theta_2, \ldots, \theta_{\widetilde{K}}}{\arg\min} \; \left\| \mathbf{y}\big(t_{\text{end}}; \boldsymbol{\theta}\big) - \mathbf{y}^* \right\|^2 + \frac{\gamma}{M} \sum_{i=1}^{M} u_i(\boldsymbol{\theta})^2. \tag{4}$$

The vector $\mathbf{y}(t_{\text{end}}; \boldsymbol{\theta})$ denotes the BMI readout at time t_{end} produced by driving the model motor cortical network with the upstream inputs specified by the motor command vector $\boldsymbol{\theta}$ (equation 2). This is calculated by integrating the recurrent network dynamics (equation 1) to obtain the motor cortical firing rates $\mathbf{r}(t; \boldsymbol{\theta})$ generated by that motor command, and decoding from the firing rates at the endpoint time, $\mathbf{r}(t_{\text{end}}; \boldsymbol{\theta})$, via equation 3. The second term in this cost function constitutes the metabolic cost of the upstream firing rates induced by the motor command $\boldsymbol{\theta}$, included in the objective function to ensure that only biologically plausible solutions are allowed. A quadratic cost is assumed for the sake of analytical and numerical tractability (see Methods). Finally, $\widetilde{K}$ denotes the number of motor variables optimized under the re-aiming strategy. We call these the *aiming variables*, as these are the ones that will be optimized by re-aiming. For simplicity, the remaining motor variables $\theta_{\widetilde{K}+1}, \theta_{\widetilde{K}+2}, \ldots, \theta_K$ that are not optimized are fixed at 0.

For a given decoding matrix $\mathbf{D}$, solutions to this equation constitute a concrete hypothesis about what subjects learn when learning to control that BMI decoder. In the following, we analyze these solutions to evaluate whether this hypothesis is consistent with empirical observations from BMI learning experiments. The question of *how* subjects might learn such re-aiming solutions is left for future work.
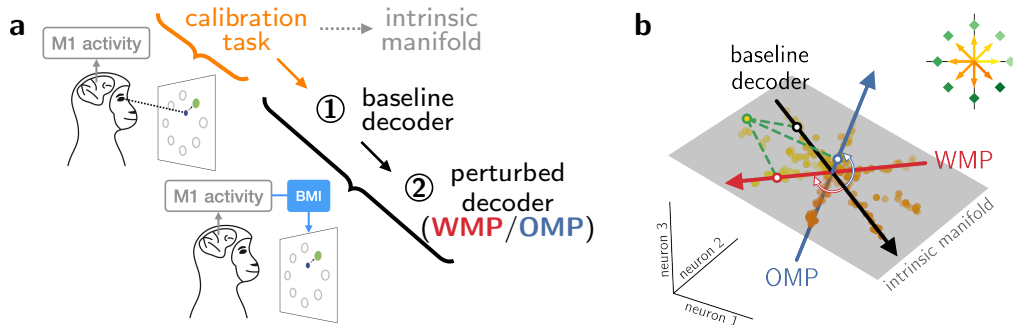
Figure 2: BMI learning task of Sadtler et al. (2014) [Sadtler et al., 2014].
(2a) Task structure. Subjects are first submitted to a "calibration task" whereby they passively observe cursor reaches on a screen. Recorded activity is used to estimate the intrinsic manifold and construct the baseline decoder. Subjects are then instructed to perform center-out cursor reaches under BMI control, first using the baseline decoder and then with a perturbed decoder constructed by perturbing the baseline decoder. This perturbation could be either a within-manifold (WMP) or outside-manifold (OMP) perturbation.
(2b) Low-dimensional illustration of intrinsic manifold and its relationship to decoding matrices in this task. Colored dots represent activity patterns recorded during different trials of the calibration task, colored by the cursor velocity presented on that trial. These stimuli are depicted by the inset in the top right, with the presented cursor velocities depicted by arrows color-matched to the activity patterns in the main figure, and the reach targets (for the subsequent center-out cursor reaching task) depicted by green diamonds. The evoked neural activity patterns reside predominantly within the two-dimensional plane depicted by the gray rectangle, the so-called intrinsic manifold in these data. Three hypothetical one-dimensional decoders are depicted by colored arrows (labelled baseline decoder, WMP, and OMP). Linear readouts from these decoders can be visualized by projecting individual activity patterns onto the corresponding decoder vector. This is illustrated for one activity pattern marked in green, whose projections onto each of the three decoders is shown. A projection far from the origin (the intersection of the three decoders) corresponds to a large absolute value of the readout. Because this activity pattern resides close to the intrinsic manifold, it yields a large readout through the baseline decoder and WMP, which are both well aligned with the intrinsic manifold. In contrast, the activity pattern's readout through the OMP is much weaker (i.e. its projection onto this decoder is much closer to the origin), since this decoder is oriented away from the intrinsic manifold. It is important to keep in mind that this illustration is only a simplified cartoon of the true task, in which more than three neurons are recorded, the intrinsic manifold is higher-dimensional (8-12D instead of 2D), and the decoders constitute two-dimensional planes (cf. figure 5a) rather than one-dimensional vectors.

It is also important to acknowledge that equation 4 constitutes an incomplete description of the true BMI learning problem. Controlling the BMI effector's movement requires taking into account the effector's dynamics (e.g. the readout $\mathbf{y}(t; \boldsymbol{\theta})$ might control the effector's velocity rather than its position) and specifying the entire sequence of readouts over time, rather than just the readout at one particular time t_{end}. Moreover, equation 4 ignores the role of sensory feedback in continuous closed-loop BMI control [Golub et al., 2015, Stavisky et al., 2017, Shanechi et al., 2017]. That said, this simplified model of re-aiming will prove useful to intuit general principles of the re-aiming learning strategy, which, as we show in chapter 3, extend to more complex settings. After all, being able to produce a target readout at a fixed future time is, in a loose sense, a pre-requisite to solving the full closed-loop control problem.

I.1.2   Re-aiming implies neural constraints on short-term learning

To illustrate the limitations of re-aiming, we begin by modelling a BMI learning task designed by Sadtler et al. (2014) [Sadtler et al., 2014]. In this task, subjects learn to perform center-out reaches with a cursor on a screen, whose velocity is controlled by the readout from a linear BMI decoder. The task comprises two phases in which two different decoding matrices $\mathbf{D}$ are used (fig. 2a). In the first phase of the task, a baseline decoder is used. This decoder is constructed based on neural responses during

a prior calibration task, in which the subject passively views 2D cursor reaches to eight radial reach targets presented on a screen. The decoding matrix is built so that the patterns of neural activity evoked by these visual stimuli would suffice to move the cursor towards the corresponding reach targets (see Methods). Importantly, neural responses during this task tend to occupy a low-dimensional subspace, termed the "intrinsic manifold" [Sadtler et al., 2014]. The construction of the baseline decoder ensures that it is highly aligned with the intrinsic manifold, so that activity patterns within this subspace can produce large readouts through the decoder (fig. 2b). Sadtler et al. found that subjects could easily perform center-out reaches with the baseline decoder, requiring only a few trials of practice to learn.

In the second phase of the task (fig. 2a), the decoding matrix $\mathbf{D}$ of the baseline decoder is perturbed, and the subject must learn to perform the same center-out cursor reaches with the resulting perturbed decoder. The key manipulation is that the decoding matrix can be perturbed in two different ways. *Within*-manifold perturbations (WMP's) consist of randomly re-orienting the baseline decoder *within* the intrinsic manifold. WMP's therefore alter how neural activity gets mapped to readouts while preserving the decoder's alignment with the intrinsic manifold (fig. 2b). *Outside*-manifold perturbations (OMP's), on the other hand, directly disrupt this alignment. Under OMP's, activity patterns within the intrinsic manifold do not suffice to move the cursor in the right direction, and new activity patterns *outside* of the intrinsic manifold must be produced by the subject in order to successfully perform the task (fig. 2b).

Sadtler et al. found that, with 1-2 hours of practice (i.e. a few hundred trials), non-human primates could generally learn to successfully perform cursor reaches with WMP's. On the contrary, such short-term learning did not seem to be possible with OMP's, under which relatively little improvement was observed over this training period. This was interpreted to reflect a "neural" constraint on learning, whereby the local motor cortical population is unable to learn to produce the activity patterns necessary to control OMP's – that is, activity patterns outside of the intrinsic manifold. Here we argue that this limitation of short-term learning could be a consequence of subjects using a re-aiming learning strategy. Such a low-dimensional learning strategy could account for the incredible efficiency with which subjects learn to control WMP's, and in the following we show that it can also explain the lack of learning under OMP's.

To simulate this BMI learning task, we begin by simulating the calibration task. We simulate motor cortical responses to the visually presented cursor reaches by driving the model motor cortical network with motor commands $\boldsymbol{\theta}$ that depend on the cursor's velocity on each trial (a constant vector pointing in the direction of the reach target on that trial). Reflecting the two-dimensional nature of these stimuli, we set the first two motor variables θ_1, θ_2 to the 2D coordinates of this cursor velocity, and set the remaining motor variables $\theta_3, \theta_4, \ldots, \theta_K$ to 0. We find that driving the network with these motor commands leads to population activity occupying an eight-dimensional subspace (fig. 3a), which we define as the intrinsic manifold. Following the methods of [Sadtler et al., 2014], we then use this subspace to construct the

baseline decoder and sample within- and outside- manifold perturbations (see Methods).

We next postulate that subjects learn to control these decoders by re-aiming with the same two motor variables driving motor cortex during the calibration task. Because the motor system has been exposed to fluctuations in only these two motor variables, we hypothesize that the motor system determines them to be the most relevant to the task at hand and therefore exploits them for re-aiming. We thus model BMI learning by optimizing θ_1 and θ_2 with respect to the re-aiming objection function (equation 4, with $\widetilde{K} = 2$), leaving the remaining motor variables fixed to 0 as in the calibration task ($\theta_3 = \theta_4 = \ldots = \theta_K = 0$). Under this learning strategy, learning can proceed extremely efficiently, as only two variables need to be optimized. The motor commands available for BMI control, however, are now severely restricted: only two motor variables are free to change, and they are furthermore bounded by the metabolic cost incurred on the upstream firing rates (cf. equation 4).

Critically, these restrictions on the upstream motor commands induce restrictions on the motor cortical activity patterns available for BMI control. To understand exactly how motor cortical activity is restricted, we consider the set of all motor cortical activity patterns $\mathbf{r}(t_{\mathrm{end}}; \boldsymbol{\theta})$ that can be reached at a fixed endpoint time t_{end} by a motor command $\boldsymbol{\theta}$ permissible under this learning strategy. We term this set the "reachable repertoire." To visualize its structure, we select a large sample of permissible motor commands and compute the activity patterns they generate at $t_{\mathrm{end}} = 1000$ms. The selected motor commands are plotted in figure 3b; recall that only the first two motor variables, θ_1, θ_2, are allowed to change, so we can plot these motor commands in $\theta_1 - \theta_2$ space. The motor cortical activity patterns they generate are plotted in figure 3c, projected down to three dimensions via Principal Components Analysis (PCA) and plotted with markers matched in size and color to the motor commands in figure 3b that generated them. We emphasize that the points in this figure should not be thought of as spatiotemporal trajectories of activity; rather, they depict activity patterns generated at the same timepoint $t_{\mathrm{end}} = 1000$ms, with different activity patterns generated by different motor commands.

Despite these motor commands being two-dimensional, the network's dynamics generate population activity occupying more than three dimensions of state space. The three-dimensional projection in figure 3c in fact contains about 80% of the variance over the plotted N-dimensional activity patterns, revealing that the reachable repertoire resides within a subspace of moderately low dimensionality, higher than that of the motor commands giving rise to it (two-dimensional) but significantly lower than that of its ambient state space (N-dimensional). In fact, the reachable repertoire turns out to be completely contained within the intrinsic manifold. This can be seen by calculating the amount of variance across activity patterns in the reachable repertoire that is contained within each dimension of the intrinsic manifold. This is plotted in figure 3a with a purple curve, which reveals that the eight dimensions of the intrinsic manifold capture almost 100% of the variance in reachable activity patterns. The reason for this is that both the activity patterns in the reachable repertoire and the activity patterns evoked by
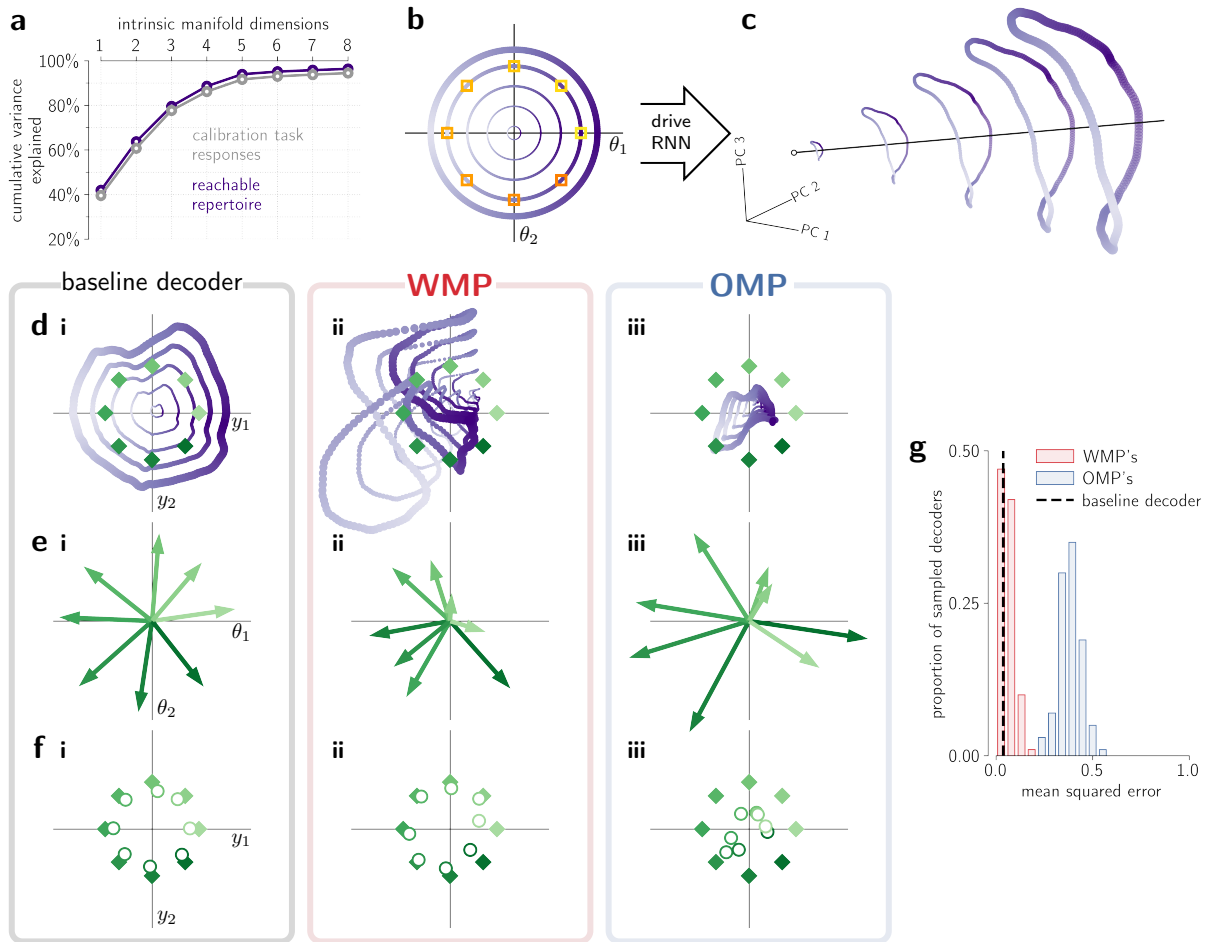
Figure 3: Re-aiming with two motor variables suffices to learn good solutions for within- but not outside- manifold perturbations.

(3a) Gray curve shows the variance of the calibration task neural response within each dimension of the intrinsic manifold. By construction, about 95% of the total variance is contained within the intrinsic manifold. Purple curve shows the variance over all activity patterns in the reachable repertoire that is contained within each dimension.

(3b) Motor commands used to generate activity patterns from the reachable repertoire. These comprise a dense range of angles on the $\theta_1 - \theta_2$ plane and 5 distinct motor command norms, $\|\boldsymbol{\theta}\|$, corresponding to each ring (see Methods for how these norms were picked). The motor commands used to simulate the calibration task are highlighted by the orange squares, picked to be unit norm, $\|\boldsymbol{\theta}\| = 1$. All other motor variables $\theta_3, \theta_4, \ldots, \theta_K$ are fixed to 0 under our model of re-aiming.

(3c) Activity patterns from the reachable repertoire at endpoint time $t_{\text{end}} = 1000$ms. Each activity pattern is color- and size- matched to the motor command in the previous panel that generated it; each ring of activity patterns is generated by the corresponding ring of motor commands in the previous panel. This ensemble of N-dimensional activity patterns is projected onto its top three principal components. Black line marks the apical axis of this conical manifold, drawn only to facilitate visualization.

(3d) Readouts from each of the reachable activity patterns plotted in figure 3c, through the baseline decoder and one example WMP and OMP. Marker colors and sizes are matched to the corresponding activity patterns in figure 3c. Green diamonds mark the eight radial target readouts for the cursor reaching task, set to the directions of the eight radial reach targets used in [Sadtler et al., 2014].

(3e) Re-aiming solutions $\hat{\boldsymbol{\theta}}(\mathbf{y}^*)$ for each decoder, plotted in $\theta_1 - \theta_2$ space. Recall that only θ_1 and θ_2 are optimized under our model of re-aiming; all other motor variables $\theta_3, \ldots, \theta_K$ are fixed to 0. The re-aiming solution to each target readout is matched in color to its corresponding diamond in figure 3d.

(3f) The colored circles show the readouts $\mathbf{y}\big(t_{\text{end}}; \hat{\boldsymbol{\theta}}(\mathbf{y}^*)\big)$ generated by the optimal re-aiming solutions shown in figure 3e, under the same three BMI decoders. The diamonds denote the eight radial target readouts as in figure 3d. Note that these readouts aren't necessarily the reachable readouts closest to the targets, as the re-aiming solutions generating them minimize the metabolic cost of the input as well the error (cf. equation 4).

(3g) Distribution of mean squared error achieved by the re-aiming solutions for 100 randomly sampled WMP's and OMP's. The mean squared error achieved by the re-aiming solutions for the baseline decoder is marked by the vertical dashed black line. Target readouts are unit norm, so a mean squared error of 1.0 is equivalent to producing a readout at the origin.

the calibration task – which define the intrinsic manifold – are generated by similarly low-dimensional motor commands $\boldsymbol{\theta}$, in which only two motor variables (θ_1, θ_2) are non-zero (fig. 3b). Consequently, all the activity patterns in the reachable repertoire reside within the same eight-dimensional subspace occupied by the calibration task responses.

The behavioral repercussions of this low-dimensional structure can be understood by looking at the BMI readouts from these reachable activity patterns – that is, the set of readouts reachable at time t_{end} by re-aiming. These readouts will, of course, depend on the BMI decoder, which specifies how activity patterns get translated to BMI readouts. In figure 3d, we visualize these reachable readouts under the baseline decoder, a WMP, and an OMP. Specifically, we calculate each decoder's readouts from each of the reachable activity patterns plotted in figure 3c and plot them in the 2D readout space. This reveals that the readouts reachable under the baseline decoder and under the WMP are much larger than those reachable under the OMP. Overlaid as solid green diamonds are the target readouts of the task, matching the eight radial reach targets in the center-out reaching task of [Sadtler et al., 2014]. While these lie firmly within the scope of the readouts reachable under the baseline decoder and the WMP, they are out of reach of any of the readouts reachable under the OMP. This entails that the re-aiming learning strategy is bound to fail for this OMP – under this decoder, none of the motor commands permissible by this learning strategy are able to produce the readouts demanded by the task at hand.

We confirm this by simulating re-aiming under each decoder. For each of the three decoders, we solve equation 4 for each target readout (using $t_{\text{end}} = 1000$ms) and drive the motor cortical network with the resulting optimal re-aiming solutions $\hat{\boldsymbol{\theta}}(\mathbf{y}^*)$. The solutions for each target readout $\mathbf{y}^*$ are plotted in figure 3e as 2D vectors in $\theta_1 - \theta_2$ space, color-matched to the target readouts in figure 3d. Note that the solutions for the perturbed decoders can be quite complex, beyond a simple rotation or scaling of the target readouts. The readouts produced by these optimal re-aiming solutions at the optimized endpoint time t_{end} are shown directly below in figure 3f. These optimal readouts necessarily reside within the manifold of reachable readouts visualized in figure 3d. As a result, only those under the baseline decoder and WMP reach the targets, with the optimal readouts under the OMP falling far short. Under this OMP, the re-aiming learning strategy evidently fails to provide an effective solution to the cursor reaching task.

The reason behind this can be immediately intuited from the aforementioned fact that the reachable repertoire only includes activity patterns within the intrinsic manifold. In other words, the reachable repertoire excludes activity patterns outside of the intrinsic manifold, which therefore are not reachable by re-aiming. Producing the target readouts under an OMP requires generating such activity patterns, and therefore re-aiming cannot succeed for these decoder perturbations. To check that this indeed holds for any OMP, we sampled 100 random WMP's and OMP's (see Methods) and repeated the above simulation for each of them. For each decoder, we quantify the success of re-aiming by the mean squared

error between the target readouts and the readouts generated by the re-aiming solutions to these targets. Figure 3g shows the distribution of this mean squared error over all the sampled WMP's and OMP's. As expected, we find that the re-aiming solutions for OMP's consistently result in higher mean squared error than do their WMP counterparts. Only the re-aiming solutions for WMP's are able to achieve mean squared error as low as that achievable under the baseline decoder.

We conclude that the re-aiming learning strategy, as formalized here, can only succeed for WMP's. This offers an explanation for why only these decoder perturbations seem to be learnable on the short timescale of a single experimental session. Because re-aiming can't succeed for OMP's, subjects must resort in these cases to an alternative – and presumably higher-dimensional – learning strategy, explaining why learning these decoders requires substantially more training [Oby et al., 2019]. We thus briefly remark that, during OMP control, subjects are likely to produce readouts differing markedly from those seen in figure 3d, which constitute predictions for what they would produce under the inadequate, and presumably quickly abandoned, re-aiming strategy.

I.1.3   Re-aiming predicts biases in short-term learning

A close look at figure 3d reveals an important difference between this WMP and the baseline decoder. Whereas the readouts reachable under the baseline decoder cover the workspace symmetrically, those reachable under the WMP are much larger in western than in eastern directions of the workspace (fig. 3dii). This explains why the optimal re-aiming solutions for the eastern targets in fact produce readouts that fall short of their targets (fig. 3fii). Such biases are not unique to this WMP, figure 4a reveals similar biases in the readouts reachable under an additional three other representative WMP's that we sampled.

This leads to an experimental prediction: if subjects are learning WMP's by re-aiming, then the BMI readouts they produce should be biased, even at the end of learning. To rigorously quantify the degree of bias predicted by our model, we calculated the maximal readout strength achievable in each target direction. We define a readout $\mathbf{y}$'s strength in a given target direction $\mathbf{y}^*$ by the alignment between these two vectors,

$$\rho\left(\mathbf{y}; \mathbf{y}^*\right) = \mathbf{y} \cdot \mathbf{y}^*, \tag{5}$$

where the vector $\mathbf{y}^*$ is taken to be unit norm ($\|\mathbf{y}^*\| = 1$). The *maximal* readout strength in that target direction is then given by the reachable readout with highest readout strength in that direction (i.e. by taking the maximum of $\rho\left(\mathbf{y}; \mathbf{y}^*\right)$ over all reachable readouts $\mathbf{y}$, cf. Methods equation 20). The strongest reachable readouts in each of the eight target directions are marked by green circles in figure 4a, color-matched to the target readouts in figure 3f. The readout strengths of the maximal readouts in each direction are then plotted below in figure 4b, revealing a smooth unimodal bias: the maximal readout
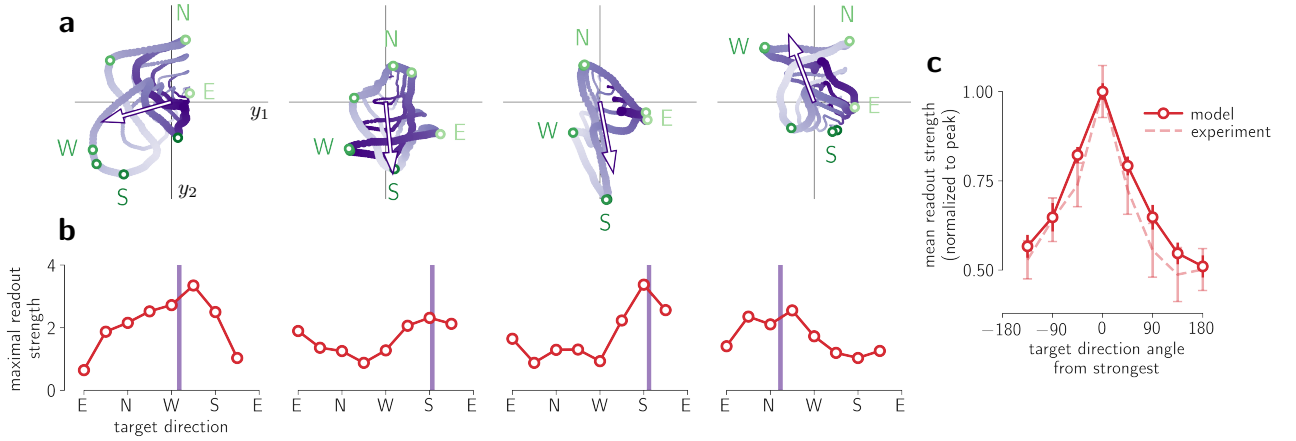
Figure 4: Re-aiming predicts biases in readouts after short-term learning of within-manifold perturbations.
(4a) Readouts reachable under four representative WMP's, using the same color conventions as in figure 3d. In each case, each of the four loops correspond to four distinct motor command norms, picked to aid visualization. The leftmost panel corresponds to the same example WMP shown in the previous figure. The reachable readouts with largest readout strengths in each target direction are overlaid, color-matched to the target readouts shown in figure 3d. The readouts with largest readout strengths in the east (E), north (N), west (W), and south (S) target directions are additionally marked by the corresponding letter. The projection of the reachable repertoire mean, $\mathbf{D}\bar{\mathbf{r}}$, is overlaid as an open arrow, arbitrarily rescaled for visibility.
(4b) Maximal readout strengths in each target direction, for each of the same four example WMP's. These are the readout strengths of the readouts marked by the green circles in the corresponding panel above. The vertical purple line indicates the direction of the projection of the reachable repertoire mean, $\mathbf{D}\bar{\mathbf{r}}$.
(4c) Solid curve is the mean maximal readout strength over all 100 sampled WMP's, obtained by horizontally centering the maximal readout curves (i.e. the curves in panel 4b, for all sampled WMP's) at their peaks and then averaging. Dashed line shows the mean readout strength during best trials of WMP control in the data of Sadtler et al. (2014), obtained by horizontally centering the readout strength curves for 46 experimental sessions and then averaging, exactly as was done with the simulated data. Both curves are normalized to their peaks to aid comparison of model and experimental data. Error bars show standard error of the mean.

strength is consistently highest for one target direction and then decays smoothly as a function of target direction angle from the strongest one. To characterize the shape of this bias for any generic WMP, we measured these maximal readout strength curves for each of the 100 WMP's we sampled, horizontally centered them at their peaks, and then averaged. The resulting curve of mean maximal readout strengths is shown in figure 4c, revealing a consistent smooth unimodal bias in the reachable readouts. In figure 9, we show that the shape of this bias is in fact highly conserved across several different models of motor cortical connectivity, indicating that it is an unavoidable consequence of the re-aiming learning strategy. The absence of such a bias in data would provide strong evidence against this theory of BMI learning.

To check whether such a bias was present in the data of [Sadtler et al., 2014], we identified the window of trials with best reaching performance under WMP control and calculated the average BMI readout strengths in each reach direction over this window (see Methods). This provided us with a curve of estimated maximal readout strengths for each session of WMP control. We then horizontally centered and averaged these curves over all sessions and subjects. The resulting curve of empirical readout strengths is overlaid in figure 4c with a dashed line. This curve reveals a smooth unimodal bias in the experimental data that is remarkably similar to that predicted by our model. Note that the decay in readout strengths from the peak is steeper in the data than in the model, which is consistent with the fact that our model's *maximal* readout strengths should provide an upper bound on the readout

strengths achievable under the re-aiming learning strategy. It is worth noting that no parameters of our model were fit to these data.

To check whether this average bias in the data arises via the same mechanisms as in the model, we next attempt to exploit our mechanistic understanding of the model to predict the direction of the bias in individual experimental sessions. In our model, biases in the readouts reachable under a given decoder stem from the geometric relationship between the decoding matrix and the reachable repertoire. This is illustrated in figure 5a using a simplified 3D caricature of the reachable repertoire. The reachable repertoire is depicted here by a purple cone, emulating its conical structure evident in figure 3c. The set of readouts reachable under a specific decoder can be thought of as a two-dimensional projection of the reachable repertoire onto the plane spanned by the rows $\mathbf{d}_1, \mathbf{d}_2$ of the decoding matrix $\mathbf{D}$. If, as in the illustration, this decoding plane is aligned with the apical axis of the cone (i.e. the axis going from the origin to the cone's base), then it is easy to appreciate that this projection will form a triangular manifold of reachable readouts, asymmetrically covering the workspace.

Crucially, this manifold need not be centered at the origin of the workspace. From equation 3, we have that the centroid of the reachable readouts is given by

$$\bar{\mathbf{y}} = \mathbf{D}\left(\bar{\mathbf{r}} - \boldsymbol{\mu}\right), \tag{6}$$

where $\bar{\mathbf{r}}$ is the reachable repertoire mean. The centroid of the reachable readouts is thus determined by the projection of the difference $\bar{\mathbf{r}} - \boldsymbol{\mu}$ onto the decoding plane. This is depicted in the illustration by the arrow going from the orange circle to the purple one. The reachable readouts will thus only be centered at the origin of the workspace if the reachable repertoire mean $\bar{\mathbf{r}}$ exactly matches the centering vector $\boldsymbol{\mu}$, or if their difference is orthogonal to the decoding plane so that its projection onto it is zero. In the illustration, $\bar{\mathbf{r}}$ and $\boldsymbol{\mu}$ differ in a direction that is parallel to the depicted decoding plane, so the centroid of the reachable readouts is non-zero, offset to the right of the workspace origin. This leads to a bias in the reachable readouts, whereby larger readouts are reachable on that side of the workspace than on the other.

If this is how the bias arises in our model, we should be able to predict the direction of the bias from the relationship between $\bar{\mathbf{r}}$ and $\boldsymbol{\mu}$. Recall that the centering vector $\boldsymbol{\mu}$ is fit to the calibration task neural responses. More specifically, it is set to the vector of mean responses over all time and stimuli (see Methods). This same centering vector is then used for all within- and outside- manifold perturbations, which perturb the decoding matrix $\mathbf{D}$ only. Recalling that the calibration task neural responses are driven by motor commands similar to those underlying the reachable repertoire (fig. 3b), it is easy to intuit that their mean $\boldsymbol{\mu}$ should bear a strong relationship to the reachable repertoire's mean $\bar{\mathbf{r}}$. This relationship is visualized in figure 5b, where we plot simulated calibration task responses together with activity patterns

from the reachable repertoire, projected down to the same subspace. Both the calibration task mean and the reachable repertoire mean point roughly in the same direction (at least within this subspace), differing mainly in their norms – because the reachable repertoire includes activity patterns generated by motor commands with stronger magnitudes than those driving the calibration task responses (fig. 3b), the reachable repertoire mean lies further away from the origin than the calibration task mean. Since $\bar{\mathbf{r}}$ and $\boldsymbol{\mu}$ point in the same direction, their difference does too, so we can write

$$\bar{\mathbf{r}} - \boldsymbol{\mu} \approx a\bar{\mathbf{r}}, \ a > 0. \tag{7}$$

Plugging this into equation 6, we have that the centroid of the reachable readouts is approximately given by

$$\bar{\mathbf{y}} \approx a\mathbf{D}\bar{\mathbf{r}}, \ a > 0. \tag{8}$$

This implies that the readouts reachable under the decoding matrix $\mathbf{D}$ will be biased in the direction of the reachable repertoire mean's projection through that decoder, $\mathbf{D}\bar{\mathbf{r}}$.

We can see that this is the case in the example WMP's shown in figure 4a. First note that, in reality, the reachable repertoire forms a higher-dimensional structure than the simple 3D cone illustrated in figure 5a. Moreover, linear decoding (equation 3) may comprise a complex mixture of reflections, scalings, and rotations in addition to the simple projection operation depicted in the illustration of figure 5a. For these reasons, the readouts reachable under each of these decoders form intricately twisted loops (fig. 4a), rather than simple circular rings like in the illustration. However, in each case the full manifold of readouts formed by the union of these loops conserves the key geometric properties caricatured in the illustration: it is asymmetric and biased to one side of the workspace. The side to which it is biased is moreover predicted by the projection of the reachable repertoire mean, marked by the arrow. The direction of this projection is marked by a vertical line in figure 4b, which reveals that the target direction with highest maximal readout strength is typically near it.

Does this also hold true in the WMP readouts produced by subjects in the experiment? Unfortunately, we cannot directly test this prediction because we have no way of measuring the reachable repertoire of a subject's motor cortex. However, we can estimate it from motor cortical activity during center-out reaches with the baseline decoder. The mean population activity generated by the re-aiming solutions for this decoder, $\bar{\mathbf{r}}_{\text{est}}$, is highly correlated with the reachable repertoire mean, $\bar{\mathbf{r}}$, so we can estimate the latter with the former. Maximal readout strengths under a given WMP thus tend to be highest for target directions pointing in the direction of the projection of $\bar{\mathbf{r}}_{\text{est}}$ through that decoder, $\mathbf{D}\bar{\mathbf{r}}_{\text{est}}$, exactly as we observed to be the case for the true repertoire mean, $\bar{\mathbf{r}}$. We can see that this is the case in figure 5c, which plots maximal readout strengths in each target direction as a function of its angle with the projection of $\bar{\mathbf{r}}_{\text{est}}$ through the decoder, for each of the 100 sampled WMP's. This reveals a strong negative correlation
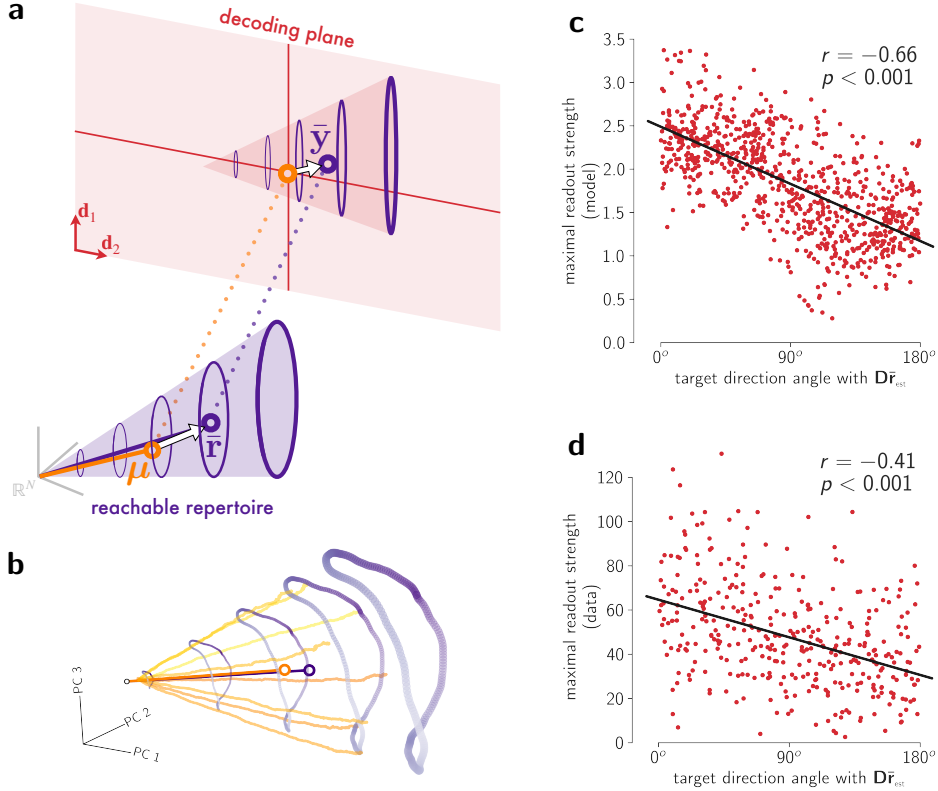
Figure 5: The geometry of the reachable repertoire predicts the direction of biases after short-term learning of within-manifold perturbations.

(5a) Illustration of how the conical structure of the reachable repertoire gives rise to biases in reachable readouts. The purple cone depicts the reachable repertoire, with purple rings depicting activity patterns generated by aiming variable pairs with constant norm, analagous to those seen in figure 3c. The red rectangle depicts the plane spanned by the two rows of the decoding matrix $\mathbf{d}_1, \mathbf{d}_2$. The reachable readouts are visualized as the projection of the reachable repertoire onto this decoding plane. The two open orange dots depict the calibration task mean, $\boldsymbol{\mu}$, and its readout, $\mathbf{D}(\boldsymbol{\mu} - \boldsymbol{\mu}) = \mathbf{0}$. Its readout marks the origin of the owrkspace, depicted by the intersection of the two red lines. The two open purple dots depict the reachable repertoire mean $\bar{\mathbf{r}}$ and its readout, $\mathbf{D}(\bar{\mathbf{r}} - \boldsymbol{\mu})$. Its readout marks the centroid of the reachable readouts, $\bar{\mathbf{y}}$. The difference $\bar{\mathbf{r}} - \boldsymbol{\mu}$ and its projection onto the decoding plane are highlighted by the white arrows.

(5b) Simulated motor cortical responses in the calibration task, color-matched to the motor commands in figure 3b driving these responses. These are plotted together with the same reachable activity patterns from figure 3c, projected onto the same three dimensions. The open dots in the interior of this conical structure show the calibration task mean $\boldsymbol{\mu}$ in orange and the reachable repertoire mean $\bar{\mathbf{r}}$ in purple. Note that, by definition, the calibration task neural responses at time $t_{\text{end}} = 1000\text{ms}$ (the last point in each trajectory) lie almost exactly on the reachable repertoire (offset slightly only because of noise in the dynamics).

(5c) Maximal readout strength in each target direction for each sampled WMP, plotted as a function of the angle between the target direction and $\mathbf{D}\bar{\mathbf{r}}_{\text{est}}$, the projection of the mean population activity during simulated reaches with the baseline decoder (see text and Methods for details). A total of 8 target directions $\times$ 100 sampled WMP's = 800 points are plotted.

(5d) Average readout strengths in each target direction over late trials of WMP control for each WMP session in the data of Sadtler et al. (2014), plotted as a function of the angle between the target direction and the projection of the mean population activity during reaches with the baseline decoder. A total of 8 target directions $\times$ 46 experimental sessions = 368 points are plotted.

between these two quantitites (Pearson $r = -0.66, p < .001$): the larger the angle between the target direction and the projection of $\bar{\mathbf{r}}_{\text{est}}$, the lower the maximal readout strength in that direction and under that decoder.

This is a prediction we can now test in the experimental data. If we assume subjects are re-aiming to perform cursor reaches with the baseline decoder during the initial phase of the experiment (fig. 2a), then the average population activity over these trials, $\bar{\mathbf{r}}_{\text{est}}$, should provide a good estimate of the reachable repertoire mean. We can then repeat the analysis outlined above to test for the same negative correlation between WMP maximal readout strengths in each target direction and target direction alignment with

the projection of $\bar{\mathbf{r}}_{\mathrm{est}}$ through the decoder. For each experimental session with a WMP, we measured $\bar{\mathbf{r}}_{\mathrm{est}}$ using the neural activity during the first phase of the task, and calculated its projection through the WMP used in the second phase. We then calculated the average readout strengths in each target direction over the best trials of WMP control, and plotted them as a function of target direction angle with this projection. The data over all experimental sessions are plotted in figure 5d, revealing a significant negative correlation akin to that observed in our model (Pearson $r = -0.41, p < .001$). This shows that empirical biases in individual sessions of WMP learning are consistent with those that would be expected from the geometric properties of the reachable repertoire in our model of re-aiming.

It is worth noting that this result is also consistent with, and predicted by, the neural re-association model of BMI learning postulated by [Golub et al., 2018]. In this model, a fixed repertoire of motor cortical activity patterns – termed the "neural repertoire" – is used for BMI control with the baseline decoder as well as with the WMP's. The mean activity during center-out reaches with the baseline decoder, $\bar{\mathbf{r}}_{\mathrm{est}}$, thus corresponds to the mean of the neural repertoire. Under any decoder, the readouts produced by activity patterns in the neural repertoire will therefore be biased in the direction of the projection of its mean, $\bar{\mathbf{r}}_{\mathrm{est}}$, through that decoder. Indeed, our model of re-aiming can be seen simply as a particular implementation of neural re-association, whereby the restricted motor commands restrict the motor cortical activity patterns available for BMI control to a fixed repertoire. Neural re-association, however, is a statistical model of how population activity changes after learning, and thus makes no claims as to the actual algorithms or mechanisms underlying these changes. Our model, on the other hand, makes a concrete claim about the learning algorithm being used by the subjects in this task, providing a mechanistic and normative rationale for the structure of the neural repertoire and the behavioral biases that arise from it. Moreover, it leads to natural extensions to other BMI tasks, which we turn to next.

I.1.4 Long-term BMI learning by generalized re-aiming

Although Sadtler et al. (2014) showed that non-human primates could not learn to control OMP's within the short time period of a single experimental session (a few hundred trials), a subsequent study by Oby et al. (2019) showed that such decoder perturbations could be learned over multiple days of training [Oby et al., 2019]. The authors observed that by the end of about 5-10 training sessions over multiple days, new motor cortical activity patterns emerged that allowed the subject to produce strong BMI readouts through the trained OMP. Could re-aiming play a role in the emergence of novel activity patterns over these longer timescales of learning?

In our simulations of short-term learning, we noted that re-aiming with only the two motor variables θ_1, θ_2 is not sufficient to produce the activity patterns required for OMP control. Because these two motor variables are the same ones driving motor cortex during the calibration task, the reachable repertoire under this learning strategy only contains activity patterns inside the intrinsic manifold (fig. 3a). Re-

aiming with a different set of motor variables, however, would completely change the reachable repertoire, possibly leading to the inclusion of new activity patterns outside of the intrinsic manifold. Here we consider the possibility that, over longer timescales of learning, subjects re-aim with $\widetilde{K} > 2$ motor variables, extending the set of motor variables used for re-aiming to additional variables beyond those evoked by the calibration task. We refer to this learning strategy as "generalized re-aiming". Because more motor variables need to be optimized under this strategy, it is natural that learning should be slower.

To simulate generalized re-aiming, we simply increase $\widetilde{K}$ in equation 4 and solve the resulting $\widetilde{K}$-dimensional optimization problem. Figure 6a quantifies the effectiveness of this learning strategy for OMP learning. The mean squared error achieved by the re-aiming solutions is plotted as a function of the number of motor variables used for re-aiming (the "aiming variables"), $\widetilde{K}$. As expected, we find that re-aiming with more motor variables leads to a lower mean squared error. For this model motor cortical network, about 15-20 aiming variables suffice to achieve a mean squared error as low as that achievable under WMP's with $\widetilde{K} = 2$. For other motor cortical models with different connectivity, fewer than 10 aiming variables suffice (fig. 9e). These values of $\widetilde{K}$ comfortably fit within the total number of extrinsic motor variables known to influence motor cortical activity [Thach, 1978, Fetz, 1992, Fetz, 2007, Scott, 2008, Omrani et al., 2017, Gallego et al., 2017, Willett et al., 2020]. However, they may be too high for naïve gradient-free optimization to succeed in solving equation 4 under biological limitations (e.g. on memory, motivation, and noise), which might explain why primates seem to be able to learn OMPs only under a highly structured incremental training paradigm [Oby et al., 2019].

Why the generalized re-aiming strategy works can be understood by looking at how increasing the number of aiming variables changes the reachable repertoire. A larger number of aiming variables permits a more diverse set of upstream inputs, which in turn implies that a more diverse set of activity patterns are reachable. This diversity is quantified in figure 6b by the participation ratio of the covariance over all reachable activity patterns (see Methods). The participation ratio measures the extent to which variability is spread out over many dimensions (high participation ratio) or concentrated to only a few (low participation ratio) [Gao et al., 2017]. Figure 6b shows that, as the number of aiming variables rises, the participation ratio of the reachable repertoire covariance increases; in other words, the reachable activity patterns occupy more and more dimensions of state space as $\widetilde{K}$ gets larger. That said, the participation ratio does begin to saturate at around $\widetilde{K} = 20$, reflecting the fact that the reachable repertoire is ultimately limited by the motor cortical circuit dynamics.

This expansion in the reachable repertoire leads to the inclusion of new activity patterns that are useful for OMP control. We can see this in figure 6c, which shows the readouts reachable under the same OMP visualized in figure 3diii. The readouts reachable under different values of $\widetilde{K}$ are shown by different colors, revealing how a larger number of aiming variables leads to reachable readouts closer to
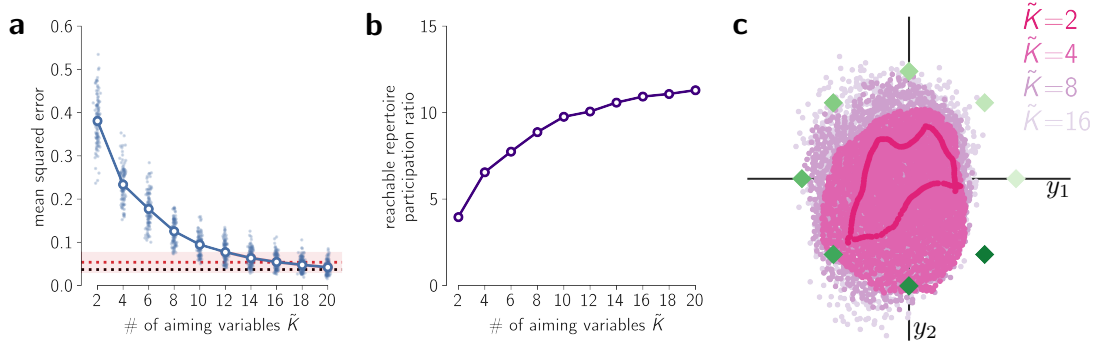
Figure 6: Generalized re-aiming yields good solutions for outside-manifold perturbations.
(6a) Mean squared error achieved by generalized re-aiming solutions to all sampled OMP's, plotted as a function of the number of motor variables used for re-aiming (the "aiming variables"), $\widetilde{K}$. Lighter blue points show the mean squared error for individual OMP's, darker points on top show the median over all sampled OMP's. For reference, dotted horizontal lines show the mean squared error achieved by re-aiming solutions with $\widetilde{K} = 2$ for the baseline decoder (black) and for WMP's (red); the red dotted line shows the median over all sampled WMP's with shading marking the upper and lower quartiles (matching the statistics of the red histogram in figure 3g).
(6b) Participation ratio of the reachable repertoire covariance as a function of the number of aiming variables (see Methods).
(6c) Readouts from the reachable repertoire through the same OMP shown in figure 3d, for different numbers of aiming variables $\widetilde{K}$. For the sake of visibility, only readouts from activity patterns genereated by motor commands with unit norm ($\|\boldsymbol{\theta}\| = 1$) are shown. The innermost ring (for $\widetilde{K} = 2$) thus corresponds to one of the rings of readouts plotted in figure 3diii.

the target readouts. As the reachable repertoire expands, more and more activity patterns occupying dimensions relevant to OMP control become reachable, ensuring the readouts reachable under this OMP cover a wider region of the workspace.

An experimental prediction of the generalized re-aiming learning strategy is that long-term BMI learning should manifest itself via a transient rise in the dimensionality of population activity. Note first that, regardless of the number of motor variables used for re-aiming, the re-aiming solutions for a 2D reaching task will always live on a low-dimensional manifold: the continuum of solutions to equation 4 for all possible 2D target readouts $\mathbf{y}^*$ necessarily lies on a manifold of intrinsic dimension 2. At the *end* of learning, motor cortical population activity should therefore remain relatively low-dimensional. But *during* learning, a much larger space of inputs may be explored en route to identifying the optimal re-aiming solutions. Consequently, motor cortical population activity should exhibit a rise in dimensionality during this time period, until it gradually coalesces to a low-dimensional manifold tailored to the geometry of the BMI decoder. This gradual refinement of motor cortical population activity is consistent with observations of long term BMI learning in primates and rodents, where it has been observed that motor cortical activity during BMI control gradually reshapes over the course of days of training in such a way that independent variability in single neuron firing rates decays as correlated variability over the population grows [Athalye et al., 2017, Athalye et al., 2018, Athalye et al., 2020]. Our results suggest that this apparently independent variability present early in learning may be a manifestation of subjects' search for generalized re-aiming solutions. An important analysis we leave for future work is to test whether a transient rise in dimensionality is present in the data of Oby et al. (2019). It is worth noting, however, that the predicted increase in dimensionality may be difficult to detect under the

incremental training paradigm used in this experiment, since population activity will converge to a low-dimensional manifold whenever a satisfactory re-aiming solution is obtained for a given incremental decoder perturbation.

I.1.5   Illusory credit assignment by generalized re-aiming

We next turn to a different class of BMI decoder perturbation, termed the credit assignment rotation perturbation [Zhou et al., 2019]. We can think of the readout from a linear BMI decoder (equation 3) as summing together the N columns of the decoding matrix $\mathbf{D}$, each one weighted by the activity of the corresponding neuron (fig. 7a). The contribution of each recorded neuron's activity to the BMI readout $\mathbf{y}$ is thus determined by its corresponding column in $\mathbf{D}$, which we refer to as its "decoding vector". Under a credit assignment rotation perturbation, the decoding vectors of a random subset of neurons (the "rotated neurons") are rotated by a given angle (fig. 7a). The contributions of each of these neurons to the BMI readout are therefore altered by this decoder perturbation. The contributions of each of the remaining "non-rotated" neurons, on the other hand, remain the same. Errors induced by this decoder perturbation can thus be corrected by selectively adjusting the responses of the rotated neurons only. But doing so requires identifying exactly which neurons' decoding vectors were rotated – a tall order given that the subject has no explicit knowledge about the BMI decoder or the few motor cortical neurons (among millions) it records from.

Despite these challenges, multiple studies have shown that non-human primates can learn to control such decoder perturbations, and that this learning is accompanied by selective changes in the responses of rotated neurons [Jarosiewicz et al., 2008, Chase et al., 2012, Zhou et al., 2019]. These studies used a center-out cursor reaching task akin to that described above (fig. 2a), in which the cursor's velocity is initially controlled by the readouts from a baseline decoder that is fit to motor cortical activity recorded during a prior calibration task. After an initial phase of cursor reaching with this decoder, the baseline decoder is perturbed by rotating a random subset of its columns, and subjects have to learn to control the cursor with this new perturbed decoder. Subjects' motor cortical activity naturally changes after learning this novel BMI decoder, and this can be quantified by comparing individual neurons' tuning to reach direction during each of the two phases of the task. Each of the aforementioned studies found that, after learning, tuning curves of both rotated (fig. 7b) and non-rotated neurons (fig. 7c) shift in a direction congruent with the rotation of the decoding vectors. For example, if the decoding vectors are rotated counter-clockwise, tuning curves shift counter-clockwise, on average. Notably, however, tuning curves of rotated neurons shift more on average than those of non-rotated neurons (compare figs. 7b and 7c). This observation has previously been interpreted to support the hypothesis that the motor system is indeed able to single out the rotated neurons, solving the so-called credit assignment problem [Minsky, 1961] by learning individual neurons' contributions to the BMI readout under each decoder
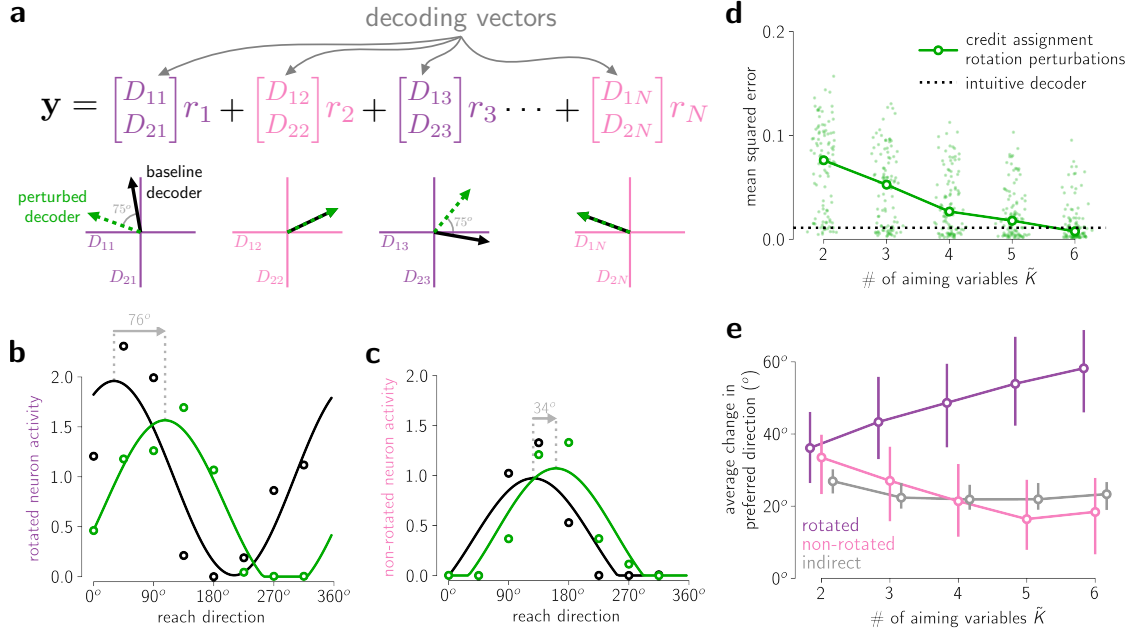
Figure 7: Generalized re-aiming solutions replicate motor cortical tuning changes observed under credit assignment rotation perturbations.

(7a) A linear BMI readout (equation 3) can be seen as summing together columns of the decoding matrix $\mathbf{D}$, each weighted by the firing rate of the corresponding neuron (the centering term $\boldsymbol{\mu}$ has been dropped here for simplicity). These columns are called the neurons' decoding vectors, and they are plotted on the axes below the equation. Under a credit assignment rotation perturbation, the decoding vectors of a subset of neurons (marked in purple) are rotated by a fixed angle (in this case, 75°). The neurons' decoding vectors under this perturbed decoder are shown by dashed green arrows. The neurons whose decoding vectors are rotated are termed "rotated" neurons (in purple), the rest of the neurons that are recorded by the BMI are termed "non-rotated" neurons (in pink). Neurons that are not recorded by the BMI (i.e. whose decoding vectors are just a vector of 0's, not depicted here) are termed "indirect" neurons.

(7b) Tuning curve of a representative example rotated neuron of our model, during cursor reaching with the baseline decoder (black) and with a credit assignment rotation perturbation (green). Each dot shows the time-averaged activity of this neuron during a simulated reach in each target direction, simulated by driving the motor cortical network with the re-aiming solutions for these targets, under the given decoder (using $\tilde{K} = 2$ for the baseline decoder and $\tilde{K} = 6$ for the perturbed decoder). Curves show tuning curves fit to these responses via linear regression (see Methods). The vertical dotted gray lines mark the preferred direction under each decoder, with an arrow labeling the change in preferred direction.

(7c) Tuning curve of a representative example non-rotated neuron of our model, under the same two decoders. All conventions exactly as in the previous panel. Note that this neuron's preferred direction changes less than that of the rotated neuron in the previous panel.

(7d) Mean squared error achieved by generalized re-aiming solutions for 100 random credit assignment rotation perturbations. Light green dots denote individual decoder perturbations, overlaid darker open dots denote medians over all 100 sampled decoder perturbations. Dotted horizontal line shows the mean squared error achieved by re-aiming solutions to the unperturbed baseline decoder with $\tilde{K} = 2$.

(7e) Average change in preferred direction of rotated, non-rotated, and indirect neurons after learning by generalized re-aiming. For each decoder perturbation, the changes in preferred direction are averaged over all neurons in each subpopulation. The lines show the median over all sampled decoder perturbations, with error bars marking the upper and lower quartiles. Positive angles indicate a counter-clockwise rotation, consistent with the direction of rotation of the decoding vectors of the rotated neurons.

[Legenstein et al., 2010]. Here we consider whether this pattern of results could instead be accounted for by generalized re-aiming, a global learning strategy entirely unconcerned with learning about individual neurons.

To simulate generalized re-aiming in this task, we simulate motor cortical activity during the calibration task, use this to construct a baseline decoder, and then sample 100 random credit assignment rotation perturbations, following closely the experimental procedures of Zhou et al. (2019) [Zhou et al., 2019] (see Methods). As in this study, each perturbed decoder is constructed by randomly rotating 50% of the columns of the decoding matrix by 75° counter-clockwise. We then simulate motor cortical activity

33

during cursor reaching with each decoder by driving the motor cortical network with the re-aiming solutions for that decoder (equation 4). By comparing neurons' directional tuning under the re-aiming solutions for the baseline decoder and under those for a given perturbed decoder, we can then predict how directional tuning would change after learning that decoder by re-aiming.

Reflecting the fact that the baseline decoder is easy to learn, we used $\widetilde{K} = 2$ to compute re-aiming solutions for it. For the perturbed decoders, we simulated generalized re-aiming with 2 to 6 aiming variables. We find that re-aiming with about $\widetilde{K} = 6$ aiming variables is necessary to achieve the same mean squared error achievable under the baseline decoder (fig. 7d). However, re-aiming with only $\widetilde{K} = 2$ aiming variables suffices to achieve a relatively low mean squared error (compare to fig. 3g), suggesting ordinary 2D re-aiming could still be a viable learning strategy for this task.

To quantify how neurons' directional tuning changes under these various re-aiming strategies, we measured neurons' preferred reach directions (i.e. the reach direction at which the tuning curve peaks, cf. figures 7b, 7c) under the re-aiming solutions for the baseline decoder and the re-aiming solutions for each perturbed decoder. For each perturbed decoder, we then calculated each neuron's change in preferred direction relative to the baseline decoder, and averaged this over all rotated and non-rotated neurons. Figure 7e shows this average change in preferred direction for all sampled perturbed decoders, as a function of the number of motor variables used for re-aiming. Consistent with the experimental results, we find that re-aiming consistently leads to a global counter-clockwise shift in motor cortical tuning curves, congruent with the direction of rotation of the decoding vectors. Importantly, we additionally find that generalized re-aiming with $\widetilde{K} > 2$ aiming variables replicates the aforementioned credit assignment effects, whereby the preferred directions of rotated neurons shift on average more than their non-rotated counterparts. This is true despite the fact that the credit assignment problem was never truly solved – no neuron-specific parameters were modified under this learning strategy.

Because we have complete access to the full population of neurons in our motor cortical model, we can also measure tuning changes in the sub-population of "indirect" neurons not recorded by the BMI (i.e. neurons whose decoding vectors in $\mathbf{D}$ comprise a vector of 0's). These are plotted in figure 7e with a gray line. Under generalized re-aiming, indirect neurons' tuning curves shift less on average than rotated neurons'. Whether they shift more or less than non-rotated neurons' tuning curves, on the other hand, depends on the specific value of $\widetilde{K}$ and varies substantially across different decoder perturbations. These two results are roughly consistent with the observations of Zhou et al. (2019) [Zhou et al., 2019]. This study found that, in two non-human primate subjects, the average change in indirect neuron tuning curves was smaller than that measured for the rotated neurons. However, this change was larger than that measured for the non-rotated neurons in one subject but smaller in the other, consistent with the variability observed in our simulations. To our knowledge, this is the only systematic study on indirect neurons' responses before and after learning a credit assignment rotation perturbation; more studies are

needed to fully test the predictions of our model.

An important additional prediction of our model is that credit assignment effects do not arise under ordinary 2D re-aiming (fig. 7e, $\widetilde{K} = 2$). This is consistent with prior modeling work showing that two-dimensional re-aiming does not suffice to account for empirically observed changes in motor cortical tuning curves [Chase et al., 2010]. Interestingly, it is also consistent with recent experimental work showing that differences between rotated and non-rotated neurons seem to arise gradually over multiple days of training [Zhou et al., 2019]. Our model suggests that this timecourse of learning might reflect a change in learning strategy, whereby subjects initially engage in low-dimensional re-aiming to rapidly reduce gross reaching errors before subsequently turning to generalized re-aiming to further refine BMI control over a longer timescale, thus giving rise to more marked credit assignment effects later in learning.

We briefly remark, however, that Zhou et al. did not observe changes in the preferred directions of non-rotated neurons after the first day of training. In our simulation, on the other hand, the preferred directions of non-rotated neurons shift back to their starting values under the baseline decoder as the number of aiming variables increases. This discrepancy between our model and the experimental data could be explained by subjects using suboptimal re-aiming solutions deviating from the optimal solution of equation 4, possibly due to the difficulty of solving this equation when $\widetilde{K} > 2$. This possibility is also consistent with the fact that the amount of change in preferred directions is significantly larger in our simulation ($30 - 60^o$) than in the experimental data ($20 - 40^o$).

I.1.6   Operant conditioning of single neurons by re-aiming

A final BMI task we consider is the operant conditioning of motor cortical neurons. In this task, the subject is rewarded for simply increasing the activity of one group of motor cortical neurons over another [Fetz, 1969, Fetz and Baker, 1973, Koralek et al., 2012, Clancy et al., 2014, Athalye et al., 2018]. Experiments showing that primates and rodents are capable of solving such tasks are often cited as evidence that the motor system can learn to specifically modulate the responses of individual neurons. Classical models of operant conditioning have argued that these changes happen via reward-modulated plasticity at their synapses [Legenstein et al., 2008]. Here we explore the extent to which these results could be explained by re-aiming instead.

We begin by investigating re-aiming solutions to the classic operant conditioning task of Fetz and Baker (1973). In this task, the subject is rewarded for increasing the firing rate of one neuron – termed the "target" neuron – while simultaneously decreasing that of another neuron – termed the "distractor" neuron. Remarkably, Fetz and Baker found that non-human primates are able to do this with only minutes of practice, regardless of which of two nearby recorded neurons is designated the target neuron. The identity of the target and distractor neurons could be flipped midway through a recording session, and the subject would subsequently adapt to this new reward contingency within tens of minutes, increasing

the activity of the neuron whose activity was previously suppressed (fig. 8a). Could low-dimensional re-aiming support the generation of these two opposing patterns of activity?

We can directly answer this question with our model by inspecting the reachable repertoire. If the reachable repertoire contains activity patterns in which neuron a is more active than neuron b as well as activity patterns in which neuron b is more active than neuron a, then good re-aiming solutions will exist for both reward contingencies. This is visualized in figure 8b, where we have plotted the firing rates of two neurons in the network across a large number of activity patterns in the reachable repertoire (following the same conventions as figure 3c for plotting the reachable repertoire). On this plane, deviations from the diagonal indicate good solutions for one of the two reward contingencies: motor commands generating activity patterns *below* the diagonal are good solutions when neuron a is the target neuron; motor commands generating activity patterns *above* the diagonal are good solutions when neuron b is the target neuron. For this particular pair of neurons, good solutions for both contingencies exist. The activity patterns generated by these solutions are marked with correspondingly colored circles.

More generally, we can think of this plane as a particular two-dimensional projection of the reachable repertoire, specified by the pair of neurons recorded in this task. A given pair of neurons thus admits good re-aiming solutions to both reward contingencies whenever the corresponding projection of the reachable repertoire covers both sides of the diagonal. Framed in this way, it is straight-forward to intuit that such solutions will generally exist for most random pairs of neurons – there is no reason to expect random two-dimensional projections of the reachable repertoire to lie exactly on the diagonal. Solving for the re-aiming solutions to 100 random pairs of neurons (see Methods), we indeed find that good solutions exist in most cases. This is shown in figure 8c, which shows the firing rates generated by the optimal re-aiming solution to each reward contingency, for each sampled pair of neurons. The marginal histograms on the top and the right of the scatter plot show the distribution of firing rates under each reward contingency, demonstrating firing rates are consistently higher when the neuron is the target than when it is the distractor.

That said, substantial variability exists – for some pairs of neurons, the optimal re-aiming solutions don't produce very high firing rates for the target neuron. These infelicitous neuron pairs are ones where the two neurons are highly correlated across all reachable activity patterns in the repertoire, such that the relevant projection hardly deviates from the diagonal and no good re-aiming solutions exist. This reveals a tight relationship between neural correlations and operant conditioning performance. Experimentally, such correlations were studied in the context of operant conditioning by Clancy et al. (2014), who trained mice on a single reward contingency of this operant conditioning task [Clancy et al., 2014]. In this experiment, "baseline correlations" between the two conditioned neurons were measured during a period of spontaneous behavior just prior to the operant conditioning task. The authors found that the degree to which individual mice successfully learned the task was in fact predictable from this baseline
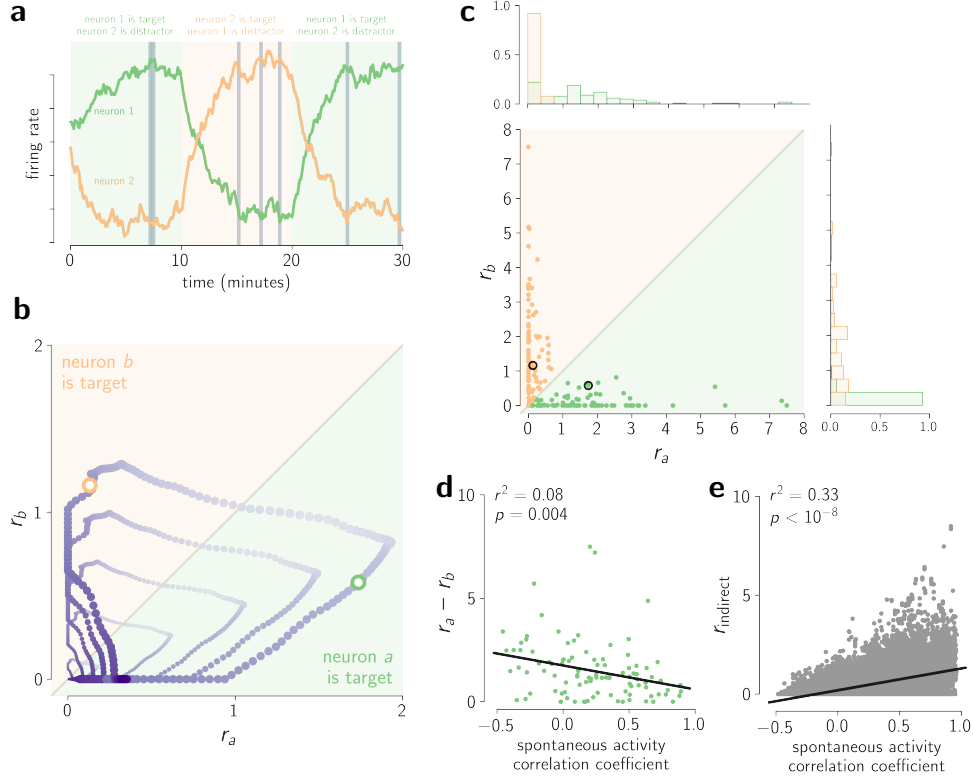
Figure 8: Operant conditioning of single neurons via re-aiming.
(8a) Caricature of operant conditioning task (these are not real data). Initially, neuron 1 is designated the "target" neuron, and the subject is rewarded when neuron 1 fires more than neuron 2. Accordingly, the subject learns to increase the firing rate of neuron 1 while simultaneously decreasing the firing rate of neuron 2, leading to reward delivery (marked by dark vertical lines). Fetz and Baker (1974) observed that non-human primates could learn to do this with only minutes of practice, and that they could appropriately adapt the firing rates of the two neurons even when the identity of the target neuron switched from neuron 1 to neuron 2 and back (transition from green to orange and back).
(8b) Activity of two neurons in activity patterns within the reachable repertoire. Each activity pattern is sized and colored according to the pair of aiming variables (θ_1, θ_2) that generated it, following the same conventions as figure 3c. In activity patterns below the diagonal, neuron a is more active than neuron b, thus providing good solutions to the reward contingency where neuron a is the target neuron; the reverse holds for activity patterns above the diagonal. The green and orange circles denote the activity patterns generated by re-aiming solutions for the two respective reward contingencies.
(8c) Activity driven by re-aiming solutions to each reward contingency, for 100 random pairs of neurons. For each pair of neurons, an orange and green point are plotted showing their activity at $t_{\mathrm{end}} = 1000$ms generated by the re-aiming solutions for each of the two reward contingencies. Following the convention in the previous panel, green points corresponds to the reward contingency in which neuron a is the target neuron and orange to the other. The axes at the top and the right show the marginal distribution of activity of neuron a and neuron b over all 100 pairs of neurons, in each reward contingency. Comparison of the green and orange histograms reveals how the activity changes from one reward contingency to the other under the re-aiming strategy.
(8d) Difference in activity (at $t_{\mathrm{end}} = 1000$ms) between the target neuron (neuron a) and the distractor neuron (neuron b) generated by the re-aiming solutions for the same 100 random pairs of neurons, plotted as a function of baseline correlation between the two neurons during simulated spontaneous behavior. Only a single reward contingency (in which neuron a is the target) is shown.
(8e) Activity of indirect neurons (at $t_{\mathrm{end}} = 1000$ms) generated by re-aiming solutions for the 100 random pairs of neurons (for the same reward contingency shown in the previous panel), plotted as a function of baseline correlation between each indirect neuron and the target neuron during simulated spontaneous behavior.

correlation: the stronger the baseline correlation, the worse the mice tended to perform the operant conditioning task. To the extent that this baseline correlation between the two neurons matches their correlation over the reachable repertoire, this observation is exactly in agreement with our model, where a stronger correlation over the repertoire leads to worse re-aiming solutions.

But under what conditions would these two correlation measurements match? One possibility is that subjects re-aim with the same motor variables that happened to be driving motor cortex during

behavior prior to the BMI task, much like we hypothesized occurs with the calibration task of Sadtler et al. (2014). Indeed, in a recent operant conditioning study with non-human primates, Engelhard et al. (2019) found that subjects seem to utilize the same patterns of activity evoked by the task they were engaged in just prior to the operant conditioning task [Engelhard et al., 2019]. If the motor variables used for re-aiming are the same ones driving motor cortical activity during the period of spontaneous behavior prior to operant conditioning, then the activity patterns measured during this period would strongly resemble those in the reachable repertoire. In this case, baseline correlations measured during spontaneous behavior would approximately match correlations over the reachable repertoire.

To quantitatively test this, we simulated motor cortical activity during spontaneous behavior by driving the model motor cortical network with random low-dimensional motor commands, constructed by assigning random values to the two aiming variables θ_1, θ_2 and fixing all other motor variables $\theta_3, \theta_4, \ldots, \theta_K$ to 0. We used these random motor commands to simulate one hundred trials of spontaneous activity and then measured the correlation between each conditioned pair of neurons over these data. Mirroring the experimental results of Clancy et al., we find that this baseline correlation is weakly but significantly predictive of re-aiming success (fig. 8d; 9% variance explained in our model vs. 8% in the experimental data, cf. supplementary figure 8a in [Clancy et al., 2014]).

Finally, we consider what happens to the remaining "indirect" neurons not included in the pair of neurons being conditioned. The activity of these indirect neurons during operant conditioning was also studied by Clancy et al., who found that, after learning, indirect neurons that had a strong baseline correlation with the target neuron remained highly active during performance of the operant conditioning task (supplementary figure 9a in [Clancy et al., 2014]). This is consistent with our model of re-aiming, in which re-aiming solutions tend to drive indirect neurons proportionally to their baseline correlation with the target neuron (fig. 8e).

Another notable observation of [Clancy et al., 2014] was that indirect neurons that showed a moderate baseline correlation with the target neuron were highly active during early stages of learning but, by the end of learning, became as silent as those with much weaker baseline correlations. This is inconsistent with our model of re-aiming: figure 8e predicts that, at the end of learning, indirect neurons with moderate baseline correlations should be more active than those with weak baseline correlations with the target neuron. A possible explanation for this inconsistency is that subjects in fact reaim with $\widetilde{K} > 2$ motor variables. Supposing that they re-aim with the same motor variables driving spontaneous behavior prior to the task, it would be hardly surprising if these encompassed more than 2 motor variables – sponteanous behavior is likely driven by independent changes in many motor variables. Here we chose $\widetilde{K} = 2$ because it allowed for a particularly efficient method for calculating re-aiming solutions and because the reachable repertoire in this case is easier to visualize. We leave for future work a more comprehensive study of $\widetilde{K}$-dimensional generalized re-aiming in operant conditioning tasks.

Discussion

Much like playing an instrument or mastering a new sport, producing movements with a brain-machine interface (BMI) constitutes a completely novel motor behavior that the motor system did not evolve to perform. Despite this, humans are able to learn to play new instruments and sports, and multiple studies have shown that multiple mammals are able to learn to control BMI's. These studies have also shown that if the BMI decoder is constructed in the right way [Gilja et al., 2012, Sadtler et al., 2014], then BMI control can in fact be learned with only hours of practice – a much shorter timescale than the years of practice it takes to learn to produce the fine movements required to play a new instrument or sport. What does this tell us about the learning strategies that might underlie BMI learning? And what might these strategies have in common with more natural forms of motor learning?

Here, we have proposed that motor cortical BMI learning proceeds via re-aiming. Under this learning strategy, internal motor variables are manipulated to control the BMI effector using the same motor cortical circuitry usually used to control the limbs. Because only a few motor variables need to be optimized for this goal, learning can proceed rapidly and flexibly. Because the motor cortical circuitry is conserved, the operation of motor cortex during natural motor control is conserved as well.

To understand the implications of this learning strategy, we formulated a mechanistic model of re-aiming in which the internal motor variables specify low-dimensional upstream inputs to motor cortex. By analyzing how these inputs get transformed into motor cortical activity patterns through the circuit's non-linear recurrent dynamics, we were able to demonstrate that re-aiming can in fact account for a wide range of experimental observations about BMI learning. This model provides a mechanistic account of the different timescales of learning required for different BMI decoders [Sadtler et al., 2014, Oby et al., 2019], selective changes in motor cortical tuning curves over learning [Jarosiewicz et al., 2008, Chase et al., 2012, Zhou et al., 2019], and the astonishing ability of mammals to flexibly modulate the activity of single neurons via operant conditioning [Fetz, 1969, Fetz and Baker, 1973, Clancy et al., 2014]. The model also makes a novel experimental prediction about behavioral biases during short-term learning of a certain class of decoders (fig. 4c), which we were able to corroborate in previously published data. Because the model is posed at the level of a neural circuit, its success at replicating these empirical phenomena allows us to understand them in terms of the biological non-linear dynamics of neural circuits.

I.2.1 Intrinsic variable learning vs. independent neuron learning

An important debate in the BMI learning literature has been whether human and non-human primates are able to precisely learn the contributions of individual neurons to BMI control. Several studies have been directly aimed at answering this question, leading to evidence in favor [Jarosiewicz et al., 2008,

Chase et al., 2010, Chase et al., 2012, Zhou et al., 2019] and against [Hwang et al., 2013, Sakellaridi et al., 2019] this "individual neuron learning" hypothesis. The alternative hypothesis is often referred to as "intrinsic variable learning" [Andersen et al., 2019], whereby subjects learn to control the BMI using the same constrained set of activity patterns usually used for natural motor control, unable to independently control individual neurons. Our model of re-aiming is an instantiation of this latter hypothesis, with the aiming variables $\theta_1, \theta_2, \ldots, \theta_{\widetilde{K}}$ acting as the so-called intrinsic variables.

Our simulations of generalized re-aiming show that many of the experimental results traditionally attributed to some form of individual neuron learning ([Legenstein et al., 2008, Legenstein et al., 2010]) can in fact can be accounted for by intrinsic variable learning. In particular, even classical single neuron operant conditioning results can be reproduced by our model. Our simulations show that recurrent dynamics within motor cortex can generate the activity patterns required by these BMI tasks, without ever needing to learn neuron-specific parameters. This suggests caution in underestimating the role of macroscopic cognitive strategies when observing what may look like highly specific microscopic changes to the activity of single neurons.

Of course, these results and conclusions by no means rule out the possibility that individual neuron learning plays an important role in BMI learning. Indeed, all the results we have replicated have been replicated in the past by various models of synaptic plasticity, in which parameters of individual neurons are explicitly learned [Legenstein et al., 2008, Legenstein et al., 2010, Wärnberg and Kumar, 2019, Engelhard et al., 2019, Feulner and Clopath, 2020]. As argued in the introduction, however, learning about individual neurons entails solving an extremely high-dimensional optimization problem with no access to explicit gradients, which could limit learning to be slow and brittle. Several of these studies worked around this problem by using small and simplified feed-forward models of motor cortex [Legenstein et al., 2010] or biologically implausible learning rules [Wärnberg and Kumar, 2019, Feulner and Clopath, 2020]. That said, a few theoretical studies have demonstrated that biologically plausible learning rules can in fact succeed in biologically relevant regimes despite these obstacles [Legenstein et al., 2008, Raman et al., 2019, Engelhard et al., 2019].

As is typically the case in complex biological phenomena like BMI learning, the true underlying mechanisms most likely comprise a mixture of both classes of hypotheses. As has been discussed at length in the literature, a variety of learning mechanisms are likely at play during motor learning of any kind [Shadmehr et al., 2010, Costa, 2011, Orsborn and Pesaran, 2017, Krakauer and Mazzoni, 2011, Athalye et al., 2020]. Many of these are likely to involve synaptic plasticity in motor cortex, which scant evidence suggests is ever switched off [Rokni et al., 2007, Peters et al., 2017]. The present study should be seen as an investigation into the properties of the solutions that can be learned by a pure re-aiming strategy. An important line of future work will be to study the solutions that can be learned by pure synaptic plasticity, so as to hopefully be able to dissociate the contributions of these two distinct learning

strategies.

Our theory may also provide a useful starting point for beginning to model and understand how re-aiming and synaptic plasticity might proceed concurrently in a coordinated and synergistic fashion. One straight-forward possibility is that synaptic plasticity operates on a much slower timescale than re-aiming. This could help explain selective changes to motor cortical responses that arise only late in learning and are not replicated by our re-aiming model. For example, Ganguly et al. (2011) observed that indirect neurons become less tuned to reach direction after days of practice with a given BMI decoder [Ganguly et al., 2011]. Jarosiewicz et al. (2008) observed a similar effect in rotated neurons after a credit assignment rotation perturbation [Jarosiewicz et al., 2008] (although note this was not replicated in subsequent experiments with such decoder perturbations [Chase et al., 2012, Zhou et al., 2019]). These selective changes in tuning strength are not reproduced by the re-aiming model presented here (fig. 12a), but can be reproduced by models of synaptic plasticity [Legenstein et al., 2010]. In the context of operant conditioning, our re-aiming model does not reproduce the experimental observation that indirect neurons moderately correlated with the target neuron become silent late in learning [Clancy et al., 2014]; it is certainly possible that concurrent synaptic plasticity would suffice to do so. Interestingly, these highly specific changes that our model does not reproduce seem to only arise late in learning, consistent with the idea that they might be brought about by a distinct learning process that operates on a slower timescale than the rapid improvements in performance that occur within a single experimental session [Kleim et al., 2004, Zhou et al., 2019].

I.2.2   What and where are the motor variables?

Our theory of re-aiming postulates the existence of a low-dimensional set of abstract motor variables that are represented in areas upstream to motor cortex. During BMI control, these motor variables can be volitionally controlled on a trial-by-trial basis to execute the desired movements, eventually converging to the learned re-aiming solutions defined in equation 4. These postulates raise three important questions.

First, what role do the motor variables play during natural motor control? In our model of motor cortex, the motor variables select and trigger particular spatiotemporal patterns of activity in the motor cortical population, which are then shaped and driven by the intrinsic recurrent dynamics of the motor cortical circuit. This is in line with modern views of motor cortex as a pattern generator, in which an upstream motor command specifies the patterns to be generated at any given time [Churchland et al., 2010, Shenoy et al., 2013, Hennequin et al., 2014, Sussillo et al., 2015, Kao et al., 2020]. In our model, the motor variables parametrize these upstream motor commands. By controlling specific dimensions of the upstream input, they can parametrically modulate the motor command to selectively engage specific modes of motor cortical operation relevant to the task at hand. Importantly, these modes of operation need not evoke overt limb movements. It is well known that motor cortical activity is not alway accompanied by

movement, such as during the preparation of a movement [Churchland et al., 2010, Kaufman et al., 2014, Vyas et al., 2018], during motor imagery [Jeannerod and Frak, 1999, Munzert et al., 2009, Rastogi et al., 2020], or during passive observation of movement [Fadiga et al., 1995, Hari et al., 1998, Tkach et al., 2007, Dushanova and Donoghue, 2010, Vigneswaran et al., 2013, Jiang et al., 2020]. It is these modes of operation that are presumably engaged by the calibration task and by the motor variables used for BMI control. This could explain how BMI control could be achieved by re-aiming without overt limb movements [Taylor et al., 2002, Carmena et al., 2003].

Under this model of natural motor control, the aiming variables corresponding to overt movement parameters will determine how these are represented in motor cortical activity during movement. In contrast to classical tuning curve models [Georgopoulos et al., 1986, Moran and Schwartz, 1999], however, this representation will generally have complex spatiotemporal structure [Churchland and Shenoy, 2007, Churchland et al., 2012, Omrani et al., 2017]. Moreover, all the aiming variables contribute non-linearly to this representation, so that the representation of movement direction, for example, could be modulated by other movement-related variables such as posture or load [Thach, 1978, Fetz, 1992, Scott, 2008]. Our simple model of upstream activity, on the other hand, really is a tuning curve model. Nevertheless, because the tuning depends non-linearly on all the aiming variables, many of these properties still hold even in that population. Complex and mixed representations of movement variables in these populations [Zhang et al., 2017] are perfectly consistent with our theory, whose strongest commitment is that these be spatiotemporally separable (even this is relaxed in the more elaborate model of re-aiming considered in chapter 2, section II.2.2).

An important limitation of this model of motor control is that it excludes the role of sensory feedback [Todorov and Jordan, 2002, Scott, 2004, Shadmehr et al., 2010]. Substantial evidence suggests that primates continuously use sensory feedback to make corrective movements in standard BMI reaching tasks [Stavisky et al., 2017, Shanechi et al., 2017], and that these corrective movements are constructed according to a forward model of the BMI cursor dynamics [Golub et al., 2015]. None of these components of BMI control can be captured by the simple open-loop model presented here. We attempt to address some of these concerns in chapter 3, where we construct a model of closed-loop feedback control in which the motor variables are continuously updated in response to sensory feedback. Because they remain low-dimensional, BMI learning remains limited much like in the open-loop model.

A second important question regards the anatomical substrates of the motor variables; that is, where in the brain the upstream inputs representing these these variables are. In primates, a good candidate area is the posterior parietal cortex, known to play an important role in processing information relevant to visually guided reaching [Mountcastle et al., 1975, Snyder et al., 1997, Batista et al., 1999, Buneo et al., 2002, Andersen and Buneo, 2002, Mulliken et al., 2008], as well as other more abstract movement-related signals [Musallam et al., 2004, Aflalo et al., 2015, Zhang et al., 2017, Andersen et al., 2019].

Importantly, parietal cortex is known to have a number of direct anatomical connections with primary motor cortex in primates [Johnson et al., 1996, Wise et al., 1997]. Another obvious candidate with strong anatomical connections to primary motor cortex is the dorsal premotor cortex (PMd) [Dum and Strick, 2002], which has been shown to contain neurons sensitive to a wide variety of movement-related variables during movement execution and preparation [Moran and Schwartz, 1999, Cisek and Kalaska, 2005, Kaufman et al., 2014, Perich et al., 2018], during observation of movements [Cisek and Kalaska, 2004], and during movement-related decisions [Wise et al., 1997, Cisek and Kalaska, 2005]. Moreover, data from simultaneous recordings of PMd and primary motor cortex indicate that changes in the inputs from PMd to primary motor cortex play an important role during short-term motor learning [Perich et al., 2018]. A final candidate area is the motor thalamus, known to be a primary driver of motor cortical activity in rodents [Sauerbrei et al., 2020] and hypothesized to play an important role in the generation of movement [Kao et al., 2020, Logiaco et al., 2019] and in learning novel motor behaviors [Logiaco and Escola, 2020, Murray and Escola, 2020].

A third and final question concerns how the re-aiming solutions are computed and stored during learning. The theory presented here only treats the question of *what* solutions subjects learn, and makes no claims about *how* they are learned and subsequently maintained. That said, a strong assumption we made in motivating the re-aiming learning strategy was that learning could operate within the low-dimensional space of the aiming variables. It is this low dimensionality that we claimed would be critical for efficient learning; if the aiming variables were learned by simply optimizing the connectivity of an upstream circuit, then the same limitations of learning that we argued apply to learning by synaptic plasticity would also apply to learning by re-aiming. One intriguing resolution to this problem would be that the aiming variables are stored and updated via the intrinsic dynamics of an upstream circuit, as in the pre-frontal cortex model of Wang et al. (2018) [Wang et al., 2018]. In this model, a recurrent neural network implicitly stores a behavioral policy in its internal state, which is adaptively updated over time by the network's internal dynamics. A similar architecture might be operating upstream of motor cortex, in which an upstream population continually stores and updates a sensorimotor policy that maps external reach targets to internal motor commands. Through instruction and practice, the dynamics of this circuit would ultimately converge to a policy implementing the optimal re-aiming solutions of equation 4. These dynamics could moreover implicitly mimic the dynamics of standard gradient-free optimization algorithms, such as the node perturbation algorithm shown by [Héliot et al., 2010] to reproduce empirical BMI learning curves. This complex learning circuit would likely encompass additional populations beyond those directly driving motor cortex, such as the basal ganglia, which are well known to play an important role in BMI learning [Koralek et al., 2012, Neely et al., 2018, Vendrell-Llopis et al., 2019, Athalye et al., 2020].

I.2.3 The "intrinsic manifold" of population activity

A simple but important takeaway from this study is that the low-dimensional structure of activity in a population depends not only on the intrinsic dynamics and connectivity within that population but also on the spatial structure of its upstream input. This is simply a mechanistic realization of the statistical theory of Gao & Ganguli (2017) [Gao et al., 2017]. The observation that population activity is confined to a low-dimensional subspace – often termed the "intrinsic manifold" [Sadtler et al., 2014, Wärnberg and Kumar, 2019, Kao and Hennequin, 2019] or the "neural modes" [Gallego et al., 2017] – thus does not necessarily indicate that the circuit connectivity prevents it from generating activity patterns outside of this subspace. It might be the case that many more activity patterns outside of this subspace are accessible, but that only a low-dimensional subset are evoked by the subjects' behavior during the recording session [Gao et al., 2017]. This may be particularly true for output areas tightly linked to behavior, like motor cortex. For these brain areas, it stands to reason that if the behavior is low-dimensional, then the upstream inputs driving it will be low-dimensional as well. In this light, the fact that motor behaviors are generally low-dimensional [d'Avella and Bizzi, 1998, d'Avella et al., 2003, Ivanenko et al., 2003, Todorov, 2004, Kutch and Valero-Cuevas, 2012, Kuppuswamy and Harris, 2014] explains why motor cortical activity recorded in the lab is often observed to be low-dimensional [Gao and Ganguli, 2015, Gallego et al., 2017], both during during BMI control [Athalye et al., 2017, Athalye et al., 2018, Golub et al., 2018] and natural motor control [Churchland et al., 2012, Kaufman et al., 2016, Gallego et al., 2018]. In short, the so-called intrinsic manifold is not just intrinsic to the circuit's structure, but also to the behavioral task evoking the activity patterns used to measure it [Gao and Ganguli, 2015, Gao et al., 2017].

This simple insight leads to a novel interpretation of the observation that substantially longer training is required to learn OMP's than WMP's [Sadtler et al., 2014, Oby et al., 2019]. Previous models of this phenomenon have assumed that the longer learning time reflects the challenge of modifying the motor cortical connectivity to allow the production of activity patterns outside the intrinsic manifold [Wärnberg and Kumar, 2019, Feulner and Clopath, 2021]. Our simulations suggest that this isn't necessary; in fact, it suffices to simply exploit additional input dimensions beyond those evoked by the calibration task (fig. 6a). Under this model of learning, the longer learning time required for OMP's reflects the fact that these new input dimensions need to be discovered from scratch, as the calibration task provides no prior information about them.

It is important to note that, to a certain extent, this particular result presupposes the specific connectivity structure we assume in our model of motor cortex. Because we use random connectivity, a rich space of activity patterns is accessible via the recurrent dynamics, so that driving the network with new input dimensions generally leads to the generation of new activity patterns. A network with low-rank connectivity, on the other hand, might not share this property [Mastrogiuseppe and Ostojic, 2018,

Wärnberg and Kumar, 2019]. We consider one such example in figure 9, where we compare simulations of re-aiming with a network with random connectivity and with a network with connectivity optimized to perform center-out reaches with a biomechanical arm model. As is often observed in networks trained to perform a single task [Schuessler et al., 2020], the optimized connectivity has low rank (fig. 9a). The network's activity is consequently constrained to a lower-dimensional subspace [Mastrogiuseppe and Ostojic, 2018], such that even re-aiming with $\widetilde{K} = 20$ aiming variables does not suffice to produce good solutions for OMP control (fig. 9e). It is important to keep in mind, however, that in reality motor cortical connectivity is likely optimized to perform a wide variety of motor behaviors, rather than a single center-out reaching task. This assumption is implicit in our choice of high-rank connectivity structure, as in several other recent models of motor cortical function [Hennequin et al., 2014, Logiaco et al., 2019, Kao et al., 2020]. We moreover find that such high-rank connectivity is necessary to produce calibration task responses with dimensionality near that observed by Sadtler et al. (2014) (fig. 9c).

Another important aspect of the intrinsic manifold that this study highlights is its non-linear structure. Because firing rates are bounded from below by 0, activity patterns are confined to the positive orthant of state space. These bounds can impose strong constraints on the structure of population activity within the intrinsic manifold, resulting in asymmetries like those present in the reachable repertoire of our model (fig. 3c). Importantly, we showed that these asymmetries can actually have behavioral repercussions for BMI control. Understanding and identifying such non-linear structure in motor cortical activity may thus prove crucial both for (1) designing better BMI decoders and (2) understanding BMI learning.

From the perspective of designing BMI decoders, it is important to realize that typical linear Gaussian models of the intrinsic manifold cannot capture the kinds of non-linear asymmetries we observed in our model. For example, classic linear dimensionality reduction methods like Principal Components Analysis and Factor Analysis assume elliptical symmetry, and are thus incapable of revealing this structure. This may limit the extent to which decoders built using such methods are able to avoid asymmetries in learning and control. In the simple case of a 2D cursor reaching task, we found that they actually were able to avoid these issues – the baseline decoder, which was built from a PCA approximation to the intrinsic manifold, was able to largely avoid the behavioral biases that WMP's suffered from (fig. 11). But this may not always be the case, particularly in more complex tasks incorporating BMI effectors with more than two degrees of freedom.

From the perspective of understanding BMI learning, our model illustrates the importance of taking into account the non-linear dynamics of motor cortex. The asymmetry of the reachable repertoire arose precisely because of the non-linear activation function we used in our recurrent network model. Linear network dynamics in fact fail to replicate the behavioral biases observed in the data, since the reachable

repertoire is elliptically symmetric in this case (fig. 14a). It was only by enforcing positive firing rates in the motor cortical dynamics that we were able to reproduce the experimental data.

Indeed, the particular structure induced by these dynamics may help explain a recent puzzling observation about WMP learning. In a re-analysis of the data of Sadtler et al. (2014), Hennig et al. (2020) observed strong transient fluctuations in motor cortical activity at behaviorally salient moments during the experiment, such as at a pause in the recording or at the moment in which the baseline decoder was perturbed [Hennig et al., 2020]. Of relevance here is that these transient fluctuations were found to have strong low-dimensional structure. They were confined to a few dimensions, termed the "neural engagement axes", that pointed in the general direction of the positive orthant of state space (i.e. they were vectors with entries of the same sign). Strong fluctuations in this direction are also present in our model of motor cortex. In figure 5b, for example, it can be easily appreciated that the calibration task responses are highly spread out in the direction of the reachable repertoire mean – a vector of non-negative entries that points in the direction of the positive orthant of state space. The high variability in this direction is due to the non-linear dynamics imposed by the positivity constraint on the firing rates. This suggests that the particular structure of the neural engagement axes measured by Hennig et al. (2020) may be a consequence of the non-linear dynamics of motor cortex. Hennig et al. (2020) moreover went on to show that fluctuations along the neural engagement axes had behavioral consequences, and that this affected how quickly particular reach directions could be learned. This result further underscores the role that non-linear dynamics might play in shaping BMI learning.

Because no changes are made to the motor cortical circuit, a key prediction of our model of re-aiming is that the intrinsic manifold measured during the calibration task should remain the same after learning. Unfortunately, the calibration task was not repeated at the end of learning in the experiments of [Sadtler et al., 2014] or [Oby et al., 2019], so this cannot be tested with existing data. However, in both of these experiments subjects did perform a "washout" block at the end of learning, where they returned to controlling the unperturbed baseline decoder. The authors found that population activity during this block was in fact changed relative to the initial pre-learning block of baseline decoder control, reflecting a "memory trace" of the decoder perturbation learned in between these two blocks of baseline decoder control [Losey et al., 2020]. This observation can be accomodated in our model through a change in the re-aiming solutions used for baseline decoder control. As discussed in the context of generalized re-aiming, many motor-relevant variables are likely represented in the upstream inputs to motor cortex. One of these may reflect a kind of "context" signal that could change after learning a new sensorimotor mapping [Heald et al., 2020], thus giving rise to a memory trace of the newly learned BMI decoder. An important future line of work will be to investigate what changes in the aiming variables could give rise to the particular changes in population activity observed in the washout block.

I.2.4 The role of the calibration task in BMI design

From a more practical perspective, our theory of re-aiming highlights the role of the calibration task in BMI design. The calibration task is typically seen as a way to calibrate the decoding parameters; that is, as a source of information for optimizing the BMI *decoder*. Here we suggest that it additionally serves as a source of information for the subject herself, for optimizing the BMI *learner* [Shenoy and Carmena, 2014, Perdikis and Millan, 2020]. For example, in modeling short-term learning of WMP's we assumed that subjects re-aimed by optimizing the two motor variables modulated by the calibration task. If any other two motor variables had been optimized instead, the re-aiming solutions would not have succeeded in solving the task – the prior information provided by the calibration task allows efficient learning by telling the subject which motor variables to re-aim with.

This predicts that careful design of the calibration task may help optimize subjects' learning. For example, the calibration task should evoke changes in as few motor variables as necessary, so that subjects subsequently re-aim by optimizing only those $\widetilde{K} < K$ motor variables and avoid wasting time exploring modifications to other motor variables. Another important factor to consider is the relationship of these motor variables to natural motor control. For example, if the calibration task requires making overt manual movements [Jarosiewicz et al., 2008, Ganguly and Carmena, 2009, Gilja et al., 2012, Shenoy and Carmena, 2014], then re-aiming with the motor variables it evokes might not allow for BMI control without simultaneous overt movements. Learning might thus be slower in this case, as the subject would have to resort to identifying new motor variables that don't evoke overt movement.

It is also important to note that other inductive biases unrelated to the calibration task may play a role in determining which motor variables are used for re-aiming. For example, because primates are accustomed to using their arms to move objects in front of their face, they might automatically resort to re-aiming with arm-related motor variables to move a BMI cursor on a screen directly in front of them. Such prior biases might in fact be particularly strong in the BMI learning studies reviewed here, in which the non-human primate subjects typically had years of experience learning and using BMIs. More generally, these considerations suggest that learning might be improved by ensuring the calibration task provides information complimentary to the subject's prior biases.

These insights into the role of the calibration task in BMI learning provide a novel perspective on the classic "two-learners" problem in BMI design [Orsborn et al., 2014, Shenoy and Carmena, 2014, Merel et al., 2015, Perdikis and Millan, 2020]. In our model, the calibration task drives both the experimenter's (or BMI designer's) "learning" of a good decoder and in the subject's (or user's) "learning" of good motor commands for controlling that BMI decoder. In the experiments simulated here, the decoder is learned only from the calibration task data and then fixed during the user's learning. But recent studies have aimed to close this loop by continuously adapting the BMI decoder while the user is learning,

with promising results [Orsborn et al., 2014, Shenoy and Carmena, 2014, Silversmith et al., 2020]. Our model provides a theoretical framework for modeling how such closed-loop decoder adaptation might improve learning by facilitating the acquisition of re-aiming solutions. As opposed to previous such models [Merel et al., 2015], ours incorporates mechanistic principles of neural circuit dynamics. Using our model to understand how these might influence the effects of closed-loop decoder adaptation is a fascinating and important direction for future research.

I.2.5   Re-aiming and naturalistic sensorimotor learning

We can think of re-aiming as learning a sensorimotor map: a mapping from externally determined sensory goals, $\mathbf{y}^*$, to internally generated motor commands, $\hat{\boldsymbol{\theta}}(\mathbf{y}^*)$. Viewed in this way, it stands to reason that the algorithms by which subjects learn the re-aiming solutions in equation 4 might be the same ones underlying sensorimotor learning more generally. Studies of sensorimotor learning might thus provide important insights into the algorithms underlying BMI learning by re-aiming.

Of particular relevance to the BMI tasks considered here are classic studies of sensorimotor adaptation, where subjects have to overcome a force-field or visuomotor transformation in reaching towards a target [Shadmehr et al., 2010, Krakauer and Mazzoni, 2011]. Learning in these motor adaptation tasks often occurs on a timescale of 100's of trials in both primates [Li et al., 2001, Perich et al., 2018, Sun et al., 2020] and rodents [Mathis et al., 2017], similar to the timescales of learning observed in short-term BMI learning tasks (e.g. WMP learning). Neural recordings during learning in motor adaptation tasks show motor cortical tuning curve shifts largely in agreement with what would be predicted by a re-aiming strategy [Li et al., 2001, Sun et al., 2020], and suggest that these shifts arise from changes to the preparatory input from dorsal pre-motor cortex to primary motor cortex [Perich et al., 2018]. Indeed, changes in the preparatory state of motor cortex (presumably set by upstream inputs [Hennequin et al., 2014, Sussillo et al., 2015, Kao et al., 2020]) have been shown to play a critical role in adaptation to visuomotor rotations both under manual and BMI control [Vyas et al., 2018, Vyas et al., 2020]. These results are consistent with the idea that inputs to motor cortex are modified during learning while motor cortex itself remains unchanged, much like in our theory of re-aiming.

So what do we know about the algorithms underlying sensorimotor adaptation? A large body of work on sensorimotor adaptation in humans points to the existence of multiple learning processes [Smith et al., 2006, Mazzoni and Krakauer, 2006, Taylor et al., 2014]. Of relevance to re-aiming is the role of explicit cognitive strategies [Taylor and Ivry, 2012, McDougle et al., 2016], which seem to drive the faster timescales of learning [McDougle et al., 2015]. Such strategies generally consist of the manipulation of a few task-relevant abstract variables. For example, using a 2D reaching task with visuomotor rotations, McDougle & Taylor (2019) showed that humans' behavior during learning is consistent with a low-dimensional strategy in which subjects memorize corrective movements in the 2D

workspace while simultaneously learning a parametric map from 2D reach directions to 2D movements [McDougle and Taylor, 2019].

Such low-dimensional learning strategies are consistent with our model of BMI learning, and evidence exists that explicit cognitive strategies play an important role during BMI learning. Sakellaridi et al. (2019) reported the use of an explicit re-aiming strategy by a tetraplegic human learning to control a BMI [Sakellaridi et al., 2019]. Additionally, fMRI studies suggest that learning to volitionally modulate activity in a target brain area is only possible when subjects are instructed to use an explicit strategy tailored to that brain area [Scharnowski et al., 2015, Sitaram et al., 2017]. In non-human subjects, the role of certain strategies could be dissociated by examining differences in learning of different kinds of decoder perturbations. For example, some WMP's may be much easier to learn than others: a WMP inducing a pure visuomotor rotation on the readouts from the baseline decoder is likely much easier to learn than the WMP's in figure 4a, which induce much more complex visuomotor transformations.

In the wider motor learning literature, a distinction is often made between sensorimotor adaptation and skill learning, whereby the latter happens on much longer timescales than the former [Krakauer and Mazzoni, 2011, Krakauer et al., 2019]. The correspondence between our model of re-aiming and sensorimotor adaptation suggests a simple hypothesis about the mechanistic underpinnings of this distinction: sensorimotor adaptation is implemented through modifications of internal motor variables whereas skill learning is implemented through synaptic plasticity within motor cortex. As already mentioned above, an important future line of work will be to flesh out the different behavioral predictions these two models of learning make, and whether these match up with the behavioral signatures of sensorimotor adaptation and skill learning.

Methods

I.3.1   Motor cortical dynamics

The motor cortical network model was simulated by integrating equation 1 using a standard 4th order Runge-Kutta method with step size 0.1ms, implemented with the torchdiffeq Python package [Chen et al., 2018]. In all simulations presented above, only 10% of the recurrent weights were set to non-zero values, which were independently sampled from a zero-mean Gaussian, $\mathcal{N}(0, 1/N)$. Input weights were all sampled from a zero-mean Gaussian, $W_{ij}^{\text{in}} \sim \mathcal{N}(0, 1/M)$. Encoding weights were sampled randomly from the standard Gaussian distribution, $U_{ij} \sim \mathcal{N}(0, 1)$ (any normalization is taken care of by the metabolic cost term in 4 when computing re-aiming solutions). Other connectivities are considered in figure 9. We used $\tau = 200$ms, as in Hennequin et al. (2014) [Hennequin et al., 2014]. To enable efficient numerical simulation, network size was set to $N = M = 256$. We found that simulations with larger networks (up

to 2048 neurons) produced similar results.

I.3.2 Computing re-aiming solutions

Concatenating the aiming variables into a $\widetilde{K}$-dimensional vector that contains only the motor variables being optimized, $\widetilde{\boldsymbol{\theta}} = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_{\widetilde{K}} \end{bmatrix}$, we can treat equation 4 as an optimization over all $\widetilde{K}$-dimensional vectors $\widetilde{\boldsymbol{\theta}}$ in $\mathbb{R}^{\widetilde{K}}$. Under the assumptions of our model of motor cortical dynamics, however, we can simplify this problem by analytically solving for the optimal magnitude of $\widetilde{\boldsymbol{\theta}}$, $\|\widetilde{\boldsymbol{\theta}}\|$, given its direction. Once we have this optimal magnitude, all that remains is an optimization over its direction – an optimization over unit vectors on the $\widetilde{K}$-dimensional unit hypersphere. This is a $(\widetilde{K}-1)$-dimensional manifold that, importantly, is bounded, so we can hope to find the optimal direction efficiently by brute force search, avoiding the difficulties of non-convex gradient-based optimization in high dimensions (e.g. local optima).

More formally, we decompose the re-aiming optimization (equation 4) into an optimization of the norm s and direction $\widetilde{\boldsymbol{\theta}}_0$ of the aiming vector $\widetilde{\boldsymbol{\theta}}$,

$$\hat{\boldsymbol{\theta}}(\mathbf{y}^*) = \hat{s} \begin{bmatrix} \hat{\widetilde{\boldsymbol{\theta}}}_0 \\ \mathbf{0} \end{bmatrix}$$

$$\hat{s}, \hat{\widetilde{\boldsymbol{\theta}}}_0 = \underset{s>0,\, \|\widetilde{\boldsymbol{\theta}}_0\|=1}{\arg\min}\ \left\| \mathbf{y}\left(t_{\text{end}}; s\widetilde{\boldsymbol{\theta}}_0\right) - \mathbf{y}^* \right\|^2 + \frac{\gamma}{M} \sum_{i=1}^{M} u_i(s\widetilde{\boldsymbol{\theta}}_0)^2, \tag{9}$$

The vector of zeros in the first equation denotes the values of all the motor variables not used for re-aiming $(\theta_{\widetilde{K}+1}, \theta_{\widetilde{K}+2}, \dots, \theta_K)$, which in our simulations are always fixed to 0. We next proceed to analytically solve for the optimal magnitude $\hat{s}$ by exploiting two simplifications afforded to us by the rectified linear activation function $\phi(\cdot)$ of the motor cortical dynamics. The first is the scale-invariance of linear rectification ($\phi(sx) = s\phi(x)$ for any $s \geq 0$), which accordingly endows the motor cortical dynamics with scale invariance,

$$\mathbf{r}(t; s\boldsymbol{\theta}) = s\mathbf{r}(t; \boldsymbol{\theta}), \quad s > 0 \tag{10}$$

It is straight forward to prove that this holds true whenever $x_i(0) = 0$, which we assumed to be the case in our simulations. The second simplification is to approximate the quadratic cost term by its large M limit

$$\frac{1}{M} \sum_{i=1}^{M} u_i(\boldsymbol{\theta})^2 \approx \lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} u_i(\boldsymbol{\theta})^2 \approx \frac{\|\boldsymbol{\theta}\|^2}{2}. \tag{11}$$

The equality holds whenever the encoding weights are randomly and independently distributed with zero mean and unit variance, as they are here, so that the central limit theorem can be invoked to replace the sum with an expectation (the factor of 1/2 arises from the fact that only half of each axis counts towards the expectation due to the rectification). Plugging these two equations into equation 9 together

with the BMI readout equation 3, we obtain

$$\hat{s}, \hat{\widetilde{\boldsymbol{\theta}}} = \underset{s>0,\, \|\widetilde{\boldsymbol{\theta}}\|=1}{\arg\min} \; \left\| s\mathbf{Dr}\big(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}\big) - \mathbf{D}\boldsymbol{\mu} - \mathbf{y}^* \right\|^2 + \frac{\gamma}{2} s^2. \tag{12}$$

It is then straight forward to solve for $\hat{s}$ in terms of $\hat{\widetilde{\boldsymbol{\theta}}}$, yielding the following closed set of equations fully specifying the re-aiming solution

$$\hat{s}(\widetilde{\boldsymbol{\theta}}) = \phi\left(\frac{(\mathbf{D}\boldsymbol{\mu} + \mathbf{y}^*) \cdot \mathbf{Dr}\big(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}\big)}{\left\| \mathbf{Dr}\big(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}\big) \right\|^2 + \frac{\gamma}{2}} \right) \tag{13a}$$

$$\hat{\widetilde{\boldsymbol{\theta}}} = \underset{\|\widetilde{\boldsymbol{\theta}}\|=1}{\arg\min} \; \left\| \hat{s}(\widetilde{\boldsymbol{\theta}})\mathbf{Dr}\big(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}\big) - \mathbf{D}\boldsymbol{\mu} - \mathbf{y}^* \right\|^2, \tag{13b}$$

where the notation $\cdot$ denotes the standard Euclidean dot product. We have thus reduced what was an optimization over all vectors in $\mathbb{R}^{\widetilde{K}}$ (equation 4) to an optimization over all vectors living on the $\widetilde{K}$-dimensional unit hypersphere (equation 13b).

We can solve this new optimization problem by brute force search, by simply uniformly sampling a large number of vectors $\widetilde{\boldsymbol{\theta}}$ on the hypersphere and identifying the one that produces the smallest value of the right-hand side of equation 13b. Evaluating $\mathbf{r}(t_{\text{end}}, \widetilde{\boldsymbol{\theta}})$ for a large number of different $\widetilde{\boldsymbol{\theta}}$ vectors can be done efficiently by using a GPU to simulate in parallel the recurrent network dynamics driven by each of the sampled motor commands. Note that, once these activity patterns have been calculated, the brute force search can be performed for any value of γ, without having to again simulate the dynamics.

For simulations with $\widetilde{K} = 2$, this sufficed to produce good re-aiming solutions. In this case, the relevant hypersphere is the unit circle, from which it is straight forward to sample densely and uniformly. For simulations of generalized re-aiming, however, we took an additional step to ensure the obtained solutions were as good as they could be even for the larger values of $\widetilde{K}$ where it becomes more difficult to sample densely from the corresponding unit hypersphere. We first performed a brute force search over 2^{17} vectors sampled uniformly from the unit hypersphere, as was just described, to obtain an initial estimate of the re-aiming solution. This initial estimate was then used as a starting point for the L-BFGS algorithm [Byrd et al., 1995], which was applied to directly optimize the raw aiming variables $\theta_1, \theta_2, \ldots, \theta_{\widetilde{K}}$. We found that this additional step was essential for values of $\widetilde{K}$ greater than around 10.

In all simulations in the results section, we used an endpoint time of $t_{\text{end}} = 1000$ms, reflecting the typical time it takes for trained primates to complete center-out reaches under BMI control [Sadtler et al., 2014, Jarosiewicz et al., 2008]. The results of simulations with other endpoint times are shown in the supplementary results section (fig. 9b). To emulate a center-out reaching task, the target readouts $\mathbf{y}^*$ were set

to eight equally spaced unit vectors on the unit circle (cf. fig. 3d). Mean squared error is calculated as

$$\text{mean squared error} = \frac{1}{8} \sum_{i=1}^{8} \left\| \mathbf{y}\left(t_{\text{end}}; \hat{\boldsymbol{\theta}}(\mathbf{y}_i^*)\right) - \mathbf{y}_i^* \right\|^2, \tag{14}$$

where $\mathbf{y}_1^*, \mathbf{y}_2^*, \ldots, \mathbf{y}_8^*$ correspond to the eight radial target readouts. In the case of operant conditioning, there is no "target readout" per se, as subjects are simply instructed to modulate firing rates as much as possible in a given direction. In this case, a different re-aiming objective was used, see the section "Simulation of operant conditioning" for details.

I.3.3   Setting the metabolic cost weight

The metabolic cost weight parameter γ was picked to ensure that low mean squared error would be achieved under the baseline decoder with $\widetilde{K} = 2$. Specifically, we calculated re-aiming solutions with $\widetilde{K} = 2$ for the baseline decoder under a wide and dense range of values of γ. We took advantage of the fact that the brute-force search algorithm outlined above allows us to easily evaluate solutions for different values of γ with only a single forward pass of the model. Once we had re-aiming solutions for each value of γ, we calculated the error achieved by these re-aiming solutions for each target readout, and found the largest value of γ that permitted an error of less than .05 for all eight targets.

I.3.4   Characterizing the reachable repertoire

The reachable repertoire is the set of activity patterns at time t_{end} that can be generated by any motor command admissible under the re-aiming strategy. We assume that these admissible motor commands are bounded, reflecting the fact that (i) actual extrinsic motor variables are finite and bounded and (ii) upstream firing rates are bounded. Formally, we enforce this by assuming an upper bound on the motor command norm, $\|\boldsymbol{\theta}\| \leq s_{\text{max}}$. In our simulations of short-term learning of WMP's and OMP's, we set the value of this bound to the maximum norm over all 2D re-aiming solutions to this equation. Specifically, we computed re-aiming solutions with $\widetilde{K} = 2$ to each target readout for all decoder perturbations (8 target readouts × (100 WMP's + 100 OMP's + baseline decoder) = 1,608 re-aiming solutions), calculated their norms, and set s_{max} to their maximum. For the randomly connected network presented above, we found this value to be approximately 1.25.

In figure 3c, we drove the motor cortical network with motor command vectors with five distinct norms, $\|\boldsymbol{\theta}\| \in \{0.1, 0.4, 0.7, 1.0, s_{\text{max}}\}$ (see fig. 3b), chosen to aid visualization of the reachable repertoire. These were constructed by picking 256 2D vectors equally spaced on the unit circle, uniformly re-scaling them by each norm, and using these values to set the two aiming variables θ_1, θ_2 (all other motor variables were set to zero). We then simulated the motor cortical network with each of these motor commands, resulting in $5 \times 256 = 1,280$ activity patterns, $\mathbf{r}(t_{\text{end}}; \boldsymbol{\theta})$. We then applied Principal Components Analysis

(PCA) to this ensemble of activity patterns, and projected them onto the top three PC's to obtain the visualization in figure 3c. The exact same projection was used to simultaneously visualize the calibration task activity in figure 5b. These same activity patterns were used in figure 3d, where their readouts through three different decoders are plotted. In figure 4a, we used only four motor command norms (thus producing four loops of readouts), equally spaced between 0.1 and s_{max}. In figure 8b, five norms equally spaced between .1 and the maximum norm of the re-aiming solutions for that pair of neurons were used. These choices were all made to aid visualization of the reachable repertoire's structure.

To obtain the mean and covariance of the reachable repertoire for $\widetilde{K} = 2$ (figs. 5b and 3a, respectively), we computed expectations over a uniform distribution on the reachable repertoire manifold. Note that this is not the same as an expectation over activity patterns generated by uniformly distributed aiming variables. Rather, these two distributions of activity patterns are related via the Jacobian of the mapping from aiming variables, $\widetilde{\boldsymbol{\theta}}$, to activity patterns, $\mathbf{r}(t_{\text{end}}; \widetilde{\boldsymbol{\theta}})$. In section I.4.4, we derive the structure of this Jacobian matrix in terms of the polar coordinates of the two aiming variables,

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = s \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix}. \tag{15}$$

Defining $\mathbf{r}_0(\varphi)$ as the activity pattern at time t_{end} generated by a pair of aiming variables θ_1, θ_2 with angle φ and unit norm $s = 1$, we obtain the following expressions for the mean $\bar{\mathbf{r}}$ and covariance $\boldsymbol{\Sigma}_r$ of the uniform distribution on the reachable repertoire manifold,

$$\bar{\mathbf{r}} = \frac{2}{3} s_{\text{max}} \frac{\int_0^{2\pi} \mathbf{r}_0(\varphi) \|\mathbf{r}_0(\varphi)\| \|\mathbf{r}_0'(\varphi)\| \sin \omega(\varphi) \, \mathrm{d}\varphi}{\int_0^{2\pi} \|\mathbf{r}_0(\varphi)\| \|\mathbf{r}_0'(\varphi)\| \sin \omega(\varphi) \, \mathrm{d}\varphi} \tag{16}$$

$$\boldsymbol{\Sigma}_r = \frac{1}{2} s_{\text{max}}^2 \frac{\int_0^{2\pi} \mathbf{r}_0(\varphi) \mathbf{r}_0(\varphi)^T \|\mathbf{r}_0(\varphi)\| \|\mathbf{r}_0'(\varphi)\| \sin \omega(\varphi) \, \mathrm{d}\varphi}{\int_0^{2\pi} \|\mathbf{r}_0(\varphi)\| \|\mathbf{r}_0'(\varphi)\| \sin \omega(\varphi) \, \mathrm{d}\varphi} - \bar{\mathbf{r}} \bar{\mathbf{r}}^T, \tag{17}$$

where $\mathbf{r}_0'(\varphi)$ denotes the partial derivative with respect to φ and $\omega(\varphi)$ denotes the angle between $\mathbf{r}_0(\varphi)$ and $\mathbf{r}_0'(\varphi)$. Because we have no analytical access to $\mathbf{r}_0(\varphi)$, we performed these integrals numerically and used a finite-differences approximation for the derivative $\mathbf{r}_0'(\varphi)$. This is how the reachable repertoire mean plotted in figure 5b (and its projection in figures 4a and 4b) was calculated. This estimate of its covariance was used for figure 3a (see equation 28).

Analagous calculations for $\widetilde{K} > 2$ quickly become numerically intractable, as a dense sample of activity patterns from the reachable repertoire are required to numerically estimate the derivatives and integrals. We thus chose to characterize the dimensionality of the reachable repertoire under generalized re-aiming by the covariance over activity patterns generated by uniformly distributed motor commands, $\boldsymbol{\Sigma}_\theta$. As noted, this is not the same as the covariance over activity patterns uniformly distributed on the reachable repertoire manifold, but these two covariances are strongly related. To calculate the covariance

over activity patterns generated by uniformly distributed motor commands, we first re-parameterized the aiming variables $\widetilde{\boldsymbol{\theta}} = s\widetilde{\boldsymbol{\theta}}_0$ in terms of a norm $s \geq 0$ and a unit vector $\widetilde{\boldsymbol{\theta}}_0 \in \mathbb{R}^{\widetilde{K}}$. Assuming a uniform distribution over s and $\widetilde{\boldsymbol{\theta}}_0$, we have that this covariance is given by

$$
\begin{aligned}
\boldsymbol{\Sigma}_\theta &= \left\langle \left\langle \mathbf{r}(t_{\mathrm{end}}; s\widetilde{\boldsymbol{\theta}}_0)\mathbf{r}(t_{\mathrm{end}}; s\widetilde{\boldsymbol{\theta}}_0)^T \right\rangle_s \right\rangle_{\widetilde{\boldsymbol{\theta}}_0} - \left\langle \left\langle \mathbf{r}(t_{\mathrm{end}}; s\widetilde{\boldsymbol{\theta}}_0) \right\rangle_s \right\rangle_{\widetilde{\boldsymbol{\theta}}_0} \left\langle \left\langle \mathbf{r}(t_{\mathrm{end}}; s\widetilde{\boldsymbol{\theta}}_0) \right\rangle_s \right\rangle_{\widetilde{\boldsymbol{\theta}}_0}^T \\
&= \left\langle s^2 \right\rangle_s \left\langle \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)^T \right\rangle_{\widetilde{\boldsymbol{\theta}}_0} - \left\langle s \right\rangle_s^2 \left\langle \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) \right\rangle_{\widetilde{\boldsymbol{\theta}}_0} \left\langle \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) \right\rangle_{\widetilde{\boldsymbol{\theta}}_0}^T \\
&= \frac{s_{\mathrm{max}}^2}{3} \left\langle \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)^T \right\rangle_{\widetilde{\boldsymbol{\theta}}_0} - \frac{s_{\mathrm{max}}^2}{4} \left\langle \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) \right\rangle_{\widetilde{\boldsymbol{\theta}}_0} \left\langle \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) \right\rangle_{\widetilde{\boldsymbol{\theta}}_0}^T,
\end{aligned} \tag{18}
$$

where we used the scale invariance of the motor cortical dynamics (equation 10) to write $\mathbf{r}(t_{\mathrm{end}}; s\widetilde{\boldsymbol{\theta}}_0) = s\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)$ in the second line, and in the third line we simply plugged in expressions for the first and second moments of s under a uniform distribution between 0 and s_{max}. The expectations in the last line were estimated numerically by sampling 2^{17} aiming variable vectors $\widetilde{\boldsymbol{\theta}}_0$ uniformly from the $\widetilde{K}$-dimensional hypersphere with unit radius, calculating the activity patterns $\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)$ generated by these aiming variables at time $t_{\mathrm{end}} = 1000\mathrm{ms}$, and then calculating sample averages over the resulting ensemble of activity patterns. In figure 6c, the readouts from a random subset of these activity patterns are shown. In figure 6b, the participation ratio of this covariance is plotted as a function of $\widetilde{K}$,

$$
\text{participation ratio of } \boldsymbol{\Sigma}_\theta = \frac{\mathrm{Tr}[\boldsymbol{\Sigma}_\theta]^2}{\mathrm{Tr}[\boldsymbol{\Sigma}_\theta^2]} = \frac{\left(\sum_{i=1}^N \lambda_i\right)^2}{\sum_{i=1}^N \lambda_i^2} \tag{19}
$$

where λ_i are the eigenvalues of $\boldsymbol{\Sigma}_\theta$.

I.3.5 Predicting behavioral bias with maximal readout strength

To quantify biases in readouts, we used the readout strength metric $\rho(\mathbf{y}; \mathbf{y}^*)$ defined in equation 5, with the target readout vector $\mathbf{y}^*$ enforced to be unit norm. This metric is exactly equivalent to the cursor progress metric used by Golub et al. (2018); we have simply renamed it for a clearer exposition in the context of our notation and terminology. This metric was chosen for consistency with previous analyses of the data of Sadtler et al. ([Golub et al., 2018, Hennig et al., 2020]). By depending both on the angle and norm of the readout direction $\mathbf{y}$ (through the Euclidean dot product), the readout strength $\rho(\mathbf{y}; \mathbf{y}^*)$ measures the extent to which the readout $\mathbf{y}$ points in the same direction as the target readout $\mathbf{y}^*$. Readout strength is highest when the readout (i) points in the same direction as the target and (ii) has a large norm. The readouts with highest readout strength in a given target direction are thus those with largest norm that point in that direction (cf. fig. 4a).

We used the reachable readouts with highest readout strengths in each target direction to quantify biases imposed by the reachable repertoire. We termed the readout strengths of these readouts the

maximal readout strengths (used in figures 4b, 4c, and 5c), formally defined as

$$\rho^{\mathrm{max}}(\mathbf{y}^*) = \max_{\|\widetilde{\boldsymbol{\theta}}\| \leq s_{\mathrm{max}}} \rho\left(\mathbf{y}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}); \mathbf{y}^*\right), \tag{20}$$

where $\widetilde{\boldsymbol{\theta}} \in \mathbb{R}^2$ is a vector containing the aiming variables θ_1, θ_2 (recall that $\widetilde{K} = 2$ in these simulations). The inequality constraint enforces that the maximization is only over readouts from activity patterns within the reachable repertoire (see section "Characterizing the reachable repertoire" for how s_{max} is set). This maximization problem was solved by again exploiting the same re-parameterization of the aiming variables used for calculating re-aiming solutions (equation 9). Specifically, we decompose the vector of aiming variables $\widetilde{\boldsymbol{\theta}}$ into a magnitude and direction $\widetilde{\boldsymbol{\theta}} = s\widetilde{\boldsymbol{\theta}}_0$ (where $\|\widetilde{\boldsymbol{\theta}}_0\| = 1$). Plugging in the readout equation (equation 3) into the definition of the readout strength (equation 5) and exploiting the scale invariance of the motor cortical dynamics (equation 10), we derive that the maximal readout strength is given by

$$\rho^{\mathrm{max}}(\mathbf{y}^*) = \max_{\|\widetilde{\boldsymbol{\theta}}_0\| = 1} \left\{ \mathbf{D}\left(\hat{s}_\rho(\widetilde{\boldsymbol{\theta}}_0)\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) - \boldsymbol{\mu}\right) \cdot \mathbf{y}^* \right\}, \tag{21}$$

where

$$\hat{s}_\rho(\widetilde{\boldsymbol{\theta}}_0) = \operatorname*{arg\,max}_{0 < s < s_{\mathrm{max}}} \left\{ s\mathbf{D}\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) \cdot \mathbf{y}^* \right\} = \begin{cases} s_{\mathrm{max}} & \text{if } \mathbf{D}\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) \cdot \mathbf{y}^* > 0 \\ 0 & \text{else} \end{cases}. \tag{22}$$

The optimization in equation 21 was performed by brute force search; recall that $\widetilde{\boldsymbol{\theta}}_0$ is just a 2D unit vector (since $\widetilde{K} = 2$ in these simulations), so that the brute force search is simply over vectors on the unit circle.

I.3.6 Re-analysis of data from Sadtler et al. (2014)

To measure maximal readout strengths in the data of Sadtler et al., we first identified the trials with largest readout strengths in each session of WMP control. Specifically, for each session we identified the 50 adjacent trials in which the center-out reaches were completed in the shortest amount of time. This provided a window of peak WMP control performance during that session. We then calculated the readout strength in the direction of the target at each timestep in each trial. Note that the direction of the target can change continuously over time as the cursor moves, so we binne together target directions in 45° bins centered at each of the eight reach target angles. This resulted in eight bins of readout strengths, across many timesteps and trials. The readout strengths in each bin were then averaged, to produce readout strength curves for each session, analogous to those plotted in figure 4b. These readout strength curves were then horizontally centered at their peaks and averaged to obtain the light dashed curve in figure 4c. The error bars in this figure reflect the standard error of the mean. Data from a total of 46 sessions were used.

For figure 5d, the average readout strength in each bin for each session is plotted as a function of the projection of $\bar{\mathbf{r}}_{\mathrm{est}}$ through the WMP decoder from that session. The estimate of the reachable repertoire mean, $\bar{\mathbf{r}}_{\mathrm{est}}$, was measured from motor cortical activity during the first phase of the task, in which the subject performed center-out cursor reaches with the baseline decoder. We calculated target-specific means by averaging motor cortical activity over all trials and time during reaches to each target, and then averaged these target-specific means over targets to obtain $\bar{\mathbf{r}}_{\mathrm{est}}$. In the case of the theoretical model, $\bar{\mathbf{r}}_{\mathrm{est}}$ was measured by simulating reaches to each target under the baseline decoder: we drove the motor cortical network with the re-aiming solutions for the baseline decoder, and then averaged the motor cortical activity over all time and over all eight reach directions to obtain $\bar{\mathbf{r}}_{\mathrm{est}}$.

I.3.7 Simulation of the calibration task

The calibration task was simulated by setting the first two motor variables θ_1, θ_2 to the coordinates of the reach direction $\mathbf{y}_i^*$ being presented in each trial (a 2D unit vector pointing in one of eight equally spaced angles, cf. equation 14), and setting all other motor variables to zero ($\theta_3 = \theta_4 = \ldots = \theta_K = 0$). To simulate noise in the neural responses, we added noise in the dynamics and in the initial conditions on each trial. At each timestep, zero mean Gaussian noise with standard deviation 0.05 was sampled and added to the single neuron potentials $x_i(t)$ and to the two motor variables θ_1, θ_2. Initial conditions in each trial were sampled randomly from a 0-mean isotropic Gaussian with standard deviation 0.1. The network was driven for 1000ms in each trial, matching the duration of each trial in the experiment.

For simulations with WMP's and OMP's, we simulated 10 trials of each reach direction, replicating the structure of the calibration task used by [Sadtler et al., 2014]. For simulations with credit assignment rotation perturbations (fig. 7), the calibration task was identical except that only a single trial of each reach direction was simulated, to mimic the decoder initialization procedure of [Zhou et al., 2019]. Note that in all cases re-aiming with $\widetilde{K} = 2$ aiming variables implies re-aiming with the same two motor variables driving the calibration task neural responses, θ_1, θ_2.

I.3.8 Within- and outside- manifold perturbations

In the BMI system used by Sadtler et al. (2014) and Oby et al. (2019), 96-channel microelectrode arrays were used to record neural activity [Sadtler et al., 2014, Oby et al., 2019]. Spikes were detected by threshold crossings in the recorded voltage signals at each electrode, resulting in a series of spike trains at each electrode. Spike trains at each electrode could therefore contain spikes from multiple neurons near the electrode site, as no spike sorting was performed. In total, about 100 neurons were likely to have been recorded, constituting a small fraction of the total population of neurons in motor cortex. To simulate this, we ensured that the BMI decoder in our simulations only had access to a linear mixture of firing rates from $N_r = 99$ neurons. This was done by first multiplying the firing rates with a $N_r \times N$

"recording matrix" $\mathbf{H}$, which had the following tri-diagonal structure

$$H_{ij} = \begin{cases} 0 & \text{if } j > N_r \\ \xi_{ij} \sim \mathrm{Unif}(0,1) & \text{if } i - 3 \leq j \leq i + 3 \\ 0 & \text{else} \end{cases} \tag{23}$$

Thus, each "neural unit" in the vector $\mathbf{Hr}$ is composed of a linear mixture of three neurons, with neural units with adjacent indices mixing together overlapping sets of neurons. Following Sadtler et al., we next z-scored the activity recorded by each neural unit with respect to its statistics during the calibration task. We concatenated mean raw firing rates during the calibration task into an N-dimensional centering vector $\boldsymbol{\mu}$, and concatenated the standard deviation of the neural units (i.e. standard deviation of each entry in $\mathbf{Hr}$) into a $N_r \times N_r$ diagonal matrix $\mathbf{S}_r$. The readout was then obtained by decoding from the N_r-dimensional vectors of z-scored mixed firing rates,

$$\mathbf{y} = \underbrace{\mathbf{D}_0 \mathbf{S}_r^{-1} \mathbf{H}}_{\mathbf{D}}(\mathbf{r} - \boldsymbol{\mu}) \tag{24}$$

We refer to the $2 \times N_r$ matrix $\mathbf{D}_0$ as the *effective* decoding matrix. It is this matrix that is perturbed by the decoder perturbations of [Sadtler et al., 2014]. Note that the resulting full decoding matrix $\mathbf{D}$ is such that only its first N_r columns are non-zero, reflecting the fact that only a subset of the full population of motor cortical neurons is recorded by the BMI.

The effective decoding matrix $\mathbf{D}_0$ was constructed following the methods of Sadtler et al., with the exception that we used PCA instead of Factor Analysis to estimate the intrinsic manifold. This choice was made purely for the sake of numerical convenience, as PCA has a closed-form solution that can be computed much more efficiently. Moreover, the neurons in our network do not have Poisson-like variability, so the added benefits of Factor Analysis are not required here [Yu et al., 2009]. The full procedure for estimating the intrinsic manifold and constructing the baseline decoder is outlined in detail in section . In short, the baseline decoder effective decoding matrix can be expressed as a product of a $2 \times \ell$ matrix $\mathbf{K}$ and an $\ell \times N_r$ matrix $\mathbf{L}$,

$$\mathbf{D}_0^{\mathrm{base}} = \mathbf{KL}, \tag{25}$$

where ℓ equals the dimensionality of the intrinsic manifold. The matrix $\mathbf{L}$ projects N_r-dimensional activity patterns down to this ℓ-dimensional manifold; its rows span the subspace defined by the intrinsic manifold (cf. section I.4.2.1). In our simulations we used $\ell = 8$, as we found that the top 8 principal components contained 95% of the variance in calibration task responses. The matrix $\mathbf{K}$ then translates the resulting ℓ-dimensional dimensionality-reduced activity patterns into 2-dimensional BMI readouts.

This matrix is fit to the calibration task data; it is fit to decode the calibration task stimuli from the dimensionality-reduced neural responses at each timestep and trial (cf. section I.4.2.2).

Within-manifold perturbations (WMP's) perturb the baseline decoder in such a way that the row space of $\mathbf{L}$ remains intact, so as to conserve the decoder's relationship with the intrinsic manifold. This is done by simply shuffling the rows of $\mathbf{L}$ without modifying them, via pre-multiplication with a random $\ell \times \ell$ permutation matrix $\mathbf{P}$,

$$\mathbf{D}_0^{\mathrm{WMP}} = \mathbf{KPL}. \tag{26}$$

Outside-manifold decoders, on the other hand, directly disrupt the row space of $\mathbf{L}$. This is done by randomly shuffling the components of each of its rows, via post-multiplication with a random $N_r \times N_r$ permutation matrix $\mathbf{P}$,

$$\mathbf{D}_0^{\mathrm{OMP}} = \mathbf{KLP}. \tag{27}$$

It is important to keep in mind that both WMP's and OMP's can change the readouts in complex ways (see fig. 4a, supplementary figure 2 in [Golub et al., 2018] for examples), beyond a simple rotation like that depicted by the cartoon in figure 2b. Once the baseline decoder was constructed, we randomly sampled 100 within-manifold and 100 outside-manifold perturbations by randomly selecting 100 $\ell \times \ell$ and $N_r \times N_r$ permutation matrices, respectively.

To minimize any differences between these two types of decoder perturbations that would go beyond their opposing relationship to the intrinsic manifold, we imposed several restrictions on the selected permutation matrices, as was done by Sadtler et al. (see section I.4.2.3) for details). First, we enforced that the mean principal angle between the row space of the baseline effective decoding matrix and the row space of each perturbed effecitve decoding matrix fell within 60° and 80°. Second, we enforced that the re-aiming solutions for the baseline decoder would generate readouts through each perturbed decoder that resulted in a mean squared error between 0.6 and 0.8. Finally, we fit tuning curves to the neural activity generated by the re-aiming solutions for the baseline decoder, and asked how much the preferred directions would need to change to produce the same readouts under the perturbed decoder, using the same procedure utilized by Sadtler et al. We enforced that this change be between 30° and 45°. We typically found that about 100-200 permutations out all possible permutation matrices satisfied these criteria. We then randomly sampled 100 of them.

Following the procedure used by Sadtler et al. with monkey L, we did not actually consider all possible $N_r \times N_r$ permutation matrices for OMP's. Rather, we grouped all N_r neural units into ℓ groups, and then considered all ℓ-dimensional permutations of these groups. In other words, rather than permuting all N_r columns of $\mathbf{L}$, ℓ groups of columns were permuted. This ensured that the total number of possible decoder perturbations was the same for WMP's and OMP's. The ℓ groups were assigned as follows: for each neural unit in $\mathbf{Hr}$, we fit a cosine tuning curve to the calibration task responses to

obtain its modulation depth (cf. section I.4.2.3). The ℓ/N_r neurons with smallest modulation depths were then assigned to one group, while the remaining neurons were randomly assigned to the remaining $\ell - 1$ groups.

In figure 3a, we define the "dimensions" of the intrinsic manifold as a set of orthonormal basis vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_\ell$ spanning the intrinsic manifold (cf. section I.4.2.1). We then calculated the variance explained by each dimension by

$$\text{variance explained by dimension } i = \frac{\mathbf{f}_i^T \mathbf{\Sigma} \mathbf{f}_i}{\text{Tr}[\mathbf{\Sigma}]}. \tag{28}$$

For the gray curve, $\mathbf{\Sigma}$ was set to the sample covariance of the simulated calibration task responses. For the purple curve, $\mathbf{\Sigma}$ was set to the reachable repertoire covariance $\mathbf{\Sigma}_r$ (defined in equation 16). In each case, the cumulative variance explained was calculated by ordering the dimensions by variance explained and then summing them in that order.

I.3.9 Credit assignment rotation perturbations

In the BMI system used by Zhou et al. (2019), the activity of 10–12 individual neurons was recorded, sorted by matching spike waveforms to identify spikes from single neurons [Zhou et al., 2019]. Importantly, each neuron had reliable tuning to reach direction during the calibration task. In our simulation, we modelled this by randomly choosing 80 neurons, fitting tuning curves to their activity during the calibration task, and keeping only those with modulation depth greater than 0.5 (see below for how this is measured). This typically resulted in 10–15 neurons being included in the BMI decoder (i.e. being assigned non-zero weights in the decoding matrix $\mathbf{D}$). For the network used in the simulations reported in section I.1.5, this selection procedure resulted in $N_r = 11$ neurons being included.

Modulation depths were extracted from tuning curves fit to the calibration task data. Specifically, tuning curves were fit to time-averaged firing rates $\bar{\mathbf{r}}_j$ in each trial $j = 1, \ldots, 8$ of the calibration task. We modelled these average firing rates as depending linearly on the corresponding reach targets $\mathbf{y}_1^*, \ldots, \mathbf{y}_8^*$ in each trial through an $N \times 3$ matrix of tuning weights $\mathbf{T}$,

$$\bar{\mathbf{r}}_j = \mathbf{T}\tilde{\mathbf{y}}_j + \epsilon, \quad \epsilon \sim \mathcal{N}(0, 1) \tag{29}$$

where ϵ is 0-mean Gaussian noise and $\tilde{\mathbf{y}}_j$ is a 3D vector with the coordinates of the direction of the jth reach target, $\mathbf{y}_j^*$, as its first two components and a constant 1 in its third component, included to model baseline tonic firing rates of each neuron. Thus, only the first two columns of the tuning weight matrix $\mathbf{T}$ model how the ith neuron's firing rate depends on the reach target's direction, whereas its third column models activity independent of the reach target. We fit these tuning weights by maximum liklihood (i.e.

least squares regression),

$$\hat{\mathbf{T}} = \arg\max_{\mathbf{T}} \sum_{j=1}^{8} \log P\left(\bar{\mathbf{r}}_j | \mathbf{T}, \tilde{\mathbf{y}}_j\right) = \arg\min_{\mathbf{T}} \sum_{j=1}^{8} \|\bar{\mathbf{r}}_j - \mathbf{T}\tilde{\mathbf{y}}_j\|^2 = \left(\sum_{j=1}^{8} \bar{\mathbf{r}}_j \tilde{\mathbf{y}}_j^T\right)\left(\sum_{j=1}^{8} \tilde{\mathbf{y}}_j \tilde{\mathbf{y}}_j^T\right)^{-1}. \quad (30)$$

To extract from these weights the directional tuning of neuron i, we take the 2D vector comprising the first two components of the ith row of $\hat{\mathbf{T}}$. We notate this 2D vector by $m_i \mathbf{p}_i$, where $\mathbf{p}_i$ is a unit vector pointing in its direction and m_i is its norm. The angle of $\mathbf{p}_i$ is neuron i's preferred direction, and m_i is its modulation depth.

Following the methods of Zhou et al., the baseline decoder was constructed from these tuning curve parameters. First, raw firing rates were baseline-subtracted and normalized by their modulation depths,

$$\mathbf{r}^{\mathrm{norm}} = \mathbf{M}\left(\mathbf{r} - \boldsymbol{\mu}\right), \quad (31)$$

where $\boldsymbol{\mu}$ is given by the third column of $\hat{\mathbf{T}}$, containing the baseline firing rates estimated from the linear regression fit (equation 30), and $\mathbf{M}$ is an $N \times N$ diagonal matrix containing the inverse modulation depths m_i^{-1} for each of the N_r neurons recorded by the BMI and a 0 otherwise. These normalized firing rates were then transformed into 2D readouts by a $2 \times N$ effective decoding matrix $\mathbf{D}_0^{\mathrm{base}}$ containing the preferred direction vectors $\mathbf{p}_i$ of each of the N_r recorded neurons in their corresponding columns and 0's everywhere else. More precisely, the ith column of $\mathbf{D}_0^{\mathrm{base}}$, $\mathbf{d}_i$, is given by

$$\mathbf{d}_i = \begin{cases} \frac{k}{N_r} \mathbf{p}_i & \text{if neuron } i \text{ is recorded by BMI} \\ \mathbf{0} & \text{else,} \end{cases} \quad (32)$$

where the scaling constant k is chosen to minimize the mean squared error between the readouts from the calibration task activity and the target readouts. This is the classic population vector algorithm (PVA) [Georgopoulos et al., 1986]. The full decoding matrix of the baseline decoder was thus $\mathbf{D} = \mathbf{D}_0^{\mathrm{base}} \mathbf{M}$.

Credit assignment rotation perturbations were constructed by simply picking a random subset of the N_r non-zero columns of $\mathbf{D}_0^{\mathrm{base}}$ and rotating them. In our simulations, we picked a random 50% of these columns and rotated them 75° counter-clockwise, as was done in the decoder perturbations used by Zhou et al. We sampled 100 random perturbations in this way, in each case rotating a different subset of columns. The normalization matrix $\mathbf{M}$ and baseline subtraction parameters $\boldsymbol{\mu}$ are kept the same for all decoders.

To measure the tuning changes predicted by re-aiming, we simulated cursor reaches with each decoder by driving the motor cortical network with the re-aiming solutions for that decoder. In each case, noise was applied to the dynamics, exactly as in the calibration task. We then fit tuning curves to each neuron's time-averaged activity, using linear regression exactly as described in equation 30, and extracted the

preferred direction of each rotated, non-rotated, and indirect neuron. For each perturbed decoder, we then determined each neuron's change in preferred direction by calculating the difference between its preferred direction under the re-aiming solutions for the perturbed decoder and its preferred direction under the re-aiming solutions for the baseline decoder (cf. figures 7b, 7c). These changes were then averaged over all neurons in each sub-population (rotated, non-rotated, or indirect).

Figure 7e shows the percentiles (median and quartiles) of the distribution of this average tuning change over all 100 sampled decoder perturbations. An analagous analysis of the changes in modulation depth is shown in figure 12a.

I.3.10   Simulation of operant conditioning

In an operant conditioning task, there is no "target readout" per se. The objective is to simply increase the activity of one neuron over another, as much as possible. We can thus express the objective as maximizing the difference in firing rate between the two neurons, which can be thought of as a one-dimensional linear readout from the population. Formally, we calculate readouts in this task by a dot product between the firing rate vector $\mathbf{r}$ and a decoding vector $\mathbf{d}$ which has a $+1$ for the target neuron, a -1 for the distractor neuron, and 0's everywhere else. This one-dimensional readout indicates how much more active the target neuron is than the distractor neuron. The goal in an operant conditioning task is to maximize this readout.

Adding in a metabolic cost, the objective function we use for re-aiming is

$$\hat{\boldsymbol{\theta}} = \underset{\theta_1, \theta_2, \ldots, \theta_{\widetilde{K}}}{\arg\max} \ \mathbf{d} \cdot \mathbf{r}\big(t_{\mathrm{end}}; \boldsymbol{\theta}\big) - \frac{\gamma}{M} \sum_{i=1}^{M} u_i(\boldsymbol{\theta})^2. \tag{33}$$

By again applying the re-parameterization in equation 9, we can decompose this optimization problem into an optimization over the direction $\widetilde{\boldsymbol{\theta}}_0$ and norm s of the aiming variable vector $\widetilde{\boldsymbol{\theta}} = \begin{bmatrix} \theta_1 & \theta_2 & \ldots & \theta_{\widetilde{K}} \end{bmatrix}$,

$$\hat{s}, \hat{\widetilde{\boldsymbol{\theta}}}_0 = \underset{s>0, \|\widetilde{\boldsymbol{\theta}}_0\|=1}{\arg\max} \ \mathbf{d} \cdot \mathbf{r}\big(t_{\mathrm{end}}; s\widetilde{\boldsymbol{\theta}}_0\big) - \frac{\gamma}{M} \sum_{i=1}^{M} u_i(s\widetilde{\boldsymbol{\theta}}_0)^2$$

$$\approx \underset{s>0, \|\widetilde{\boldsymbol{\theta}}_0\|=1}{\arg\max} \ s\mathbf{d} \cdot \mathbf{r}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0\big) - \frac{\gamma}{2}s^2, \tag{34}$$

where in the second line we plugged in equation 10 and used the mean-field approximation from equation 11. This approximation allows us to analytically solve for the optimal magnitude $\hat{s}$,

$$\hat{s}(\widetilde{\boldsymbol{\theta}}_0) = \underset{s>0}{\arg\max} \ s\mathbf{d} \cdot \mathbf{r}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0\big) - \frac{\gamma}{2}s^2 = \frac{1}{\gamma}\phi\left(\mathbf{d} \cdot \mathbf{r}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0\big)\right), \tag{35}$$

which in turn allows us to solve for the optimal direction $\hat{\widetilde{\boldsymbol{\theta}}}_0$ via optimization over the $\widetilde{K}$-dimensional

unit hypersphere,

$$\hat{\widetilde{\boldsymbol{\theta}}}_0 = \underset{\|\widetilde{\boldsymbol{\theta}}_0\|=1}{\arg\max} \; \hat{s}(\widetilde{\boldsymbol{\theta}}_0)\mathbf{d} \cdot \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0) = \underset{\|\widetilde{\boldsymbol{\theta}}_0\|=1}{\arg\max} \; \phi\left(\mathbf{d} \cdot \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)\right). \tag{36}$$

In all our operant conditioning simulations, we used $\widetilde{K} = 2$. We thus solved the above optimization by brute force search over the unit circle. As in previous simulations, we used $t_{\mathrm{end}} = 1000$ms.

Note that the optimal readout achieved by this re-aiming solution is

$$\mathbf{d} \cdot \mathbf{r}(t_{\mathrm{end}}; \hat{\boldsymbol{\theta}}) = \frac{1}{\gamma}\phi\left(\mathbf{d} \cdot \mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}_0)\right)^2. \tag{37}$$

Changing the metabolic cost weight γ thus has no other effect on re-aiming other than re-scaling the re-aiming solutions and the readouts they produce. We thus simply set it to 1 in all our simulations.

In classic operant conditioning experiments [Fetz and Baker, 1973], neurons selected for operant conditioning had to be active prior to the conditioning task to be identified by the recording electrode. We imposed a similar constraint in our simulation by first driving the network with 500 random K-dimensional motor commands, and identifying the 50% of neurons with highest average firing rate over motor commands and time. The neuron pairs used for operant conditioning were sampled from this sub-population.

To simulate a baseline period of spontaneous behavior prior to operant conditioning, we used a similar procedure but now driving the network with 500 random $\widetilde{K}$-dimensional motor commands, with $\widetilde{K} = 2$. This allowed us to ask whether operant conditioning performance under re-aiming could be predicted from correlations arising during behavior driven by the same motor variables used for re-aiming. In figure 8d, correlations were measured by correlation coefficient between the two conditioned neurons. In figure 8e, correlation coefficients between each indirect neuron and the target neuron are plotted, against the firing rate of the indirect neuron at $t_{\mathrm{end}} = 1000$ms when driving the network with the re-aiming solution.

Supplementary Methods and Results

I.4.1 Simulations with different motor cortical connectivity

In addition to the randomly connected network architecture used in the results presented above, we simulated re-aiming in the task of Sadtler et al. (2014) with alternative motor cortical connectivity profiles, described below. In each case, we simulated the calibration task, sampled decoder perturbations, and computed re-aiming solutions exactly as described above (subject to minor exceptions noted below). Results of these simulations are shown in figure 9.
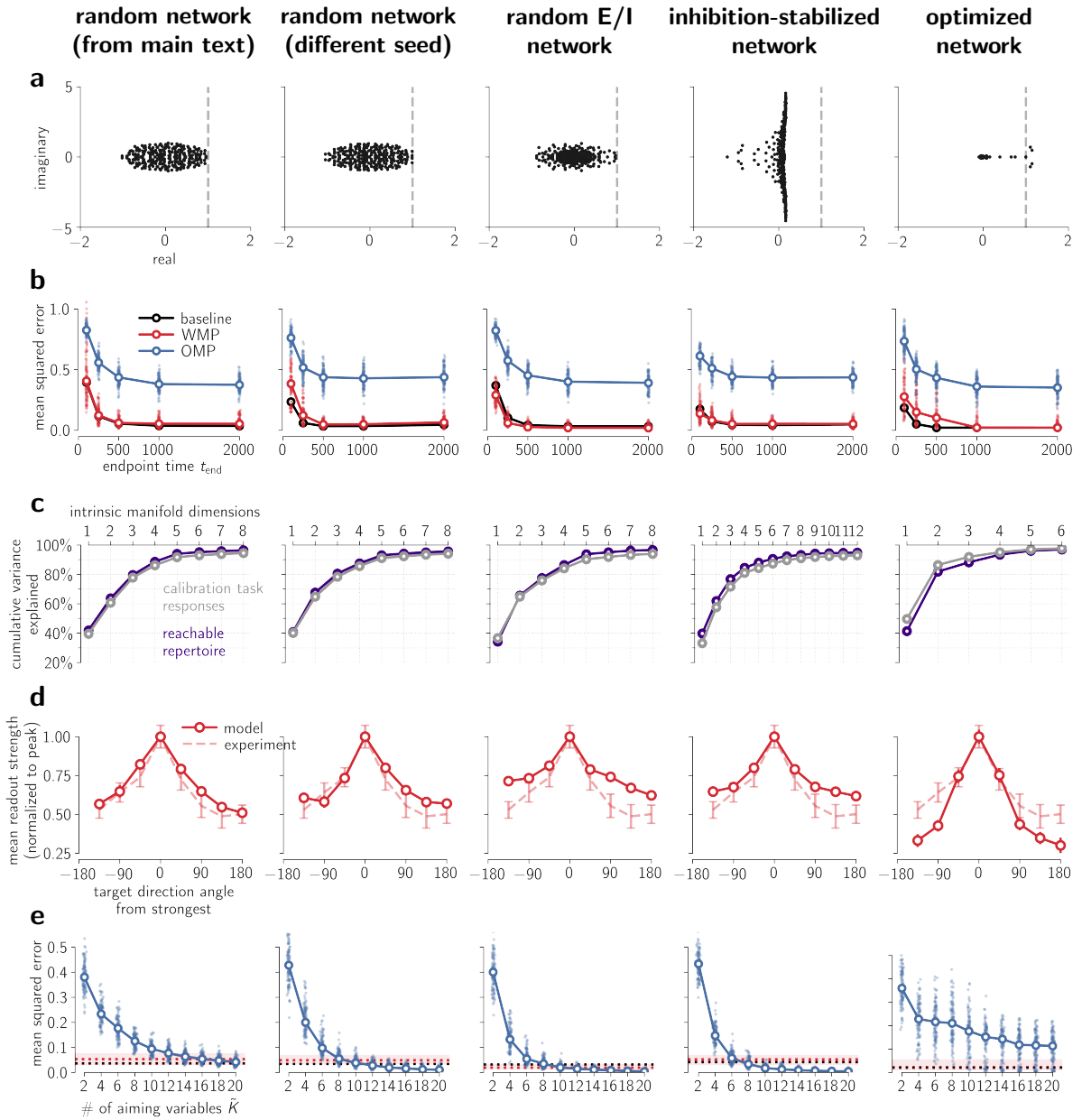
Figure 9: Each column shows simulation results for a different network: (1) the randomly connected network used in the main text; (2) another randomly connected network, with weights sampled in exactly the same way; (3) a network with random E/I connectivity; (4) a network with inhibition-optimized E/I connectivity; (5) network with connectivity optimized for delayed center-out reaching.

(9a) Eigenspectra of recurrent weight matrices of each network, plotted on the complex plane. Note that the optimized network has low-rank connectivity: almost all eigenvalues are clustered at 0.

(9b) Mean squared error achieved by re-aiming solutions for different endpoint time, t_{end}, for each decoder. Lighter markers correspond to individual decoders, darker open markers (connected by lines) show medians over all decoders. Note that the metabolic cost weight, γ, is chosen to ensure low error under the baseline decoder for $t_{\text{end}} = 1000$ms only (cf. Methodssection "Setting the metabolic cost weight")

(9c) Cumulative variance of calibration task responses and reachable repertoire in each dimension of the intrinsic manifold. Intrinsic manifold was found to be about 12-dimensional with stability-optimized connectivity and about 6-dimensional with connectivity optimized for delayed center-out reaching.

(9d) Average maximal readout strength, as in figure 4c. Experimental data overlaid for reference.

(9e) Mean squared error achievable with different numbers of aiming variables, for each OMP. Same conventions followed as in figure 6a.

Random excitatory/inhibitory (E/I) connectivity We constructed a random sparse and balanced

E/I recurrent connectivity matrix following the sampling procedure described in [Hennequin et al., 2012].

In short, all excitatory weights had the same strength, all inhibitory had the same strength (re-scaled relative to the excitatory weights to account for the different number of excitatory and inhibitory neurons), and each row of the weight matrix was enforced to be 0 mean. We used a sparsity of 10% (i.e. only 10% of weights were non-zero), with 80% of the neurons in the population being excitatory. Input and encoding weights were sampled randomly as for the randomly connected network (cf. section I.3.1).

Inhibition-optimized excitatory/inhibitory connectivity   We constructed a spare and balanced E/I recurrent connectivity matrix following the optimization procedure described in [Hennequin et al., 2014]. In short, the excitatory weights were initialized to be very strong, and then inhibitory weights were optimized to ensure the dynamics were stable (by minimizing the spectral abscissa of the full connectivity matrix). Half of the neurons were assigned to be excitatory, and the inhibitory weights were enforced to be on average three times stronger than the excitatory weights. The only difference with [Hennequin et al., 2014] was that we initialized the weight matrices with a spectral radius of 5. This was necessary as we found that using an initial spectral radius of 10 (as was used in the original study) lead to chaotic dynamics under the rectified linear activation function and sustained input. Input and encoding weights were sampled randomly as for the randomly connected network (cf. section I.3.1).

Because of their highly non-normal dynamics, we found that these networks remained highly sensitive to changes in initial conditions, even with the reduced initial spectral radius. We therefore reduced the standard deviation of the initial conditions by half when simulating the calibration task (see Methods). We also found that these networks produced much higher-dimensional calibration task responses than the randomly connected network. We therefore used a 12-dimensional intrinsic manifold (i.e. $\ell = 12$) for constructing WMP's and OMP's.

Connectivity optimized for delayed center-out reaching   Finally, we constructed a network optimized to output joint torques for delayed center-out reaching with a biomechanical arm model. We used an architecture and optimization scheme inspired by [Sussillo et al., 2015], in which the recurrent network is driven by two distinct inputs. The first input is a one-dimensional signal reflecting a go cue that indicates when the reach should be performed (go time). This is constructed by setting θ_K to 1 at the start of the trial and then setting it to 0 at go time, 1000ms after trial start. The other input is a two-dimensional signal reflecting the visual presentation of the target to reach towards, presented prior to go time to prepare the subject (or network) to perform the delayed center-out reach. This is constructed by setting θ_1, θ_2 to the coordinates of the reach target at a randomly sampled target presentation time before the go cue, and then setting it back to 0 at the same time the go cue input is shut off. All other motor variables are set to 0 ($\theta_3 = \theta_4 = \ldots = \theta_{K-1} = 0$). We chose to use the very last motor variable θ_K for the go cue to reflect the hypothesis that subjects would not re-aim with this non-directional motor variable.

Two joint torques were read out from the network through a set of readout weights, which were optimized along with the network parameters (input, recurrent, and encoding weights). These parameters were optimized to produce the joint torques required to move the endpoint of a planar two-link arm model (taken from [Todorov and Li, 2003]) to the cued reach target, in 500ms with a bell-shaped speed profile. Following the methods of [Kao et al., 2020], these target joint torques were computed by backpropagating through the arm model dynamics to minimize mean squared error between the arm endpoint velocity and the desired velocity profile for each reach target. We then trained the network's parameters so that in each trial it would produce 0 torque until go time, followed by the optimal reaching torque corresponding to the reach target on that trial.

The loss function minimized was a combination of the mean squared error plus L2 regularization on all parameters as well as on network firing rates (to encourage naturalistic solutions to this task [Sussillo et al., 2015]). This was minimized via stochastic gradient descent using the Adam optimization algorithm (initial learning rate set to .001) [Kingma and Ba, 2017]. Since only three motor variables were non-zero during this task, only the first two columns of the encoding weights U_{ij} were optimized by this procedure. The remaining columns were thus set randomly, as for the randomly connected network (cf. section I.3.1).

We found that the resulting optimized network produced much lower-dimensional calibration task responses than randomly connected networks did. We therefore used a 6-dimensional intrinsic manifold (i.e. $\ell = 6$) for constructing WMP's and OMP's in simulations with this network. This meant only $6! - 1 = 719$ possible decoder perturbations existed (as opposed to 40,319), so far fewer decoder perturbations satisfied the stringent criteria outlined in section I.4.2.3 for WMP's and OMP's. We thus loosened these criteria to include WMP's and OMP's with mean squared error going up to 1.2. These networks were also found to be highly sensitive to noise, so we reduced the standard deviation of the noise in the dynamics and in the initial conditions during simulation of the calibration task to 0.02 and 0.005, respectively.

I.4.2 Construction of Sadtler et al. (2014) BMI decoders

I.4.2.1 Estimating the intrinsic manifold

To estimate the intrinsic manifold, we fit a Probabilistic PCA (PPCA) model [Tipping and Bishop, 1999] to the mixed and z-scored calibration task responses,

$$\mathbf{r}_i^{\text{mixed}} = \mathbf{S}_r^{-1}\mathbf{H}(\mathbf{r}_i - \boldsymbol{\mu}). \tag{38}$$

Here, i indexes a particular timestep and trial during the calibration task. The PPCA generative model assumes that each of these data points are generated from a corresponding set of ℓ uncorrelated latent

variables $\mathbf{z}_i = \begin{bmatrix} z_{i1} & z_{i2} & \dots & z_{i\ell} \end{bmatrix}$ as follows,

$$\mathbf{z}_i \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right), \tag{39a}$$

$$\mathbf{r}_i^{\text{mixed}}|\mathbf{z}_i \sim \mathcal{N}\left(\mathbf{F}\mathbf{z}_i, \sigma^2\mathbf{I}\right). \tag{39b}$$

The model thus assumes that the activity patterns $\mathbf{r}_i^{\text{mixed}}$ are concentrated within the column space of the factor loading matrix $\mathbf{F}$ – it is the columns of this matrix that define the intrinsic manifold.

However, this defines the dimensions of the intrinsic manifold in mixed and z-scored neural activity space (i.e. the space defined by the coordinates of the $\mathbf{r}_i^{\text{mixed}}$ vectors). To convert these to dimensions of the full N-dimensional state space, where each coordinate corresponds to the activity of an individual neuron (i.e. the space defined by the coordinates of the $\mathbf{r}_i$ vectors), we invert equation 38 to obtain

$$\mathbf{r}_i = \mathbf{H}^{-1}\mathbf{S}_r\mathbf{r}_i^{\text{mixed}} + \boldsymbol{\mu},$$

where we define $\mathbf{H}^{-1}$ as the $N \times N_r$ matrix containing the inverse of the tri-diagonal component of $\mathbf{H}$ in its first N_r rows and 0's filling all subsequent rows. We then apply this linear transformation to the columns of $\mathbf{F}$ to obtain an analogous $N \times \ell$ factor loading matrix $\mathbf{F}_r$ defined in the full N-dimensional state space,

$$\mathbf{F}_r = \mathbf{H}^{-1}\mathbf{S}_r\mathbf{F}.$$

Note that, since the bottom $N - N_r$ rows of $\mathbf{H}^{-1}$ are filled with 0's, those same rows of $\mathbf{F}_r$ are also filled with 0's. This reflects the fact that the intrinsic manifold is orthogonal to the dimensions of activity corresponding to neurons not recorded in the experiment. Finally, we defined an orthonormal basis $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_\ell \in \mathbb{R}^N$ for the intrinsic manifold by taking the left singular vectors of $\mathbf{F}_r$. These are the vectors used in equation 28 for figure 3a.

The parameters $\mathbf{F}$ and σ^2 are set to their maximum likelihood estimates under the calibration task data, given by

$$\mathbf{F} = \begin{bmatrix} \sqrt{\lambda_1 - \sigma^2}\mathbf{v}_1 & \sqrt{\lambda_2 - \sigma^2}\mathbf{v}_2 & \dots & \sqrt{\lambda_\ell - \sigma^2}\mathbf{v}_\ell \end{bmatrix}$$

$$\sigma^2 = \frac{1}{N_r - \ell} \sum_{i=\ell+1}^{N_r} \lambda_i$$

where $\lambda_1, \lambda_2, \dots, \lambda_{N_r}$ are the eigenvalues of the sample covariance of the mixed and z-scored calibration task activity $\{\mathbf{r}_i^{\text{mixed}}\}$ ordered from largest to smallest (i.e. λ_1 is the largest eigenvalue), and $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N_r}$ are their associated eigenvectors (i.e. the principal components, ordered from most to least variance explained).

This method for estimating the intrinsic manifold is almost the same as that used by Sadtler et al., which differs only in that a Factor Analysis model was used rather than a PPCA model. In that case, the maximum likelihood estimates of the model parameters cannot be evaluated in closed form and must be computed via an iterative optimization algorithm (the Expectation Maximization algorithm). We found that using a Factor Analysis model instead of PPCA had no noticeable effects on our results, so we reported only results with the more easily fit PPCA model.

I.4.2.2   Construction of the baseline decoder

As mentioned in the Methods section, the baseline decoder has the following form

$$\mathbf{D}_0^{\text{base}} = \mathbf{KL}$$

$$\mathbf{K} \in \mathbb{R}^{2 \times \ell}, \quad \mathbf{L} \in \mathbb{R}^{\ell \times N_r}.$$

We term $\mathbf{L}$ the *dimensionality reduction matrix* and $\mathbf{K}$ the *velocity readout matrix*. Here we describe in greater detail how these two matrices are fit to the calibration task data. Unless otherwise noted, these procedures are exactly as those described in [Sadtler et al., 2014] and [Golub et al., 2018].

The dimensionality reduction matrix $\mathbf{L}$ is chosen by asking what setting of the latent variables $\hat{\mathbf{z}}_i$ under the PPCA model (equation 39) would have been most likely to have generated a given mixed and z-scored activity pattern $\mathbf{r}_i^{\text{mixed}}$; that is, the mode of the posterior distribution $P(\mathbf{z}_i | \mathbf{r}_i^{\text{mixed}})$. It is straight forward to derive that, under the PPCA generative model, this is given by

$$\hat{\mathbf{z}}_i = \underbrace{\left(\mathbf{F}^T \mathbf{F} + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{F}^T}_{\hat{\mathbf{L}}} \mathbf{r}_i^{\text{mixed}}.$$

The $N_r \times \ell$ matrix $\hat{\mathbf{L}}$ thus yields a linear transformation from N_r dimensions to ℓ dimensions. The z-scored and mixed activity patterns $\{\mathbf{r}_i^{\text{mixed}}\}$ from the calibration task can thus be reduced to ℓ dimensions via multiplication with $\hat{\mathbf{L}}$, resulting in a corresponding set of dimensionality-reduced activity patterns $\{\hat{\mathbf{z}}_i\}$ (as above, here and in the rest of this section the index i jointly indexes a timestep and trial of the calibration task).

To complete the construction of the dimensionality reduction matrix $\mathbf{L}$, these dimensionality-reduced activity patterns are then z-scored. The standard deviations of each component of the $\hat{\mathbf{z}}_i$ vectors are calculated over all timesteps and trials of the calibration task, and collected in a diagonal matrix $\mathbf{S}_z$. Note that mean subtraction is not necessary since the activity vectors $\mathbf{r}_i^{\text{mixed}}$ have already been z-scored so are mean 0. The final dimensionality reduction matrix is then given by

$$\mathbf{L} = \mathbf{S}_z^{-1} \hat{\mathbf{L}}$$

This second z-scoring step is necessary to ensure that controlling the BMI does not require neurons to produce firing rates beyond the range exhibited during the calibration task.

The dimensionality reduction matrix used by Sadtler et al. differed from ours in that $\hat{\mathbf{L}}$ was constructed from the posterior distribution under a Factor Analysis generative model, rather than a PPCA generative model. Like in PPCA, the mode of the posterior distribution of a Factor Analysis model can also be expressed as a linear transformation of $\mathbf{r}_i^{\mathrm{mixed}}$, yielding a very similar expression for $\hat{\mathbf{L}}$.

The velocity readout matrix $\mathbf{K}$ is also chosen by maximum likelihood fit of a generative model. In this case, we assume that the z-scored dimensionality-reduced activity patterns from the calibration task, $\hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}} = \mathbf{L}\mathbf{r}_i^{\mathrm{mixed}}$, depend on the observed cursor velocities $\mathbf{y}_i$ via the following latent Gaussian state space model,

$$\mathbf{y}_i | \mathbf{y}_{i-1} \sim \mathcal{N}\left(\mathbf{y}_{i-1}, \mathbf{Q}\right)$$

$$\hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}} | \mathbf{y}_i \sim \mathcal{N}\left(\mathbf{B}\mathbf{y}_i, \mathbf{R}\right)$$

where $i-1$ indexes the previous timestep in the same trial. Note that the cursor velocities $\mathbf{y}_i$ are constant within each trial of the calibration task. As was done in the original experiment of Sadtler et al., we set

$$\mathbf{Q} = 2k^2\mathbf{I},$$

where $k = 1/.15$ denotes the ratio of the cursor speeds used in our simulation ($\|\mathbf{y}_i\| = 1$) and the cursor speeds used in the original experiment ($\|\mathbf{y}_i\| = .15$ m/s). Maximum likelihood estimates of the remaining parameters are given by

$$\mathbf{B} = \left(\sum_i \hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}} \mathbf{y}_i^T\right)\left(\sum_i \mathbf{y}_i \mathbf{y}_i^T\right)^{-1}$$

$$\mathbf{R} = \frac{1}{T}\sum_i \hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}} \hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}\,T} - \hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}}(\mathbf{B}\mathbf{y}_i)^T$$

where T denotes the total number of data points in the calibration task data: the number of timesteps in each trial times the total number of trials.

The velocity readout matrix is then obtained by asking what cursor velocity $\hat{\mathbf{y}}_i$ would have most likely generated a given low-dimensional activity pattern $\hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}}$, given observations of all previous activity patterns; that is, the mode of the posterior distribution $P(\mathbf{y}_i | \hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}}, \hat{\mathbf{z}}_{i-1}^{\mathrm{z\text{-}scored}}, \hat{\mathbf{z}}_{i-2}^{\mathrm{z\text{-}scored}}, \ldots)$, where the ellipses go back to the first timestep of the given trial. We use the posterior distribution at steady state, whose mode is given by

$$\hat{\mathbf{y}}_i = \left(\mathbf{I} - \mathbf{K}\mathbf{B}\right)\hat{\mathbf{y}}_{i-1} + \mathbf{K}\hat{\mathbf{z}}_i^{\mathrm{z\text{-}scored}},$$
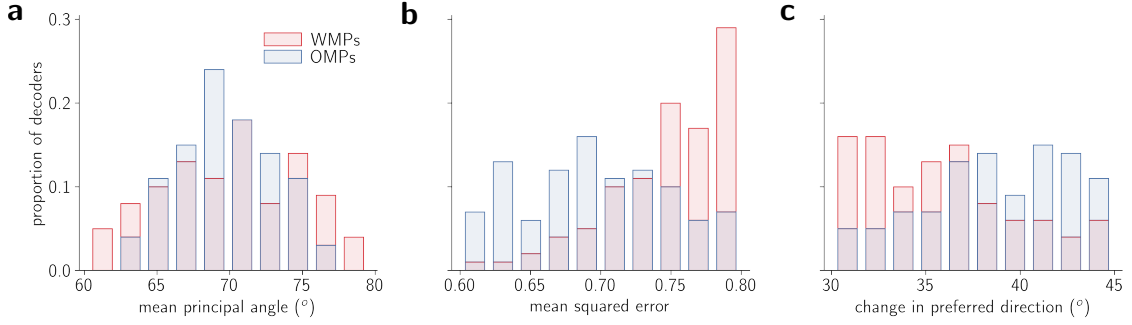
Figure 10: Differences between sampled decoder perturbations and the baseline decoder.
(10a) Distribution of mean principal angle between row space of baseline decoder and each decoder perturbation.
(10b) Distribution of mean squared error achieved by mean calibration task responses, under each decoder perturbation.
(10c) Distribution of minimal absolute change in preferred direction needed to produce the same readouts with each decoder perturbation as with the baseline decoder.

where $\mathbf{K}$ is the so-called steady-state Kalman gain matrix. This matrix is given by

$$\mathbf{K} = \mathbf{\Sigma}_{\mathrm{ss}}\mathbf{B}^T(\mathbf{B}\mathbf{\Sigma}_{\mathrm{ss}}\mathbf{B}^T + \mathbf{R})^{-1}$$

where $\mathbf{\Sigma}_{\mathrm{ss}}$ is the steady-state posterior covariance, given by the solution to the discrete-time algebraic Riccatti equation

$$0 = \mathbf{\Sigma}_{\mathrm{ss}}\mathbf{B}^T(\mathbf{B}\mathbf{\Sigma}_{\mathrm{ss}}\mathbf{B}^T + \mathbf{R})^{-1}\mathbf{B}\mathbf{\Sigma}_{\mathrm{ss}} - \mathbf{Q}$$

The $2 \times \ell$ velocity readout matrix used for the baseline decoder is thus set to the steady-state Kalman gain matrix, $\mathbf{K}$.

I.4.2.3   Subsampling WMP's and OMP's

As mentioned in the methods, we attempted to minimize any differences between within- and outside-manifold perturbations that would go beyond their opposing relationship to the intrinsic manifold. To do this, we first calculated every possible WMP and OMP, corresponding to each ℓ-dimensional permutation. Since we set $\ell = 8$, this resulted in $\ell! - 1 = 40,319$ decoder perturbations of each type (minus 1 to exclude the identity permutation). We then quantified how different each of these perturbations were from the baseline decoder with three different metrics, and eliminated all decoder perturbations for which one or more of these metrics fell outside a specific range.

The first metric is the angle between the perturbed decoder's row space and the baseline decoder's. For each decoder perturbation, $\mathbf{D}_0^{\mathrm{WMP}}$ or $\mathbf{D}_0^{\mathrm{OMP}}$, we calculated the two principal angles [Björck and Golub, 1973] between its row space and that of the baseline decoder effective decoding matrix, $\mathbf{D}_0^{\mathrm{base}}$, and averaged these two angles. Any decoder perturbations for which this mean principal angle was greater than 80° or less than 60° was eliminated (fig. 10a).

The second metric is the mean squared error that would be achieved if the subject were to simply reproduce the neural activity from the calibration task. Analagous to the procedure followed by Sadtler

et al., we averaged the calibration task responses over time and over trials for each reach target,

$$\bar{\mathbf{r}}_j^{\text{calib}} = \langle \mathbf{r}_i \rangle_{i \in \text{time points in calibration task trials with reach target } j}$$

and then computed the readouts from these time- and trial- averaged firing rate vectors under each decoder perturbation, $\mathbf{D}_0^{\text{WMP}}$ or $\mathbf{D}_0^{\text{OMP}}$. We then discarded all decoder perturbations where the mean squared error between these readouts and the target readouts was greater than 0.8 or less than 0.6 (fig. 10b).

The third metric is to ask how much the mean calibration task responses would have to change to produce the same readouts under the perturbed decoder as under the baseline decoder. We first calculated the time- and trial- averaged z-scored and mixed firing rates from the calibration task

$$\bar{\mathbf{r}}_j^{\text{mixed}} = \mathbf{S}_r^{-1} \mathbf{H}(\bar{\mathbf{r}}_j^{\text{calib}} - \boldsymbol{\mu}).$$

For each perturbed decoder, $\widetilde{\mathbf{D}}_0 = \mathbf{D}_0^{\text{WMP}}$ or $\mathbf{D}_0^{\text{OMP}}$, we then computed the activity patterns closest to $\bar{\mathbf{r}}_j^{\text{mixed}}$ that would produce the same readouts through that decoder as the original activity patterns would through the baseline decoder,

$$\hat{\mathbf{r}}_j^{\text{mixed}} \left(\widetilde{\mathbf{D}}_0 \right) = \arg\min_{\mathbf{r}} \|\mathbf{r} - \bar{\mathbf{r}}_j^{\text{mixed}}\|^2 \quad \text{subject to} \quad \widetilde{\mathbf{D}}_0 \mathbf{r} = \mathbf{D}_0^{\text{base}} \bar{\mathbf{r}}_j^{\text{mixed}}$$
$$= \bar{\mathbf{r}}_j^{\text{mixed}} + \widetilde{\mathbf{D}}_0^T \left(\widetilde{\mathbf{D}}_0 \widetilde{\mathbf{D}}_0^T \right)^{-1} \left(\mathbf{D}_0^{\text{base}} - \widetilde{\mathbf{D}}_0 \right) \bar{\mathbf{r}}_j^{\text{mixed}}.$$

We then quantified the difference between $\bar{\mathbf{r}}_j^{\text{mixed}}$ and $\hat{\mathbf{r}}_j^{\text{mixed}}(\widetilde{\mathbf{D}}_0)$ by fitting tuning curves and asking how much the preferred direction changed. Tuning curves were fit by least-squares regression, exactly as described in Methodsequation 30 (but with $\bar{\mathbf{r}}_j^{\text{mixed}}$ or $\hat{\mathbf{r}}_j^{\text{mixed}}(\widetilde{\mathbf{D}}_0)$ plugged in for $\bar{\mathbf{r}}_j$), and preferred directions were extracted from the fitted tuning weights as described in that section. For each decoder perturbation, we then computed the mean absolute difference between the preferred directions fitted to the computed activity patterns $\{\hat{\mathbf{r}}_j^{\text{mixed}}(\widetilde{\mathbf{D}}_0)\}_{j=1}^8$ and the preferred directions fitted to the true calibration task mean responses $\{\bar{\mathbf{r}}_j^{\text{mixed}}\}_{j=1}^8$. Any perturbed decoders that resulted in a mean absolute difference of more than 45° or less than 30° were discarded (fig. 10c).

We typically found that about 100-200 permutations out all possible decoder perturbations satisfied these criteria. We then randomly sampled 100 of them. The distributions of these three merics for the 100 sampled WMP's and OMP's used in sections I.1.2, I.1.3, I.1.4 are shown in figures 10a, 10b, and 10c.
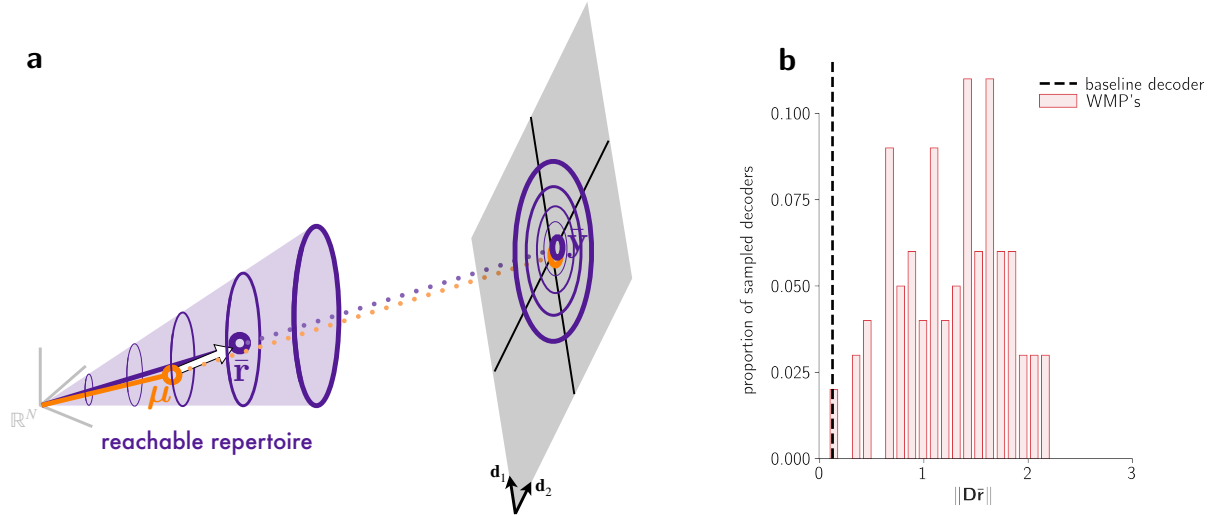
Figure 11: Geometry of the readouts reachable under the baseline decoder.
(11a) Cartoon of the reachable repertoire, the decoding plane of the baseline decoder, and the readouts reachable under this decoder. All formatting conventions exactly as in figure 5a, see caption of that figure for further details.
(11b) Histogram of the norm of the projection of reachable repertoire mean through each of the 100 sampled WMP's. The dashed black line shows the same for the baseline decoder.

I.4.3 Geometry of the readouts reachable under the baseline decoder

In our analysis of biases in the readouts reachable under WMP's, we completely ignored the fact that the readouts reachable under the baseline decoder do not seem to be biased at all (fig. 3di). Rather, the readouts reachable under this decoder are both centered at the origin and symmetrically distributed around it. Why is this the case for the baseline decoder, but not for WMP's?

Recall that the centroid of the reachable readouts depends on the difference between the reachable repertoire mean and the centering vector (equation 6), and that in our model this difference is non-zero, $\bar{\mathbf{r}} - \boldsymbol{\mu} \neq \mathbf{0}$ (fig. 5b). For the centroid to be at the origin (i.e. $\bar{\mathbf{y}} \approx \mathbf{0}$), then, it must be that this difference is orthogonal to the decoding plane of the baseline decoder, so that its projection onto the decoding matrix is near 0, $\bar{\mathbf{y}} = \mathbf{D}(\bar{\mathbf{r}} - \boldsymbol{\mu}) \approx \mathbf{0}$ (equation 6). This is illustrated in the cartoon in figure 11a, which depicts the relationship between the reachable repertoire (a purple cone, as in figure 5a) and the decoding plane of the baseline decoder (the subspace spanned by the two rows $\mathbf{d}_1, \mathbf{d}_2$ of the decoding matrix $\mathbf{D}$). The purple and orange circles on the left depict the reachable repertoire mean, $\bar{\mathbf{r}}$, and the centering vector, $\boldsymbol{\mu}$, respectively. Note that their difference – depicted by the open arrow – is largely orthogonal to the decoding plane. As a result, the centroid of the reachable readouts, $\bar{\mathbf{y}}$, lies almost exactly at the origin, and the rings of reachable activity patterns generated by different motor command magnitudes form concentric rings of readouts around it.

To understand why this specific geometric arrangement would hold for the baseline decoder, we must turn to the details of how it is constructed. The baseline decoder is built to decode from the calibration task neural responses the identities of the stimuli driving them (cf. section I.4.2.2). It thus ignores any

dimensions of calibration task activity that do not provide information about the stimulus. One of these dimensions happens to be the reachable repertoire mean, $\bar{\mathbf{r}}$. This can be appreciated from looking at the projection of the calibration task responses visualized in figure 5b, which shows that the trajectories of activity during different trials of the calibration task all evolve identically along this direction, despite being evoked by different stimuli. This means that decoding from this dimension is useless for decoding stimulus identity from the neural responses, so the baseline decoder ignores it. This is confirmed in figure 11b, where we plot the norm of the projection of the reachable repertoire mean, $\mathbf{D}\bar{\mathbf{r}}$, through the baseline decoder and all sampled WMP's. The reachable repertoire mean is evidently almost orthogonal to the baseline decoder, producing a projection with near 0 norm (i.e. $\|\mathbf{D}\bar{\mathbf{r}}\| \approx 0$). Since the centroid of the reachable readouts is proportional to this projection (equation 8), this implies that the reachable readouts will be centered at the origin of the workspace.

The reachable repertoire mean does, on the other hand, have a strong projection through most WMP's (i.e. $\|\mathbf{D}\bar{\mathbf{r}}\| > 0$), explaining why there are strong biases in readouts reachable under these decoders (fig. 4c). That this is the case can also be traced back to the geometry of the calibration task responses. While the direction of the reachable repertoire mean may not contain information about the stimuli driving these responses, it still contains a lot of their variance. Consequently, this dimension forms part of the intrinsic manifold. Since WMP's are essentially randomly oriented within the intrinsic manifold, they are likely to decode from it by chance.

I.4.4 Derivation of reachable repertoire mean and covariance

To calculate expectations over a uniform distribution on the reachable repertoire manifold, we require an expression for the probability density function of this distribution. This is given by the volume element of the manifold divided by its total volume.

The reachable repertoire can be formalized as the following manifold,

$$\mathcal{R} = \{\mathbf{r}(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}) : \|\widetilde{\boldsymbol{\theta}}\| \leq s_{\max}\}$$

That is, the set of all activity patterns that can be generated by a $\widetilde{K}$-dimensional motor command $\widetilde{\boldsymbol{\theta}} = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_{\widetilde{K}} \end{bmatrix}$ with norm less than or equal to $s_{\max}$. The function $\mathbf{r}(t_{\text{end}}; \widetilde{\boldsymbol{\theta}})$ can be thought of as a mapping from $\widetilde{K}$-dimensional aiming variable vectors, $\widetilde{\boldsymbol{\theta}} \in \mathbb{R}^{\widetilde{K}}$, to N-dimensional firing rate vectors, $\mathbf{r} \in \mathbb{R}^N$, on the reachable repertoire. Mathematically, the mapping $\widetilde{\boldsymbol{\theta}} \to \mathbf{r}(t_{\text{end}}; \widetilde{\boldsymbol{\theta}})$ essentially defines the manifold, constituting the so-called inverse coordinate map from Euclidean space onto the manifold [Lee, 2012]. The volume element of this manifold, $dV(\widetilde{\boldsymbol{\theta}})$, can be derived from the Jacobian, $\mathbf{J}$, of this inverse coordinate map,

$$dV(\widetilde{\boldsymbol{\theta}}) = \sqrt{\det[\mathbf{J}(\widetilde{\boldsymbol{\theta}})^T \mathbf{J}(\widetilde{\boldsymbol{\theta}})]},$$

where $\det[\cdot]$ denotes the matrix determinant and the $N \times \widetilde{K}$ Jacobian matrix is given by

$$J_{ij}(\widetilde{\boldsymbol{\theta}}) = \frac{\partial}{\partial \theta_j} r_i(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}).$$

The total volume of the manifold is given by the integral of the volume element over all allowable aiming variables, i.e. all aiming variable vectors $\widetilde{\boldsymbol{\theta}} \in \mathbb{R}^{\widetilde{K}}$ satisfying $\|\widetilde{\boldsymbol{\theta}}\| < s_{\text{max}}$. The probability density function of the uniform distribution on $\mathcal{R}$ is given by the volume element divided by the total volume.

For the case of $\widetilde{K} = 2$, these quantities can be easily expressed in terms of the polar coordinates of the aiming variables,

$$\widetilde{\boldsymbol{\theta}} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = s \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}.$$

To facilitate the treatment of this change of variables, we introduce the notation

$$\mathbf{r}\left(t_{\text{end}}; s, \varphi\right) := \mathbf{r}\left(t_{\text{end}}; s \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}\right),$$

i.e. $\mathbf{r}(t_{\text{end}}; s, \varphi)$ denotes the activity pattern at time t_{end} generated by a pair of aiming variables θ_1, θ_2 with angle φ and norm s. By the scale invariance of the RNN dynamics (equation 10), we have that

$$\mathbf{r}(t_{\text{end}}; s, \varphi) = s \underbrace{\mathbf{r}(t_{\text{end}}; 1, \varphi)}_{\mathbf{r}_0(\varphi)} = s\mathbf{r}_0(\varphi)$$

where we define $\mathbf{r}_0(\varphi)$ as the activity generated by a pair of aiming variables with angle φ and unit norm. In polar coordinates, then, the reachable repertoire manifold is given by

$$\mathcal{R} = \{\mathbf{r}(t_{\text{end}}; s, \varphi) : s \in [0, s_{\text{max}}], \varphi \in [0, 2\pi]\},$$

The Jacobian of the inverse coordinate map, $(s, \varphi) \to \mathbf{r}(t_{\text{end}}; s, \varphi)$, is the $N \times 2$ matrix

$$\mathbf{J}(s, \varphi) = \begin{bmatrix} \frac{\partial}{\partial s}\mathbf{r}(t_{\text{end}}; s, \varphi) & \frac{\partial}{\partial \varphi}\mathbf{r}(t_{\text{end}}; s, \varphi) \end{bmatrix} = \begin{bmatrix} \mathbf{r}_0(\varphi) & s\mathbf{r}_0'(\varphi) \end{bmatrix} = \begin{bmatrix} \mathbf{r}_0(\varphi) & \mathbf{r}_0'(\varphi) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}$$

where $\mathbf{r}_0'(\varphi)$ denotes the derivative of $\mathbf{r}_0(\varphi)$ evaluated at φ. The volume element is then given by

$$dV(s,\varphi) = \sqrt{\det \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix} \det \begin{bmatrix} \|\mathbf{r}_0(\varphi)\|^2 & \mathbf{r}_0(\varphi) \cdot \mathbf{r}_0'(\varphi) \\ \mathbf{r}_0(\varphi) \cdot \mathbf{r}_0'(\varphi) & \|\mathbf{r}_0'(\varphi)\|^2 \end{bmatrix} \det \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}}$$

$$= s\sqrt{\|\mathbf{r}_0(\varphi)\|^2\|\mathbf{r}_0'(\varphi)\|^2 - (\mathbf{r}_0(\varphi) \cdot \mathbf{r}_0'(\varphi))^2}$$

$$= s\|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\|\sqrt{1 - \cos^2\omega(\varphi)}$$

$$= s\|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|$$

where $\omega(\varphi)$ is the angle between $\mathbf{r}_0(\varphi)$ and its derivative there $\mathbf{r}_0'(\varphi)$. The total volume V of the manifold is thus given by

$$V = \int_0^{s_{\max}} \int_0^{2\pi} s\|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\, d\varphi\, ds$$

$$= \frac{1}{2}s_{\max}^2 \int_0^{2\pi} \|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\, d\varphi$$

The probability density function of the uniform distribution on the reachable repertoire for $\widetilde{K} = 2$ is thus given by

$$p\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi)\right) = \frac{dV(s,\varphi)}{V}$$

The mean of this distribution, $\bar{\mathbf{r}}$, is thus

$$\bar{\mathbf{r}} = \int_0^{s_{\max}} \int_0^{2\pi} \mathbf{r}(t_{\mathrm{end}}; s, \varphi)p\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi)\right)\, ds\, d\varphi$$

$$= \frac{1}{V} \int_0^{s_{\max}} \int_0^{2\pi} s\mathbf{r}_0(\varphi)dV(s,\varphi)\, ds\, d\varphi$$

$$= \frac{1}{V} \int_0^{s_{\max}} \int_0^{2\pi} s^2\mathbf{r}_0(\varphi)\|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\, ds\, d\varphi$$

$$= \frac{1}{V}\frac{1}{3}s_{\max}^3 \int_0^{2\pi} \mathbf{r}_0(\varphi)\|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\, d\varphi$$

$$= \frac{2}{3}s_{\max}\frac{\int_0^{2\pi} \mathbf{r}_0(\varphi)\|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\, d\varphi}{\int_0^{2\pi} \|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\, d\varphi}$$

The covariance is

$$\boldsymbol{\Sigma} = \overline{\mathbf{r}\mathbf{r}^T} - \bar{\mathbf{r}}\bar{\mathbf{r}}^T,$$

where $\overline{\mathbf{r}\mathbf{r}^T}$ is the matrix of second moments,

$$
\begin{aligned}
\overline{\mathbf{r}\mathbf{r}^T} &= \int_0^{s_{\max}} \int_0^{2\pi} \mathbf{r}(t_{\mathrm{end}}; s, \varphi)\mathbf{r}(t_{\mathrm{end}}; s, \varphi)^T p\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi)\right) \mathrm{d}s\,\mathrm{d}\varphi \\
&= \frac{1}{V} \int_0^{s_{\max}} \int_0^{2\pi} s^2 \mathbf{r}_0(\varphi)\mathbf{r}_0(\varphi)^T dV(s, \varphi)\,\mathrm{d}s\,\mathrm{d}\varphi \\
&= \frac{1}{V} \int_0^{s_{\max}} \int_0^{2\pi} s^3 \mathbf{r}_0(\varphi)\mathbf{r}_0(\varphi)^T \|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\,\mathrm{d}s\,\mathrm{d}\varphi \\
&= \frac{1}{V}\frac{1}{4}s_{\max}^4 \int_0^{2\pi} \mathbf{r}_0(\varphi)\mathbf{r}_0(\varphi)^T \|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\,\mathrm{d}\varphi \\
&= \frac{1}{2}s_{\max}^2 \frac{\int_0^{2\pi} \mathbf{r}_0(\varphi)\mathbf{r}_0(\varphi)^T \|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\,\mathrm{d}\varphi}{\int_0^{2\pi} \|\mathbf{r}_0(\varphi)\|\|\mathbf{r}_0'(\varphi)\||\sin\omega(\varphi)|\,\mathrm{d}\varphi}
\end{aligned}
$$

Under linear dynamics, we have an explicit expression for the reachable repertoire so we can actually analytically evaluate these moments. From equation 45, we have that, with $\widetilde{K} = 2$ aiming variables, the activity patterns on the reachable repertoire can be expressed in the polar coordinates of the aiming variables as follows,

$$
\mathbf{r}(t_{\mathrm{end}}; s, \varphi) = \mathbf{N}_{\mathrm{end}}\mathbf{r}(0) + s\mathbf{M}_{\mathrm{end}}\begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}.
$$

The Jacobian of the inverse coordinate map is given by[1]

$$
\mathbf{J}(s, \varphi) = \begin{bmatrix} \mathbf{M}_{\mathrm{end}}\begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} & s\mathbf{M}_{\mathrm{end}}\begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix} \end{bmatrix} = \mathbf{M}_{\mathrm{end}}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix},
$$

yielding the volume element

$$
dV(s, \varphi) = \det\begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}\det\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}\sqrt{\det \mathbf{M}_{\mathrm{end}}^T\mathbf{M}_{\mathrm{end}}} = s\sigma_1^{\mathbf{M}}\sigma_2^{\mathbf{M}},
$$

where $\sigma_1^{\mathbf{M}}, \sigma_2^{\mathbf{M}}$ are the singular values of the reachable repertoire matrix $\mathbf{M}_{\mathrm{end}}$. The volume of the reachable repertoire is thus

$$
V = \int_0^{s_{\max}} \int_0^{2\pi} s\sigma_1^{\mathbf{M}}\sigma_2^{\mathbf{M}}\,\mathrm{d}\varphi\,\mathrm{d}s = \pi s_{\max}^2 \sigma_1^{\mathbf{M}}\sigma_2^{\mathbf{M}},
$$

yielding the following probability density function for the uniform distribution on the reachable repertoire,

$$
p\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi)\right) = \frac{dV(s, \varphi)}{V} = \frac{s}{\pi s_{\max}^2}
$$

---

[1]Note that if $\mathbf{r}(0) = \mathbf{0}$ then this exactly matches the above expression in terms of $\mathbf{r}_0(\varphi)$.
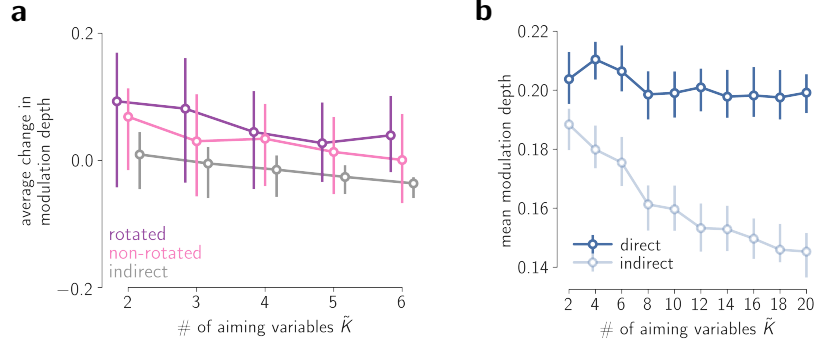
Figure 12: Changes in modulation depth under generalized re-aiming.
(12a) Average change in modulation depth of rotated, non-rotated, and indirect neurons after learning by generalized re-aiming. As in figure 7e, the changes in modulation depth under each perturbed decoder are averaged over all neurons in each sub-population. The lines show the median over all perturbed decoders, with error bars marking the upper and lower quartiles.
(12b) Average modulation depth of direct neurons (i.e. neurons recorded by the BMI, with non-zero decoding weights in **D**) and indirect neurons (neurons not recorded by the BMI) under generalized re-aiming solutions for OMP's. Modulation depths were extracted following the same procedure used to fit tuning curves in our simulation of re-aiming under credit assignment rotation perturbations (cf. equation 30 and surrounding text). For each OMP, modulation depths were averaged over all neurons in each sub-population; plotted markers show the median of these averages over all 100 sampled OMP's, error bars show upper and lower quartiles.

It is then straight-forward to derive that the mean and covariance of this distribution are given by

$$
\bar{\mathbf{r}} = \int_0^{s_{\max}} \int_0^{2\pi} \mathbf{r}(t_{\mathrm{end}}; s, \varphi) p\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi)\right) \, \mathrm{d}s \, \mathrm{d}\varphi
$$

$$
= \mathbf{N}_{\mathrm{end}}\mathbf{r}(0) + \frac{s_{\max}}{3\pi}\mathbf{M}_{\mathrm{end}} \int_0^{2\pi} \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \mathrm{d}\varphi
$$

$$
= \mathbf{N}_{\mathrm{end}}\mathbf{r}(0)
$$

$$
\boldsymbol{\Sigma} = \int_0^{s_{\max}} \int_0^{2\pi} \left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi) - \bar{\mathbf{r}}\right)\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi) - \bar{\mathbf{r}}\right)^T p\left(\mathbf{r}(t_{\mathrm{end}}; s, \varphi)\right) \, \mathrm{d}s \, \mathrm{d}\varphi
$$

$$
= \frac{1}{\pi s_{\max}^2} \int_0^{s_{\max}} \int_0^{2\pi} s^3 \mathbf{M}_{\mathrm{end}} \begin{bmatrix} \cos^2\phi & \cos\phi\sin\phi \\ \cos\phi\sin\phi & \sin^2\phi \end{bmatrix} \mathbf{M}_{\mathrm{end}}^T \, \mathrm{d}s \, \mathrm{d}\varphi
$$

$$
= \frac{s_{\max}^2}{4\pi} \mathbf{M}_{\mathrm{end}} \int_0^{2\pi} \begin{bmatrix} \cos^2\phi & \cos\phi\sin\phi \\ \cos\phi\sin\phi & \sin^2\phi \end{bmatrix} \mathrm{d}\varphi \, \mathbf{M}_{\mathrm{end}}^T
$$

$$
= \frac{s_{\max}^2}{4} \mathbf{M}_{\mathrm{end}}\mathbf{M}_{\mathrm{end}}^T
$$

This result is referenced in section II.2.3.

I.4.5   Learning-related changes in modulation depth

In a study employing credit assignment rotation perturbations in a 3D cursor reaching task, it was observed that both non-rotated and rotated neurons reduced their modulation depth after learning the perturbed decoder, and that rotated neurons reduced their modulation depth more [Jarosiewicz et al., 2008]

(see text surrounding Methodsequation 30 for how modulation depth is defined and measured). Figure 12a reveals that our model of generalized re-aiming does not reproduce this result, at least for the values of $\widetilde{K}$ we tested. Generalized re-aiming with up to 6 motor variables seems to lead to slight increases in the modulation depths of both rotated and non-rotated neurons, with marginal differences between rotated and non-rotated neurons. It is possible that using a larger number of aiming variables would lead to decreases in modulation depth consistent with the findings of this experiment; we did not test this.

Another possible resolutions to this inconsistency is that changes in modulation depth arise via synaptic plasticity, which we do not model here. This explanation is consistent with the separate observation that indirect neurons show selective decreases in modulation depth after days of practice with a given BMI decoder, but not within a single session [Ganguly et al., 2011] (although we show in figure 12b that, at least with OMP's, this phenomenon can in fact be replicated by generalized re-aiming). That changes in modulation depth only arise on longer timescales of training may also account for the fact that this phenomenon has been reported by some studies employing credit assignment rotation perturbations [Jarosiewicz et al., 2008] but not others [Chase et al., 2012, Zhou et al., 2019]. Note that differential changes in preferred direction between rotated and non-rotated neurons did replicate across these various studies. The fact that changes in modulation depth did not replicate but changes in preferred direction did possibly suggests that two distinct learning mechanims might underlie these two types of changes in motor cortical tuning curves.

Chapter 2

In this chapter, we provide a deeper mathematical analysis of our results on re-aiming in the BMI learning task of Sadtler et al. (2014). An important difficulty in doing this is that the differential equations defining our model's dynamics (equation 1) cannot be solved analytically. However, if we remove the rectified non-linearity in the dynamics, $\phi(\cdot)$, and replace it with an identity activation function so that $r_i = x_i$, these differential equations become purely linear in the dynamical variables and can be solved. This allows us to derive exact analytical expressions for the instantaneous population activity, $\mathbf{r}(t)$, the network's reachable repertoire and intrinsic manifold, and the re-aiming solutions to any given target readout and decoder (equation 4). This in turn will permit a transparent analysis of the properties of the re-aiming learning strategy in the context of this task.

Of course, an identity activation function isn't biologically realistic for a number of reasons, the most obvious of which being that it admits negative firing rates. Addressing the extent to which this shortcoming affects our results will in fact be a primary goal of our analysis in this chapter – this is explicitly discussed in section II.2.3. Analyzing the simplified linear network dynamics will in fact provide valuable insight into how exactly the rectified linear activation function used in chapter 1 influences the network's reachable repertoire, intrinsic manifold, and, by extension, the re-aiming learning strategy.

A second shortcoming of the identity activation function – which is actually shared with the rectified linear activation function – is that firing rates don't saturate and can therefore be unbounded. This shortcoming is easily addressed by ensuring that firing rates in our model are bounded within a reasonable range, simply done by enforcing stable network dynamics and bounded inputs. In this regime, the network provides a reasonable model of biological firing rates in the unsaturated regime of activity, which is typically that observed in *in vivo* recordings like those in the BMI experiments relevant here.

Analysis of linear networks has a rich tradition in theoretical neuroscience [Hennequin, 2019, Dayan and Abbott, 2001a]. Linear network dynamics have been successfully employed to model and interpret neural dynamics across a wide variety of brain areas and behavioral tasks. Various brain areas modeled with linear systems include the oculomotor integrator [Seung, 1996], the lateral intraparietal area [Ganguli et al., 2008a], and, of course, primary motor cortex [Churchland et al., 2010, Hennequin et al., 2014]. Linear networks have also been used to study the dynamics of circuits with excitatory and inhibitory populations [Murphy and Miller, 2009, Hennequin et al., 2012], to model the circuit mechanisms underlying short term memory [White et al., 2004, Ganguli et al., 2008b, Goldman, 2009, Ganguli and Latham, 2009] and

evidence integration [Ganguli et al., 2008a], and to study computational bottlenecks in reservoir computing [Susman et al., 2021].

Standing on these shoulders, in this chapter we will use linear networks to study certain properties of the re-aiming learning strategy. The mathematical muscle afforded to us by this simplification of the motor cortical dynamics will allow us to better understand some of the results in chapter 1 as well as understand how they extend to more complex settings. The chapter's results are organized into three sections, in which we will analyze the re-aiming solutions of a linear network to respectively

1. provide mathematical grounding to the the arguments presented in chapter 1 for why re-aiming fails for OMP's but succeeds for WMP's

2. connect the concept of the reachable repertoire to classical control-theoretic analyses of linear time-invariant systems, in particular to the controllability Gramian

3. illustrate the role of non-linear dynamics in producing the behavioral biases after WMP learning

Preliminary notation

Before going ahead with our analysis, we first introduce some new notation for the linear network, following standard control-theoretic conventions for linear time-invariant systems [Kao and Hennequin, 2019]. Plugging in an identity activation function, $\phi(x_i) = x_i$, into the dynamics (eqn. 1) yields the following firing rate dynamics:

$$\tau \frac{\mathrm{d}r_i}{\mathrm{d}t} = -r_i + \sum_{j=1}^{N} W_{ij}^{\mathrm{rec}} r_j + \sum_{j=1}^{M} W_{ij}^{\mathrm{in}} u_j \tag{40}$$

Putting this in matrix/vector notation, we arrive at the linear system

$$\tau \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = -\mathbf{r} + \mathbf{W}^{\mathrm{rec}}\mathbf{r} + \mathbf{W}^{\mathrm{in}}\mathbf{u}. \tag{41}$$

We next plug in $\phi(x_i) = x_i$ into the inputs equation (eqn. 2) to obtain, in vector notation,

$$\mathbf{u}(\boldsymbol{\theta}) = \mathbf{U}\boldsymbol{\theta}.$$

We next recall that, when re-aiming with $\widetilde{K}$ motor variables, we assume that $\theta_i = 0$ for all $i > \widetilde{K}$. We can therefore reduce this expression by dropping all the components of $\boldsymbol{\theta}$ that are zero, along with the corresponding columns of $\mathbf{U}$. We define the $\widetilde{K}$-dimensional vector $\widetilde{\boldsymbol{\theta}}$ that contains only the aiming variables $\theta_1, \theta_2, \ldots, \theta_{\widetilde{K}}$, and the $M \times \widetilde{K}$ matrix $\widetilde{\mathbf{U}}$ obtained by dropping columns $\widetilde{K}+1, \widetilde{K}+2, \ldots, K$ of $\mathbf{U}$, to write

$$\mathbf{u}(\boldsymbol{\theta}) = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}. \tag{42}$$

Plugging this into equation 41, we arrive at the following linear system driven by a $\widetilde{K}$-dimensional motor command

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{A}\mathbf{r} + \mathbf{B}\widetilde{\boldsymbol{\theta}} \tag{43}$$
$$\mathbf{A} = \frac{1}{\tau}\left(\mathbf{W}^{\mathrm{rec}} - \mathbf{I}\right)$$
$$\mathbf{B} = \frac{1}{\tau}\mathbf{W}^{\mathrm{in}}\widetilde{\mathbf{U}}.$$

Results

II.2.1   Properties of the reachable repertoire relevant for re-aiming

We begin by solving equation 43 to obtain

$$\mathbf{r}(t; \widetilde{\boldsymbol{\theta}}) = \underbrace{e^{\mathbf{A}t}}_{\mathbf{N}(t)}\mathbf{r}(0) + \underbrace{\left(e^{\mathbf{A}t} - \mathbf{I}\right)\mathbf{A}^{-1}\mathbf{B}}_{\mathbf{M}(t)}\widetilde{\boldsymbol{\theta}}, \tag{44}$$

where $e^{\mathbf{A}t}$ denotes the matrix exponential (a matrix of the same shape as $\mathbf{A}$). Given a fixed initial condition $\mathbf{r}(0)$, we can thus see that the reachable repertoire at time t_{end} comprises a $\widetilde{K}$-dimensional plane centered at $\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(0)$,

$$\mathbf{r}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}) = \underbrace{\mathbf{N}(t_{\mathrm{end}})}_{\mathbf{N}_{\mathrm{end}}}\mathbf{r}(0) + \underbrace{\mathbf{M}(t_{\mathrm{end}})}_{\mathbf{M}_{\mathrm{end}}}\widetilde{\boldsymbol{\theta}}. \tag{45}$$

Under silent initial conditions, $\mathbf{r}(0) = \mathbf{0}$, the dimensions occupied by the reachable repertoire are those spanned by the $\widetilde{K}$ columns of $\mathbf{M}_{\mathrm{end}}$, which we thus term the *reachable repertoire matrix*. In this case, then, the reachable repertoire's dimensionality exactly matches the number of motor variables used for re-aiming (see fig. 14b). We briefly note that this differs from the non-linear network dynamics used in chapter 1 (compare to fig. 5b), in which we found that the reachable repertoire's dimensionality exceeded the number of aiming variables (fig. 3a, in which the reachable repertoire is about eight-dimensional with $\widetilde{K} = 2$). This highlights how the non-linear dynamics can warp the reachable repertoire to occupy more dimensions than there are aiming variables.

Armed with an analytical expression for the reachable repertoire, we now set out to formalize the conditions under which the re-aiming strategy can succeed under linear dynamics. We first use equation 43 to evaluate the BMI readout at the endpoint time t_{end},

$$\mathbf{y}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}) = \underbrace{\mathbf{D}\mathbf{M}_{\mathrm{end}}}_{\boldsymbol{\Lambda}_{\mathrm{end}}}\widetilde{\boldsymbol{\theta}} + \underbrace{\mathbf{D}\left(\mathbf{N}_{\mathrm{end}}\mathbf{r}(0) - \boldsymbol{\mu}\right)}_{\mathbf{y}(t_{\mathrm{end}}; \mathbf{0})}. \tag{46}$$

We have defined here two quantities. The 2D vector $\mathbf{y}(t_{\text{end}}; \mathbf{0})$ is the BMI readout at the endpoint time in the absence of any input, $\widetilde{\boldsymbol{\theta}} = \mathbf{0}$ (in which case $\mathbf{r}(t_{\text{end}}; \widetilde{\boldsymbol{\theta}}) = \mathbf{N}(t_{\text{end}})\mathbf{r}(0)$). The $2 \times \widetilde{K}$ matrix $\boldsymbol{\Lambda}_{\text{end}}$ is the *alignment matrix*, composed of the product between the decoding matrix $\mathbf{D}$ and the reachable repertoire matrix $\mathbf{M}_{\text{end}}$. Its name derives from the fact that each of its elements are dot products between the dimensions of the reachable repertoire (the columns of $\mathbf{M}_{\text{end}}$, which we notate by $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{\widetilde{K}}$) and the decoding dimensions of the BMI decoder (the rows of $\mathbf{D}$, which we notate by $\mathbf{d}_1, \mathbf{d}_2$),

$$\Lambda_{\text{end},ij} = \mathbf{d}_i \cdot \mathbf{m}_j.$$

Thus, if the reachable repertoire spans dimensions orthogonal to the decoder's, the elements of the alignemnt matrix will be 0. On the other hand, if they are aligned, then the elements of $\boldsymbol{\Lambda}_{\text{end}}$ will be large. This alignment is thus critical for the re-aiming strategy to succeed: the $\widetilde{K}$-dimensional motor command $\widetilde{\boldsymbol{\theta}}$ only influences the BMI readout at time t_{end} via the alignment matrix $\boldsymbol{\Lambda}_{\text{end}}$, so if the alignment is low then these low-dimensional motor commands won't be able to sufficiently modulate the readouts to solve the task at hand.

To quantify this precisely in the context of a specific task, we next evaluate the mean squared error achieved by the optimal re-aiming solutions for a center-out reaching task. We first plug in equations 46 and 42 into equation 4 to solve for the re-aiming solution to a given target readout $\mathbf{y}^*$,

$$
\begin{aligned}
\widehat{\widetilde{\boldsymbol{\theta}}}(\mathbf{y}^*) &= \arg\min_{\widetilde{\boldsymbol{\theta}}} \; \left\| \boldsymbol{\Lambda}_{\text{end}}\widetilde{\boldsymbol{\theta}} + \mathbf{y}(t_{\text{end}}; \mathbf{0}) - \mathbf{y}^* \right\|^2 + \frac{\gamma}{M}\|\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}\|^2 \\
&= \left(\boldsymbol{\Lambda}_{\text{end}}^T \boldsymbol{\Lambda}_{\text{end}} + \frac{\gamma}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \right)^{-1} \boldsymbol{\Lambda}_{\text{end}}^T \left(\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0}) \right) \\
&= \left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \right)^{-1} \boldsymbol{\Lambda}_{\text{end}}^T \left(\boldsymbol{\Lambda}_{\text{end}} \left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \right)^{-1} \boldsymbol{\Lambda}_{\text{end}}^T + \gamma\mathbf{I} \right)^{-1} \left(\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0}) \right), \quad (47)
\end{aligned}
$$

where the last line follows from the Woodbury Matrix Inversion Lemma. The squared error achieved by this optimal re-aiming solution is then given by

$$
\begin{aligned}
\text{error w.r.t. } \mathbf{y}^* &= \left\| \boldsymbol{\Lambda}_{\text{end}}\widehat{\widetilde{\boldsymbol{\theta}}}(\mathbf{y}^*) + \mathbf{y}(t_{\text{end}}; \mathbf{0}) - \mathbf{y}^* \right\|^2 \\
&= \left\| \left(\boldsymbol{\Lambda}_{\text{end}} \left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \right)^{-1} \boldsymbol{\Lambda}_{\text{end}}^T \left(\boldsymbol{\Lambda}_{\text{end}} \left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \right)^{-1} \boldsymbol{\Lambda}_{\text{end}}^T + \gamma\mathbf{I} \right)^{-1} - \mathbf{I} \right) \left(\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0}) \right) \right\|^2 \\
&= \left\| -\left(\frac{1}{\gamma} \underbrace{\boldsymbol{\Lambda}_{\text{end}} \left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \right)^{-1} \boldsymbol{\Lambda}_{\text{end}}^T}_{\boldsymbol{\Sigma}_y} + \mathbf{I} \right)^{-1} \left(\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0}) \right) \right\|^2.
\end{aligned}
$$

Averaging this quantity over target readouts on the unit circle (reflecting the radial arrangement of the

reach targets in a center-out reaching task), we obtain the mean squared error,

$$\text{mean squared error} = \left\langle \left\| - \left(\frac{1}{\gamma}\boldsymbol{\Sigma}_y + \mathbf{I}\right)^{-1} (\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0})) \right\|^2 \right\rangle_{\mathbf{y}^*}$$

$$= \left\langle \left\| \left(\frac{1}{\gamma}\boldsymbol{\Sigma}_y + \mathbf{I}\right)^{-1} \mathbf{y}^* \right\|^2 \right\rangle_{\mathbf{y}^*} + \left\| \left(\frac{1}{\gamma}\boldsymbol{\Sigma}_y + \mathbf{I}\right)^{-1} \mathbf{y}(t_{\text{end}}; \mathbf{0}) \right\|^2$$

$$= \frac{1}{2}\left\| \left(\frac{1}{\gamma}\boldsymbol{\Sigma}_y + \mathbf{I}\right)^{-1} \right\|_{\mathcal{F}}^2 + \left\| \left(\frac{1}{\gamma}\boldsymbol{\Sigma}_y + \mathbf{I}\right)^{-1} \mathbf{y}(t_{\text{end}}; \mathbf{0}) \right\|^2$$

where $\| \cdot \|_{\mathcal{F}}$ denotes the Frobenius norm. In the second line and third lines above we respectively used the following properties of the uniform distribution on the unit circle:

$$\langle \mathbf{y}^* \rangle = \mathbf{0}$$

$$\left\langle \mathbf{y}^* \mathbf{y}^{*T} \right\rangle = \frac{1}{2}\mathbf{I}.$$

Noting that the mean squared error is minimized when the initial conditions are zero (so that the second term above goes to 0), we can derive the following lower bound on the mean squared error, in terms of the eigenvalues σ_1^2, σ_2^2 of the matrix $\boldsymbol{\Sigma}_y$,

$$\text{mean squared error} \geq \frac{1}{2}\left\| \left(\frac{1}{\gamma}\boldsymbol{\Sigma}_y + \mathbf{I}\right)^{-1} \right\|_{\mathcal{F}}^2 = \frac{1}{2}\sum_{i=1}^{2} \frac{1}{\left(\frac{1}{\gamma}\sigma_i^2 + 1\right)^2}. \tag{48}$$

In other words, the larger the eigenvalues of $\boldsymbol{\Sigma}_y$, the lower the mean squared error.

What does this mean in terms of the reachable repertoire and its alignment with the decoder? We can get some intuition for this equation by exploiting the random structure of the encoding weight matrix $\mathbf{U}$ to simplify our expression for $\boldsymbol{\Sigma}_y$. When the encoding weights are identically and independently distributed with 0 mean and unit variance (as they were in our simulations), the central limit theorem gives us that

$$\frac{1}{M}\left[\widetilde{\mathbf{U}}^T \widetilde{\mathbf{U}} \right]_{ij} = \frac{1}{M}\sum_{k=1}^{M} \widetilde{U}_{ki}\widetilde{U}_{kj} = \langle \widetilde{U}_{ki}\widetilde{U}_{kj} \rangle + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right) = \delta_{ij} + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right),$$

where the angular brackets $\langle \cdot \rangle$ here denote an expectation over the distribution of the encoding weights and δ_{ij} denotes the Kronecker delta ($\delta_{ij} = 1$ if $i = j$, else $\delta_{ij} = 0$). In the large M limit, then, we can make the approximation $\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \approx \mathbf{I}$, which implies

$$\boldsymbol{\Sigma}_y \approx \boldsymbol{\Lambda}_{\text{end}}\boldsymbol{\Lambda}_{\text{end}}^T$$

In other words, the eigenvalues of $\boldsymbol{\Sigma}_y$ are approximately equal to the squared singular values of the alignment matrix $\boldsymbol{\Lambda}_{\text{end}}$. Thus, the larger the singular values of the alignment matrix – that is, the larger

the alignment between $\mathbf{D}$ and $\mathbf{M}_{\mathrm{end}}$ – the lower the mean squared error achievable by the re-aiming strategy. This is simply a mathematical formalization of the same argument expressed in chapter 1: because the reachable repertoire is limited to a subspace of the full population activity state space, it must be well aligned with the decoding matrix for re-aiming to succeed.

Another way to interpret $\boldsymbol{\Sigma}_y$ is as the covariance of the readouts reachable with $\widetilde{K}$ aiming variables. We can quantify this by the covariance of readouts generated by 0-mean independently distributed aiming variables with unit variance, $\mathrm{cov}[\widetilde{\boldsymbol{\theta}}] = \mathbf{I}$, in which case we have that

$$\mathrm{cov}\left[\mathbf{y}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}\big)\right] = \mathbf{D}\left(\mathrm{cov}\left[\mathbf{r}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}\big)\right]\right)\mathbf{D}^T = \mathbf{D}\mathbf{M}_{\mathrm{end}}\mathbf{M}_{\mathrm{end}}^T\mathbf{D}^T = \boldsymbol{\Lambda}_{\mathrm{end}}\boldsymbol{\Lambda}_{\mathrm{end}}^T \approx \boldsymbol{\Sigma}_y. \tag{49}$$

The eigenvalues of $\boldsymbol{\Sigma}_y$ thus approximately quantify how spread out the reachable readouts are over the workspace. More precisely, σ_1, σ_2 approximately equal the standard deviation of the reachable readouts along two orthogonal dimensions of the workspace (given by the eigenvectors of $\boldsymbol{\Sigma}_y$). If the reachable readouts are spread over a large region of the workspace, then σ_1 and σ_2 will be large and the right-hand side of equation 48 will be small – the widely spread out reachable readouts will include the target readouts so the minimal mean squared error will be low. If, on the other hand, the reachable readouts are concentrated in a small region of the workspace, the standard deviations σ_1 and σ_2 will be small and the minimal mean squared error will be high – the limited set of reachable readouts does not include the target readouts, which therefore can't be reached.

Equation 48 accounts for why low-dimensional re-aiming fails for OMP's, which are less aligned with the reachable repertoire and therefore lead to reachable readouts confined to a small region of the workspace (cf. fig. 3d in chapter 1). In figure 13a we quantify this alignment using the singular values of the alignment matrix, σ_1, σ_2, which we refer to as the reachable readout standard deviation (as per the discussion in the preceding paragraph). This figure shows the reachable readout standard deviation for the baseline decoder and several sampled WMP's and OMP's (fit to a randomly connected linear network, see figure caption for further details), under re-aiming with $\widetilde{K} = 2$ motor variables. As expected, the reachable repertoire is more aligned with WMP's than OMP's, such that the reachable readouts under OMP's have much smaller standard deviation, well below those under the baseline decoder alignment matrix. This entails that, with OMP's, much stronger inputs are needed to reach the target readouts. As a result, given a fixed metabolic cost γ, the minimal mean squared error achievable by re-aiming is necessarily higher for OMP's than for the baseline decoder or WMP's. This is quantified in figure 13b, where we plot the theoretical lower bound on the MSE given by equation 48 as a function of γ, for the baseline decoder, an example WMP, and an example OMP (whose alignment matrix singular values are marked by stars in figure 13a). The dashed gray line marks the value of γ for which the minimal mean squared error under the baseline decoder is equal to .05; the minimal mean squared error at this value

of γ is substantially higher for the OMP than for the WMP and baseline decoder.

What happens under generalized re-aiming? Formally, as we increase the number of motor variables used for re-aiming, $\widetilde{K}$, we add columns to the $N \times \widetilde{K}$ reachable repertoire matrix $\mathbf{M}_{\mathrm{end}}$. This corresponds to an expansion in the subspace occupied by the reachable repertoire, thus increasing its overlap with the rows of any given decoding matrix $\mathbf{D}$. This is shown in figure 13c, where we plot the reachable readout standard deviation under each sampled OMP as a function of $\widetilde{K}$. Evidently, the standard deviation of the reachable readouts increases monotonically with the number of motor variables used for re-aiming, $\widetilde{K}$. This means that, at high values of $\widetilde{K}$, low mean squared error can be achieved for these OMP's. This is illustrated in figure 13d, which shows the resulting change in the lower bound on the mean squared error for the same example OMP shown in figure 13b.

We briefly remark that the theoretical bounds on the mean squared error shown in figures 13b and 13d are approximate, as we approximated $\mathbf{\Sigma}_y$ with $\mathbf{\Lambda}_{\mathrm{end}} \mathbf{\Lambda}_{\mathrm{end}}^T$ when evaluating equation 48. As discussed above, this approximation is only exact in the limit of $M \to \infty$. However, we verified in simulations (not shown) that this approximation is a good one even for finite M (we used $M = 256$), at least when the encoding weights are identically and independently distributed, $U_{ij} \sim \mathcal{N}(0, 1)$. We did not test other distributions of encoding weights. The encoding weights effectively determine the directional tuning curves of the neurons, so this assumption is equivalent to assuming a uniform distribution of preferred directions in the motor cortical population. When encoding weights are not uniformly distributed, some individual aiming variables can incur higher costs than others (through the quadratic penalty term $\|\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}\|^2 = \widetilde{\boldsymbol{\theta}}^T \widetilde{\mathbf{U}}^T \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}$), in which case the reachable repertoire covariance, $\mathbf{\Sigma}_y$, – which is evaluated as an expectation over a uniform distribution on the reachable repertoire – no longer suffices to explain the mean squared error. Concetpually, however, all of the above insights remain true.

Finally, we note that these theoretical results immediately extend to any spatiotemporally separable inputs of the form

$$\mathbf{u}(\boldsymbol{\theta}) = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}a(t), \tag{50}$$

where $a(t)$ determines the temporal profile of each input. In our simulations, this was a constant, $a(t) = 1$. But other reasonable choices might be a transient delta function $a(t) = \delta(t)$ (effectively setting the initial conditions of the network [Hennequin et al., 2014, Sussillo et al., 2015]), or a slower transient like a difference of exponentials $a(t) = e^{-t/\tau_1} - e^{-t/\tau_2}$ (resembling the shape of a post-synaptic current [Dayan and Abbott, 2001b]). Ultimately, however, the choice of $a(t)$ only influences the reachable repertoire matrix, but does not alter equation 48 – the lower bound on the mean squared error still depends on the alignment between the decoder and the reachable repertoire in exactly the same way. To see this, note that the only step that changes in the above derivation is the solution to the RNN
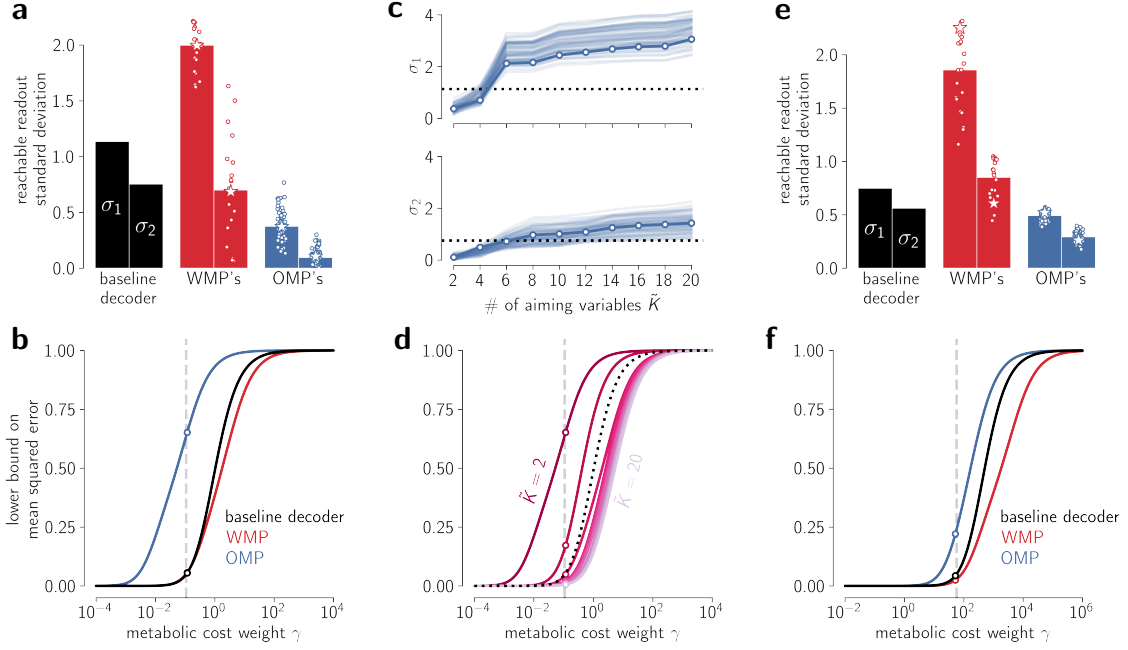
Figure 13: Theoretical bounds on re-aiming under linear dynamics.
(13a) Reachable readout standard deviation under the baseline decoder, 20 WMP's, 60 OMP's, with $\widetilde{K} = 2$ aiming variables (square root of eigenvalues of $\mathbf{\Sigma}_y$ under large M approximation $\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \approx \mathbf{I}$). We order the standard deviations by their magnitude, so in each case σ_1 is larger and σ_2 is smaller. For WMP's and OMP's, the bar height shows the median and the overlaid data points correspond to individual decoders. The stars mark example representative decoders analyzed in subsequent panels. All decoders were fit to the same randomly connected network using exactly the same procedure described in chapter 1. The network was constructed exactly as that used in the results section of chapter 1, but with an identity activation function, $\phi(x) = x$, and a smaller time constant, $\tau = 50$ms, chosen to ensure the linear dynamics yield more realistic calibration task responses (slower time constants lead to extremely low-dimensional responses under linear dynamics). These responses were significantly lower-dimensional than those of the non-linear network, so we used $\ell = 5$ dimensions for defining the intrinsic manifold and correspondingly adjusted the criteria for subsampling WMP's and OMP's (cf. section I.4.2.3). This lower dimensionality of the intrinsic manifold meant that there were fewer possible within- and outside- manifold perturbations to choose from, so we found only 20 WMP's and 60 OMP's satisfying these criteria.
(13b) Theoretical lower bound on mean squared error (MSE) in a center-out reaching task (equation 48), for the baseline decoder, one WMP, and one OMP (whose reachable readout standard deviations are marked by stars in previous panel), with $\widetilde{K} = 2$. This lower bound is plotted as a function of the metabolic cost weight γ. The vertical gray dashed line shows the metatbolic cost weight at which the minimal MSE is .05 for the baseline decoder. The open circles on this line mark the minimal MSE achievable under each decoder at this value of γ. Note that this is much lower for the WMP than the OMP.
(13c) Reachable readout standard deviation for OMP's, with different numbers of aiming variables, $\widetilde{K}$. As in previous panel σ_1 denotes the larger of the two standard deviations, and σ_2 the smaller one. Each line corresponds to a different OMP, with the darker line marking the example OMP shown in previous panel.
(13d) Theoretical lower bound on mean squared error for example OMP (same one as in previous panels), under different numbers of aiming variables, $\widetilde{K}$. The minimal mean squared error for the baseline decoder with $\widetilde{K} = 2$ is overlaid as a dotted black curve. The vertical gray line is the same as that shown in panel 13b.
(13e) Reachable readout standard deviation under the same decoders as in previous panels, with $\widetilde{K} = 2$ *time-varying* aiming variables (square root of eigenvalues of $\widetilde{\mathbf{\Sigma}}_y$ under large M approximation $\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \approx \mathbf{I}$). Stars mark the same example decoders as in previous panels.
(13f) Theoretical lower bound on mean squared error under time-varying aiming variables (equation 60), for the same three example decoders. The vertical gray dashed line again shows the metabolic cost weight at which the minimal MSE is .05 for the baseline decoder, now a much larger value than in panel 13b because allowing the aiming variables to vary freely in time permits much more energy-efficient control.

dynamics (eqn. 44), which now takes the form

$$\mathbf{r}\left(t; \widetilde{\boldsymbol{\theta}}\right) = e^{\mathbf{A}t}\mathbf{r}(0) + \left(\int_0^t e^{\mathbf{A}(t-t')}a(t')\mathrm{d}t'\right)\mathbf{B}\widetilde{\boldsymbol{\theta}}.$$

Thus, we simply re-define the reachable repertoire matrix by $\mathbf{M}_{\mathrm{end}} = \int_0^{t_{\mathrm{end}}} e^{\mathbf{A}(t_{\mathrm{end}}-t')}a(t')\mathrm{d}t'\mathbf{B}$ and the

rest of the derivation follows.

II.2.2 Time-varying motor variables: the reachable repertoire and the controllability Gramian

Equation 50 is a severe constraint on the upstream input to motor cortex. It assumes that each input u_i has exactly the same temporal profile, and that learning comprises optimizing only the spatial pattern of input (determined by the motor variables in $\widetilde{\boldsymbol{\theta}}$), while keeping its temporal evolution fixed. This helps keep the dimensionality of the learning problem low: if the full timecourses of each motor variable or input had to be individually optimized, there would be many more parameters to learn in addition to the $\widetilde{K}$ aiming variables parameterizing the spatial distribution of the input. We reiterate the fact that high-dimensional optimization is particularly difficult in the BMI learning setting, in which the motor system has to learn to control a novel effector for which no explicit information is available (unlike in natural motor learning, in which case the motor system has available to it vast information about its effectors from centuries of evolution, years of experience controlling them, and instantaneous proprioceptive feedback).

That said, we consider here the case where we allow each optimized motor variable to independently vary in time, to obtain a more complete understanding of the limits of the re-aiming learning strategy. This will moreover reveal connections between our formalism of the reachable repertoire and standard mathematical formalisms in optimal control.

When we allow the aiming variables to vary in time, our re-aiming optimization problem (eq. 4), which originally required optimizing over the space of $\widetilde{K}$-dimensional *vectors*, now becomes a variational optimization problem, in which we optimize over the space of $\widetilde{K}$-dimensional *functions* of time,

$$\widetilde{\hat{\boldsymbol{\theta}}}(t; \mathbf{y}^*) = \underset{\widetilde{\boldsymbol{\theta}}(t)}{\arg\min} \; \left\| \mathbf{y}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}\big) - \mathbf{y}^* \right\|^2 + \frac{\gamma}{M} \sum_{i=1}^{M} \frac{1}{t_{\mathrm{end}}} \int_0^{t_{\mathrm{end}}} u_i\big(\widetilde{\boldsymbol{\theta}}(t)\big)^2 \mathrm{d}t, \tag{51}$$

where $\mathbf{y}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}\big)$ now indicates the readout at time t_{end} after driving the motor cortical network with the full timecourses of the aiming variables, $\widetilde{\boldsymbol{\theta}}(t) = \begin{bmatrix} \theta_1(t) & \theta_2(t) & \dots & \theta_{\widetilde{K}}(t) \end{bmatrix}$ (as before, we will assume that any motor variables not used for re-aiming are fixed to 0, i.e. $\theta_{\widetilde{K}+1}(t) = \theta_{\widetilde{K}+2}(t) = \dots = \theta_K(t) = 0$). Note that we have changed the metabolic cost term to be an average over time, imposing a constraint on their entire timecourses. We can solve this problem using the calculus of variations. The result is Pontryagin's minimum principle, which states that the solution can be expressed in terms of the so-called control Hamiltonian,

$$H\left(\mathbf{r}, \widetilde{\boldsymbol{\theta}}, \boldsymbol{\lambda}\right) = \boldsymbol{\lambda}^T \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} + \frac{\gamma}{t_{\mathrm{end}}M} \|\mathbf{u}(\widetilde{\boldsymbol{\theta}})\|^2.$$

The first term depends on the network's dynamics (eqn. 1), and the second term depends on the time-dependent components of the cost function (namely, the second term on the right-hand side of equation

51). The N-dimensional vector $\boldsymbol{\lambda}(t)$ is a vector of time-varying Lagrange multipliers called the costate vector, with dynamics given by

$$-\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{r}} = \left(\frac{\partial}{\partial \mathbf{r}}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right)^T \boldsymbol{\lambda}, \tag{52}$$

and boundary conditions determined by the terminal cost component of the cost function (the first term on the right-hand side of equation 51, i.e. the endpoint error),

$$\boldsymbol{\lambda}(t_{\mathrm{end}}) = \frac{\partial}{\partial \mathbf{r}}\left[\left\|\mathbf{y}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}\big) - \mathbf{y}^*\right\|^2\right]. \tag{53}$$

Pontryagin's minimum principle states that the solution to equation 51 is given by the global minimum of the control Hamiltonian at each point in time. If the control Hamiltonian has a unique optimum, we can find it by solving the equation

$$\left.\frac{\partial H}{\partial \widetilde{\boldsymbol{\theta}}}\right|_{\mathbf{r}(t;\hat{\widetilde{\boldsymbol{\theta}}}),\hat{\widetilde{\boldsymbol{\theta}}}(t),\hat{\boldsymbol{\lambda}}(t)} = \mathbf{0}, \tag{54}$$

where $\mathbf{r}(t;\hat{\widetilde{\boldsymbol{\theta}}})$ and $\hat{\boldsymbol{\lambda}}(t)$ correspond to the network and costate dynamics under the optimal aiming variables, $\hat{\widetilde{\boldsymbol{\theta}}}(t)$.

Under linear dynamics, the control Hamiltonian indeed has a unique optimum and we can derive it analytically. Plugging in the linear dynamics (eqn. 43) and inputs (eqn. 42) into the control Hamiltonian, we can derive from equations 52, 53, and 54 the following system of equations,

$$-\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}t} = \mathbf{A}^T \boldsymbol{\lambda} \tag{55a}$$

$$\boldsymbol{\lambda}(t_{\mathrm{end}}) = 2\mathbf{D}^T\big(\mathbf{y}\big(t_{\mathrm{end}};\widetilde{\boldsymbol{\theta}}\big) - \mathbf{y}^*\big) \tag{55b}$$

$$\hat{\widetilde{\boldsymbol{\theta}}}(t) = -\frac{t_{\mathrm{end}}}{2\gamma}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T\hat{\boldsymbol{\lambda}}(t), \tag{55c}$$

which, together with the equation for the network dynamics (eqn. 43), form a complete set of self-consistent equations that we can solve for $\hat{\widetilde{\boldsymbol{\theta}}}(t)$. We first solve equation 55a to obtain

$$\hat{\boldsymbol{\lambda}}(t) = e^{\mathbf{A}^T(t_{\mathrm{end}}-t)}\hat{\boldsymbol{\lambda}}(t_{\mathrm{end}}) \tag{56}$$

$$\Rightarrow \hat{\widetilde{\boldsymbol{\theta}}}(t) = -\frac{t_{\mathrm{end}}}{2\gamma}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T e^{\mathbf{A}^T(t_{\mathrm{end}}-t)}\hat{\boldsymbol{\lambda}}(t_{\mathrm{end}}). \tag{57}$$

To evaluate the boundary condition, $\hat{\boldsymbol{\lambda}}(t_{\mathrm{end}})$, we derive an expression for $\mathbf{y}\big(t_{\mathrm{end}};\hat{\widetilde{\boldsymbol{\theta}}}\big)$ in terms of $\hat{\boldsymbol{\lambda}}(t_{\mathrm{end}})$, which we can then plug into equation 55b to obtain a self-consistent equation in $\hat{\boldsymbol{\lambda}}(t_{\mathrm{end}})$ that we can solve. Plugging in our expression for the aiming variables (eqn. 57) into the network dynamics (eqn. 43)

and solving the resulting ordinary differential equation, we obtain

$$\mathbf{r}(t_{\text{end}}; \hat{\tilde{\boldsymbol{\theta}}}) = \mathbf{N}_{\text{end}}\mathbf{r}(0) - \frac{t_{\text{end}}}{2\gamma} \underbrace{\int_0^{t_{\text{end}}} e^{\mathbf{A}(t_{\text{end}}-t)}\mathbf{B}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T e^{\mathbf{A}^T(t_{\text{end}}-t)}\mathrm{d}t}_{\mathbf{P}} \,\hat{\boldsymbol{\lambda}}(t_{\text{end}}). \tag{58}$$

The matrix $\mathbf{P}$ can be exactly evaluated by solving a Lyapunov equation[2]. Plugging this into the BMI readout equation (eqn. 3) to obtain $\mathbf{y}(t_{\text{end}}; \hat{\tilde{\boldsymbol{\theta}}})$, plugging this into equation 55b, and solving for $\hat{\boldsymbol{\lambda}}(t_{\text{end}})$, we then obtain

$$\hat{\boldsymbol{\lambda}}(t_{\text{end}}) = 2\mathbf{D}^T\left(\frac{t_{\text{end}}}{\gamma}\mathbf{D}\mathbf{P}\mathbf{D}^T + \mathbf{I}\right)^{-1}\left(\mathbf{y}(t_{\text{end}}; \mathbf{0}) - \mathbf{y}^*\right). \tag{59}$$

We can then plug this into equation 57 to obtain our final solution for $\hat{\tilde{\boldsymbol{\theta}}}(t)$.

To quantify the effectiveness of these optimal inputs, we next derive the minimal mean squared error achievable in a center-out reaching task, as we did in the previous section. We first derive an expression for $\mathbf{y}(t_{\text{end}}; \hat{\tilde{\boldsymbol{\theta}}})$ in terms of the target readout, $\mathbf{y}^*$, by plugging in equation 59 into equation 58 and plugging this into the BMI readout equation (eqn. 3). Plugging this into the squared error, we obtain

$$\text{error w.r.t. } \mathbf{y}^* = \left\|\left(\frac{t_{\text{end}}}{\gamma}\mathbf{D}\mathbf{P}\mathbf{D}^T\left(\frac{t_{\text{end}}}{\gamma}\mathbf{D}\mathbf{P}\mathbf{D}^T + \mathbf{I}\right)^{-1} - \mathbf{I}\right)(\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0}))\right\|^2$$

$$= \left\| -\left(\frac{t_{\text{end}}}{\gamma}\underbrace{\mathbf{D}\mathbf{P}\mathbf{D}^T}_{\widetilde{\boldsymbol{\Sigma}}_y} + \mathbf{I}\right)^{-1}(\mathbf{y}^* - \mathbf{y}(t_{\text{end}}; \mathbf{0}))\right\|^2$$

Averaging this quantity over radial reach targets (cf. derivation of equation 48), we arrive at the following lower bound on the mean squared error,

$$\text{mean squared error} \geq \frac{1}{2}\left\|\left(\frac{t_{\text{end}}}{\gamma}\widetilde{\boldsymbol{\Sigma}}_y + \mathbf{I}\right)^{-1}\right\|_{\mathcal{F}}^2 = \frac{1}{2}\sum_{i=1}^{2}\frac{1}{\left(\frac{t_{\text{end}}}{\gamma}\sigma_i^2 + 1\right)^2}, \tag{60}$$

where σ_1^2, σ_2^2 are the eigenvalues of the 2×2 matrix $\widetilde{\boldsymbol{\Sigma}}_y$. This is an analogous result to equation 48, but with the eigenvalues of the matrix $\widetilde{\boldsymbol{\Sigma}}_y = \mathbf{D}\mathbf{P}\mathbf{D}^T$ taking the place of those of the matrix $\boldsymbol{\Sigma}_y = \boldsymbol{\Lambda}_{\text{end}}\boldsymbol{\Lambda}_{\text{end}}^T$

---

[2]Using the change of variables $t' = t_{\text{end}} - t$, we can express the matrix $\mathbf{P}$ as

$$\mathbf{P} = \int_0^{t_{\text{end}}} e^{\mathbf{A}t'}\mathbf{B}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T e^{\mathbf{A}^T t'}\mathrm{d}t'$$

$$= \underbrace{\int_0^{\infty} e^{\mathbf{A}t'}\mathbf{B}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T e^{\mathbf{A}^T t'}\mathrm{d}t'}_{\mathbf{P}_\infty} - \int_{t_{\text{end}}}^{\infty} e^{\mathbf{A}t'}\mathbf{B}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T e^{\mathbf{A}^T t'}\mathrm{d}t'$$

$$= \mathbf{P}_\infty - \int_0^{\infty} e^{\mathbf{A}(t''+t_{\text{end}})}\mathbf{B}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T e^{\mathbf{A}^T(t''+t_{\text{end}})}\mathrm{d}t''$$

$$= \mathbf{P}_\infty - e^{\mathbf{A}t_{\text{end}}}\mathbf{P}_\infty e^{\mathbf{A}^T t_{\text{end}}},$$

where in the penultimate line we used the substitution $t'' = t' - t_{\text{end}}$. The matrix $\mathbf{P}_\infty$ is then given by the solution to the Lyapunov equation

$$\mathbf{A}\mathbf{P}_\infty + \mathbf{P}_\infty\mathbf{A}^T + \mathbf{B}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T = \mathbf{0},$$

which we solved using the Python function *solve_lyapunov* from the scipy *linalg* subpackage.

(note also the presence of t_{end} on the right-hand side, which we will return to below).

This result accordingly carries an analagous interpretation in terms of the readouts reachable with $\widetilde{K}$ variables. Recall that the matrix $\mathbf{\Sigma}_y$ is approximately given by the covariance of the readouts reachable with $\widetilde{K}$ spatiotemporally separable aiming variables, quantified by averaging over readouts produced by 0-mean independently distributed aiming variables with a fixed temporal profile (eqn. 49). The matrix $\widetilde{\mathbf{\Sigma}}_y$ is its analog in the case of aiming variables varying freely in time: the covariance of the readouts reachable with $\widetilde{K}$ time-varying aiming variables, quantified by averaging over readouts produced by aiming variables following independent 0-mean white noise processes. To see that this is the case, note that, if we take the approximation $\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}} \approx \mathbf{I}$, then the $N \times N$ matrix $\mathbf{P}$ is the so-called finite-time controllability Gramian. Under the dynamics in equation 43, this matrix is equal to the covariance of the network activity at time t_{end}, $\mathbf{r}(t_{\mathrm{end}};\widetilde{\boldsymbol{\theta}})$, when driven by independent white noise processes $\theta_1(t), \theta_2(t), \ldots, \theta_{\widetilde{K}}(t)$ [Kwakernaak and Sivan, 1972, Kao and Hennequin, 2019]. The matrix $\mathbf{DPD}^T$ is thus the covariance of the readouts at that time, $\mathbf{y}(t_{\mathrm{end}};\widetilde{\boldsymbol{\theta}})$.

The controllability Gramian $\mathbf{P}$ thus takes the place of the matrix $\mathbf{M}_{\mathrm{end}}\mathbf{M}_{\mathrm{end}}^T$. This latter matrix can be seen as the covariance over activity patterns in the reachable repertoire (eqn. 49; see also section I.4.4 for a more formal derivation of this fact). The matrix $\mathbf{P}$ is thus its analog when the aiming variables are allowed to vary freely in time, providing an extension of the reachable repertoire formalism to this more complex setting.

Recently, Kao & Hennequin (2019) proposed to use the controllability Gramian to quantify the "directions in state space that the network activity is most inclined to visit." Our analysis reveals that using the controllability Gramian to model the reachable repertoire implicitly assumes that no constraints are placed on the spatiotemporal structure of the upstream inputs. This may not be the case during learning, in which such constraints may play a critical role in reducing the number of parameters to be learned to facilitate efficient learning. In our model of re-aiming, this is done by enforcing that the aiming variables be spatiotemporally separable, in which case the matrix $\mathbf{M}_{\mathrm{end}}\mathbf{M}_{\mathrm{end}}^T$, rather than $\mathbf{P}$, provides the correct description of the directions encompassed by the network's reachable repertoire. Similarly, the spatiotemporal structure of the inputs driving motor cortex during the calibration task will affect the shape of the so-called intrinsic manifold estimated from the population activity evoked by this task. Since the calibration task typically comprises viewing spatiotemporally simple stimuli or executing simple behaviors, it stands to reason that the inputs will also be spatiotemporally simple. This suggests the controllability Gramian might not be a good model of the intrinsic manifold either.

We can see how removing constraints on the aiming variables affects the re-aiming strategy by comparing the readouts reachable with aiming variables constant in time (equation 48) and aiming variables varying freely in time (equation 60). Figure 13e shows the standard deviation of the readouts reachable with $\widetilde{K} = 2$ time-varying aiming variables, under the same baseline and perturbed decoders used in figure
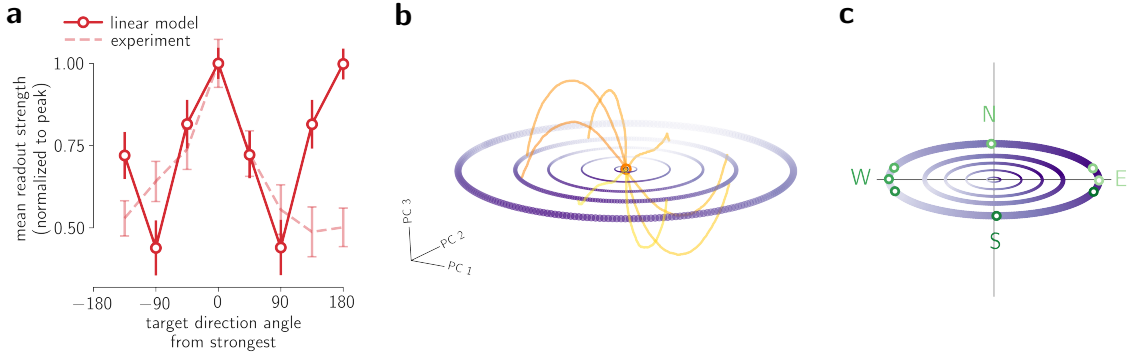
Figure 14: Re-aiming under linear dynamics does not replicate empirical biases in short-term learning.
(14a) Mean maximal readout strength over all 20 sampled WMP's, obtained following the same procedure used in figure 4c. Dashed line shows the mean readout strength during best trials of WMP control in the data of Sadtler et al. (2014), and curves are normalized to their peaks to aid comparison of model and experimental data, exactly as in figure 4c. Error bars show standard error of the mean.
(14b) Activity patterns from the reachable repertoire with $\widetilde{K} = 2$ aiming variables, at endpoint time $t_{\mathrm{end}} = 1000$ms (following the same procedure as in chapter 1 for selecting the maximum accessible motor command magnitude s_{max}, cf. section I.3.4). Calibration task responses to each of the eight radial reach stimuli are overlaid in shades of orange. Formatting conventions exactly as in figures 3c and 5b. These N-dimensional activity patterns are projected onto the top two principal components of the reachable repertoire (PC1 and PC2) and the orthogonal dimension capturing the most calibration task response variance (PC3). The reachable repertoire is exactly 2D, so PC1 and PC2 capture 100% of the variance in activity patterns within it (cf. equation 45). The small open black circle at the center marks the origin of the state space. The orange and purple open circles at the origin mark the calibration task mean $\boldsymbol{\mu}$ the reachable repertoire mean $\bar{\mathbf{r}}$ in purple, respectively. The latter is barely visible because they overlap almost completely.
(14c) Readouts reachable through the example WMP from figure 13. The reachable readouts with largest readout strengths in each target direction are marked in green, following the same conventions as figure 4a. Maximal readout strengths are evidently much higher for the western and eastern target directions than the northern and southern targets, yielding a bimodal bias.

13a, now quantified by the square roots of the eigenvalues of $\widetilde{\boldsymbol{\Sigma}}_y$ instead of $\boldsymbol{\Sigma}_y$. Note that the reachable readout standard deviation under OMP's is more similar to that under the baseline decoder in this case (compare to fig. 13a). This reflects the fact that the extra flexibility in the spatiotemporal structure of the aiming variables permits the production of activity patterns outside the intrinsic manifold even with only $\widetilde{K} = 2$ aiming variables. Recall that the intrinsic manifold is measured from activity generated by the calibration task stimuli, which are modelled by driving the network with constant motor variables mirroring the spatiotemporal structure of the stimuli. Permitting the motor variables to vary in time for re-aiming naturally allows the production of novel activity patterns not evoked by the calibration task. The behavioral outcome is that the minimum achievable mean squared error with OMP's is substantially lower with time-varying aiming variables than with constant aiming variables (compare blue curves in figures 13b and 13f).

Note as well that lower mean squared error is achievable with much higher metabolic cost weights (compare horizontal axes of figures 13b and 13f). Mathematically, this is due to the presence of the factor of t_{end} in the denominator on the right-hand side of equation 60. Conceptually, this reflects the fact that the aiming variables are being used much more efficiently. The extra flexibility provided by allowing the aiming variables to vary freely in time permits much more efficient control of the network's dynamics.

II.2.3 Linear dynamics predict biases not present in empirical data

These simple linear network dynamics are able to successfully model differences in WMP and OMP learning. Can they also reproduce the biases observed in short term learning of WMP's? It turns out the answer is no: instead of a unimodal bias, the linear network produces a bimodal bias (fig. 14a). We dedicate this section to discussing why, as the reasons behind this reveal the key geometric properties of the reachable repertoire needed to replicate the unimodal bias observed in the data. These properties highlight the important role played by the rectified linear activation function, $\phi(\cdot)$, in the dynamics of the motor cortical model presented in chapter 1 (equation 1).

We begin by recalling that the unimodal bias observed in chapter 1 arises from a particular geometric relationship between the reachable repertoire and the calibration task responses (via equations 6 and 7). Because firing rates are strictly positive under the rectified linear activation function, these two ensembles of activity patterns reside strictly within the all-positive orthant of state space. Their means, $\bar{\mathbf{r}}$ and $\boldsymbol{\mu}$, are thus non-zero vectors, also residing within this orthant. Both point in this similar direction but differ in their magnitudes (fig. 5b) because the reachable repertoire includes activity patterns generated by motor commands with stronger magnitudes (fig. 3b).

Under linear dynamics, in contrast, activity patterns are not isolated to a single orthant of state space. Because firing rates can be positive or negative, activity patterns can be displaced from the origin along any direction in state space, making it possible for them to average to 0. In fact, under silent initial conditions ($\mathbf{r}(0) \approx 0$), both the reachable repertoire and calibration task responses are mean 0[3]. Notably, this is the case independently of the magnitudes of the motor commands used to generate the activity patterns in the reachable repertoire. This is visualized in figure 14b, in which we plot a PCA projection of the reachable repertoire under linear dynamics and silent initial conditions, with the network's responses to each calibration task stimulus overlaid. Both of these ensembles of activity are centered at the origin, with stronger motor command magnitudes producing activity patterns further away from the origin but still evenly distributed around it.

The result is that the two means match, $\bar{\mathbf{r}} \approx \boldsymbol{\mu} \approx \mathbf{0}$ (cf. footnote 3), so the reachable readouts are centered at the origin (equation 6),

$$\bar{\mathbf{y}} = \mathbf{D} \left(\bar{\mathbf{r}} - \boldsymbol{\mu} \right) \approx \mathbf{0}$$

---

[3] From our above expression for the reachable repertoire (equation 45) it is straight-forward to see that the reachable repertoire mean is centered at the propagated initial conditions,

$$\bar{\mathbf{r}} = \mathbf{N}(t_{\text{end}})\mathbf{r}(0)$$

(this is more rigorously derived in section I.4.4). The mean of the calibration task responses is given by

$$\boldsymbol{\mu} = \langle\langle \mathbf{r}(t; \widetilde{\boldsymbol{\theta}}) \rangle_t \rangle_{\widetilde{\boldsymbol{\theta}}} = \langle \mathbf{N}(t) \rangle_t \mathbf{r}(0) + \langle \mathbf{M}(t) \rangle_t \langle \widetilde{\boldsymbol{\theta}} \rangle_{\widetilde{\boldsymbol{\theta}}} = \langle \mathbf{N}(t) \rangle_t \mathbf{r}(0)$$

where $\langle \cdot \rangle_t$ is an expectation over the duration of each calibration task trial (from $t = 0$ to $t = 1000$ms), and $\langle \cdot \rangle_{\widetilde{\boldsymbol{\theta}}}$ is an expectation over the 2D motor commands elicited by the calibration task stimuli. These are given by the coordinates of the eight radial reach targets, which are evenly distributed on the unit circle (see Methods) and therefore have zero mean, $\langle \widetilde{\boldsymbol{\theta}} \rangle_{\widetilde{\boldsymbol{\theta}}} = \mathbf{0}$. It thus follows that, if $\mathbf{r}(0) \approx \mathbf{0}$, then $\bar{\mathbf{r}} \approx \boldsymbol{\mu} \approx \mathbf{0}$.

This can be seen in figure 14c, where we plot the readouts reachable through one example WMP. Unlike with the non-linear network analyzed in chapter 1, the readouts reachable under linear dynamics are centered at the origin (compare to figs. 4a, 5a), eliminating the source of the unimodal bias observed in chapter 1. However, it is evident in figure 14c that biases still exist, as larger readouts can be produced in some directions than others[4]. But note the elliptical symmetry of the reachable readouts: the readout strengths that can be produced in one direction match those in the opposite direction. Indeed, it is easy to verify that, under linear dynamics and silent initial conditions, the maximal readout strength satisfies[5]

$$\rho^{\max}\left(\mathbf{y}^*\right) = \rho^{\max}\left(-\mathbf{y}^*\right).$$

Any biases in maximal readout strength are thus necessarily bimodal, with peaks separated by 180°. This is exactly what we see in the maximal readout strength curve in figure 14a.

In sum, the absence of a rectified linearity permits activity patterns to live anywhere in state space, which in turn means that the reachable repertoire and calibration task responses both average to 0 under silent initial conditions. This eliminates the source of the unimodal bias exhibited by the non-linear dynamics analyzed in chapter 1. Our model of re-aiming only captures the empirical pattern of biases during short-term WMP learning when positive firing rates are enforced via a rectified linear activation function in the dynamics. This restricts the generated activity patterns to the positive orthant of state space, ensuring that the reachable repertoire mean and calibration task mean differ when the magnitudes of the motor commands underlying these two ensembles of activity do, shifting the centroid of the reachable readouts away from the origin.

Discussion

As the saying goes, "all models are wrong, but some are useful" [Box, 1976]. In this chapter we analyzed a simple model that we knew was wrong – a model of motor cortical population activity that admitted negative firing rates. But due to its simple linear dynamics, we were able to deeply understand it through mathematical analysis. First, this allowed us to more precisely state and derive the geometric arguments made in chapter 1 about why the re-aiming strategy succeeds for WMP's but not for OMP's. Second,

---

[4]We can quantify this by the covariance of the reachable readouts, $\mathbf{\Lambda}_{\mathrm{end}}\mathbf{\Lambda}_{\mathrm{end}}^T$ (equation 49), which tells us how spread out the readouts are along each direction in the workspace; the directions with strongest/weakest readouts are the directions with highest/lowest spread. These two directions are the eigenvectors of the covariance, and the spread of the readouts along these two directions are the (square roots of the) associated eigenvalues, σ_1, σ_2. If these two eigenvalues are the same, then the readouts are equally spread out along the two directions, and equally strong readouts can be produced in both direction. If, on the other hand, they are different, then the readouts are more spread out in one direction than the other, meaning that stronger readouts can be produced in that direction. Indeed, we find that the two eigenvalues of the reachable readout are typically very different under WMP's, but less so under the baseline decoder (fig. 13a).

[5]
$$\rho^{\max}\left(\mathbf{y}^*\right) = \max_{\|\widetilde{\boldsymbol{\theta}}\| \leq s_{\max}} \mathbf{y}^* \cdot \mathbf{y}(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}) = \max_{\|\widetilde{\boldsymbol{\theta}}\| \leq s_{\max}} \mathbf{y}^* \cdot \mathbf{\Lambda}_{\mathrm{end}}\widetilde{\boldsymbol{\theta}} = s_{\max}\|\mathbf{\Lambda}_{\mathrm{end}}^T\mathbf{y}^*\| = s_{\max}\| -\mathbf{\Lambda}_{\mathrm{end}}^T\mathbf{y}^*\| = \rho^{\max}\left(-\mathbf{y}^*\right)$$

The first equality follows by definition (equation 20) and the second equality from equation 46 with $\mathbf{r}(0) = 0$. The third equality assumes that $\mathbf{\Lambda}_{\mathrm{end}}$ has full rank 2.

this allowed us to easily extend the concept of the reachable repertoire to a more complex setting, in which the aiming variables are allowed to fluctuate in time. In addition to extending the validity of our arguments from chapter 1 to this new setting, mathematical analysis of the re-aiming solutions in this case revealed deep connections to classical control-theoretic concepts and previous approaches to modeling the intrinisc manifold [Kao and Hennequin, 2018, Kao and Hennequin, 2019].

Most importantly, analysis of this simplified linear model actually provided insight into the properties of the non-linear model from chapter 1. By analyzing the reachable reperotire under linear dynamics, we were able to tease out the precise influence of the rectified linear activation function on its geometry. This revealed the importance of modeling the non-negativity of firing rates for capturing the behavioral biases present in the WMP learning data from Sadtler et al. (2014).

It is important to acknowledge, however, that many more assumptions beyond a linear activation function were implicit in our analyses in this chapter. One of the most important such assumptions was that of silent initial conditions, $\mathbf{r}(0) = 0$. This assumption made both in the linear analyses in this chapter as well as in our simulations of the non-linear network in chapter 1. Assuming completely silent initial conditions is a strong assumption, but it is a well motivated one: the primate does not know what the reach target will be at the start of the trial (so the initial conditions should be roughly the same across all trials), and, under linear dynamics, the mean squared error is minimized when the initial conditions are silent (cf. derivation of inequality 48). On a more practical note, numerical and analytical treatment of the non-linear network becomes substantially more difficult under non-zero initial conditions (cf. equation 10), so we leave a deeper analysis along these lines for future work.

Chapter 3

All of the models we have considered so far are models of *open-loop* control: once the optimal aiming variables are specified (i.e. by solving equation 4), they are used to drive the motor cortical population for the duration of the movement, unchanged until the pre-specified endpoint time t_{end}. Any errors encountered along the way – either due to noise or suboptimal specification of the motor commands – are thus ignored. A better strategy would be *closed-loop* control, wherein errors observed through sensory feedback are used to adaptively modify the aiming variables online. Under this strategy, errors that are encountered along the way can be corrected, thus improving the accuracy of the desired BMI output. Such closed-loop control strategies are well known to be optimal in the presence of noise [Kwakernaak and Sivan, 1972], and humans motor behavior exhibits hallmarks of (often optimal) closed-loop control [Keele and Posner, 1968, Todorov and Jordan, 2002, Todorov, 2004, Scott, 2004, Shadmehr et al., 2010].

Of particular relevance here is evidence that non-human primates engage in closed-loop control in BMI tasks [Golub et al., 2015, Stavisky et al., 2017, Shanechi et al., 2017]. This is perhaps most evident in the experiments done by Stavisky et al. (2017), who studied motor cortical dynamics following a sudden perturbation of the cursor position during a center-out BMI reaching task. The authors observed that motor cortical activity was modulated about 80ms after the perturbation, in a perturbation- and decoder- specific manner. This indicates that sensory feedback indeed influences motor cortical activity in a closed-loop manner over the course of a BMI reach. Moreover, the authors found that the way in which this sensory information was fed back to motor cortex was tailored specifically to the task at hand: the population response to the cursor perturbation was initially isolated to a subspace orthogonal to the BMI decoder, requiring about 50ms to align itself with the BMI decoder's decoding space so as to produce a corrective movement of the BMI cursor. These observations provide strong evidence that the motor system is indeed exploiting continuous sensory feedback to adaptively modify the BMI readouts.

In this chapter, we consider re-aiming-based models of closed-loop control, in which the aiming variables are continuously updated in response to sensory feedback. We first derive a model of closed-loop control from our original re-aiming objective, yielding a simple and intuitive architecture for closed-loop re-aiming, under linear dynamics. We will find that the model derived from this objective in fact suffers from severe suboptimalities, which can be readily understood from its transparent analytical derivation. In the next part, we attempt to address these suboptimalities by parameterizing the closed-

loop architecture of the first model and optimizing its parameters for center-out reaching. This will allow us to build models of closed-loop control under non-linear dynamics that are tailored to a given BMI decoder. We will evaluate these models of closed-loop re-aiming on the BMI learning task of Sadtler et al. (2014), and find that the same differences between WMP and OMP learning identified in the previous chapters hold even in this more complex – and biologically relevant – setting.

Results

III.1.1 The limit of $t_{\text{end}} \to 0$: myopic closed-loop re-aiming

Consider what happens after executing a movement according to our simple open-loop control model of re-aiming. First, the subject's motor system computes the optimal motor command, $\hat{\boldsymbol{\theta}}(\mathbf{y}^*)$, according to equation 4. This motor command is then used to drive the motor cortical population, leading to the production of the BMI readout $\mathbf{y}(t_{\text{end}}; \hat{\boldsymbol{\theta}}(\mathbf{y}^*))$ at the endoint time t_{end}. Suppose now that, due to noise or metabolic constraints, this readout does not exactly match the target readout, $\mathbf{y}^*$. A reasonable action to take at this point might be to compute a new motor command that will produce a readout closer to the target. One possible strategy is to simply re-aim again: solve equation 4 again, but now with the initial conditions, $\mathbf{r}(0)$, fixed to the current motor cortical firing rates, $\mathbf{r}(t_{\text{end}})$. This will yield a new motor command that guarantees to produce a readout at time $2t_{\text{end}}$ near the target, hopefully closer than it was at time t_{end}. Iterating this over and over again leads to a kind of intermittent closed-loop control strategy, whereby motor commands are updated every t_{end}ms. We refer to this strategy as *intermittent re-aiming*. There is evidence that humans engage in such forms of intermittent closed-loop control [Gross et al., 2002, Gawthrop et al., 2011].

To get a better handle of what this strategy entails, we consider the case of linear dynamics (i.e. an identity activation function, $\phi(x) = x$), for which we have an explicit expression for the optimal aiming variables, $\hat{\tilde{\boldsymbol{\theta}}}(\mathbf{y}^*) = \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_2 & \dots & \hat{\theta}_{\widetilde{K}} \end{bmatrix}$. Specifically, we have that, for a given target $\mathbf{y}^*$ and initial condition $\mathbf{r}(0)$, the optimal aiming variables take the form

$$\hat{\tilde{\boldsymbol{\theta}}} = \mathbf{G}(t_{\text{end}}) \Big(\mathbf{y}^* - \mathbf{D} \left(\mathbf{N}(t_{\text{end}}) \mathbf{r}(0) - \boldsymbol{\mu} \right) \Big), \tag{61}$$

where the $N \times N$ matrix $\mathbf{N}(t_{\text{end}})$ and $\widetilde{K} \times 2$ matrix $\mathbf{G}(t_{\text{end}})$ depend on the endpoint time t_{end}, the network's intrinsic dynamics, and the BMI decoding matrix (see equations 47, 44). From a biological perspective, this solution contains two distinct components:

1. a sensory input component, $\mathbf{G}(t_{\text{end}})\mathbf{y}^*$, that depends only on the visually presented reach target, and

2. a neural input term, $\mathbf{G}(t_{\text{end}})\mathbf{D}(\mathbf{N}(t_{\text{end}})\mathbf{r}(0) - \boldsymbol{\mu})$, that depends on the current motor cortical state

prior to driving the network, $\mathbf{r}(0)$.

In the open-loop control model discussed previously, the neural input term can be treated as effectively constant, since we assume that the motor cortical initial conditions, $\mathbf{r}(0)$, are fixed across trials. Learning the full mapping from reach targets $\mathbf{y}^*$ to optimal aiming variables $\hat{\tilde{\boldsymbol{\theta}}}$ thus requires learning a low-dimensional sensorimotor mapping, $\mathbf{G}(t_{\mathrm{end}})$, together with a single low-dimensional vector of constants, $\mathbf{G}(t_{\mathrm{end}})\mathbf{D}(\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(0) - \boldsymbol{\mu})$. Such a low-dimensional learning problem is amenable to gradient-free optimization, permitting efficient learning in the context of BMI control.

Under the intermittent re-aiming strategy, on the other hand, motor cortical firing rates change as new re-aiming solutions are computed, so the neural input term $\mathbf{DN}(t_{\mathrm{end}})\mathbf{r}(0)$ can no longer be treated as constant. More concretely, the intermittent re-aiming strategy requires computation of the following sequence of aiming variables:

$$
\hat{\tilde{\boldsymbol{\theta}}}(\mathbf{y}^*) = \begin{cases}
\mathbf{G}(t_{\mathrm{end}})\Big(\mathbf{y}^* - \mathbf{D}\left(\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(0) - \boldsymbol{\mu}\right)\Big) & \text{for } t \in [0, t_{\mathrm{end}}) \\[2mm]
\mathbf{G}(t_{\mathrm{end}})\Big(\mathbf{y}^* - \mathbf{D}\left(\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(t_{\mathrm{end}}) - \boldsymbol{\mu}\right)\Big) & \text{for } t \in [t_{\mathrm{end}}, 2t_{\mathrm{end}}) \\[2mm]
\mathbf{G}(t_{\mathrm{end}})\Big(\mathbf{y}^* - \mathbf{D}\left(\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(2t_{\mathrm{end}}) - \boldsymbol{\mu}\right)\Big) & \text{for } t \in [t_{\mathrm{end}}, 3t_{\mathrm{end}}) \\[2mm]
\vdots
\end{cases}
$$

At each iteration of re-aiming, the optimal motor command needs to take into account the current state of the system in order to drive the network towards a better state. Information about the current state comes in the form of the neural input term, $\mathbf{G}(t_{\mathrm{end}})\mathbf{D}(\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(t) - \boldsymbol{\mu})$. Learning the full mapping from reach targets $\mathbf{y}^*$ to optimal aiming variables $\hat{\tilde{\boldsymbol{\theta}}}$ thus now requires additionally learning a high-dimensional mapping from motor cortical firing rates to aiming variables, the $\widetilde{K} \times N$ matrix $\mathbf{G}(t_{\mathrm{end}})\mathbf{DN}(t_{\mathrm{end}})$. This increases the dimensionality of the learning problem by a factor of N, making efficient learning much more challenging.

A simple – yet naïve, as we'll see below – fix to this problem is to update the motor commands continuously in time, rather than updating them intermittently every t_{end}ms. To see why, note the precise role played by the neural input term in a given iteration of re-aiming. The 2D vector $\mathbf{D}(\mathbf{N}(t_{\mathrm{end}})\mathbf{r}(t) - \boldsymbol{\mu})$ on which it depends is the readout that would be produced at time $t + t_{\mathrm{end}}$ in the absence of input (cf. equation 46 and surrounding text). The neural input term thus in some sense "looks ahead" to compensate (or exploit) the readout that would be autonomously produced t_{end}ms from now through the motor cortical network's dynamics alone, without any upstream input. In the limit of $t_{\mathrm{end}} \to 0$, however, such prescient information is obviated, as one need only look at the instantaneous state of the BMI readout, $\mathbf{y}(t)$, to see what it will be at time $t + t_{\mathrm{end}}$. Indeed, in this limit, $\mathbf{N}(t_{\mathrm{end}}) = e^{\mathbf{A}t_{\mathrm{end}}} \to \mathbf{I}$, so that the neural feedback term becomes purely sensory, depending only on the current BMI readout,

$\mathbf{y}(t)$,

$$\lim_{t_{\text{end}} \to 0} \mathbf{G}(t_{\text{end}})\mathbf{D}(\mathbf{N}(t_{\text{end}})\mathbf{r}(t) - \boldsymbol{\mu}) = \mathbf{GD}(\mathbf{r}(t) - \boldsymbol{\mu}) = \mathbf{Gy}(t)$$

The optimal sensorimotor map in this limit, $\mathbf{G}$, is given by[6]

$$\mathbf{G} = \lim_{t_{\text{end}} \to 0} \mathbf{G}(t_{\text{end}}) = \frac{1}{\gamma\tau}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\widetilde{\mathbf{U}}^T\mathbf{W}^{\text{in}\,T}\mathbf{D}^T,$$

where $\widetilde{\mathbf{U}}$ is the $M \times \widetilde{K}$ matrix including only the first $\widetilde{K}$ columns of the full $M \times K$ encoding matrix $\mathbf{U}$. Plugging these limits back into equation 61 yields the following optimal time-varying aiming variables:

$$\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*) = \mathbf{G}(\mathbf{y}^* - \mathbf{y}(t)). \tag{62}$$

The result is an intuitive error feedback control strategy, whereby the instantaneous error, $\mathbf{y}^* - \mathbf{y}(t)$, is continually fed back to the motor system to modify the aiming variables online. Learning this strategy simply requires learning a low-dimensional sensorimotor mapping, the $\widetilde{K} \times 2$ matrix $\mathbf{G}$, which maps 2D sensory inputs (the displacement of the current readout from the target readout) to $\widetilde{K}$ aiming variables. We refer to this strategy as *myopic closed-loop re-aiming*; the rationale behind this name will become evident below.

Results from simulation of this control strategy are shown in figure 16. In figure 15a, we plot the squared error achieved by myopic closed-loop re-aiming with $\widetilde{K} = 2$ aiming variables, for a single target readout and under each BMI decoder. Under the baseline decoder, this control strategy succeeds: the squared error quickly falls to a low value and remains low for the duration of this time window. Under OMP's, on the other hand, the error never falls below about 0.3, indicating that the target readout is never reached. Under WMP's, we observe an entirely different phenomenon: the error quickly falls to a low value but subsequently rises to very high values above 1.0. This indicates that, under these decoders, this target readout can be reached, but it cannot be maintained beyond a short period of time. To confirm that this is the case for all other target readouts as well, we calculate the minimal error achieved over this time window for each of the eight radial target readouts, and then average. This

---

[6]To derive this, we first use the Taylor series expansion of the reachable repertoire matrix (equation 44) around $t_{\text{end}} = 0$ to write

$$\mathbf{M}(t_{\text{end}}) = t_{\text{end}}\mathbf{B} + \mathcal{O}(t_{\text{end}}^2) \Rightarrow \boldsymbol{\Lambda}_{\text{end}} = t_{\text{end}}\mathbf{DB} + \mathcal{O}(t_{\text{end}}^2)$$

Plugging this into the re-aiming solution (equation 47), we get that, to first order in t_{end}, the optimal sensorimotor map is given by

$$\mathbf{G}(t_{\text{end}}) = t_{\text{end}}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T\mathbf{D}^T\left(\gamma\mathbf{I} + \mathcal{O}(t_{\text{end}}^2)\right)^{-1} + \mathcal{O}(t_{\text{end}}^2)$$

In the limit of $t_{\text{end}} \to 0$, then $\mathbf{G}(t_{\text{end}}) \to \mathbf{0}$. This is because, as $t_{\text{end}} \to 0$, infinitely strong input is needed to move the readout towards the target in t_{end}ms. Under any non-zero metabolic cost weight, $\gamma > 0$, then, the metabolic cost term will dominate the loss function and the minimum will be at $\dot{\widetilde{\boldsymbol{\theta}}} = \mathbf{0}$. A simple way to get around this is to let the metabolic cost weight scale with t_{end}, so that it gets infinitesimally small as we take the limit of $t_{\text{end}} \to 0$. This results in

$$\mathbf{G} = \lim_{t_{\text{end}} \to 0}\left\{t_{\text{end}}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T\mathbf{D}^T\left(t_{\text{end}}\gamma\mathbf{I} + \mathcal{O}(t_{\text{end}}^2)\right)^{-1} + \mathcal{O}(t_{\text{end}}^2)\right\} = \frac{1}{\gamma}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\mathbf{B}^T\mathbf{D}^T$$

97
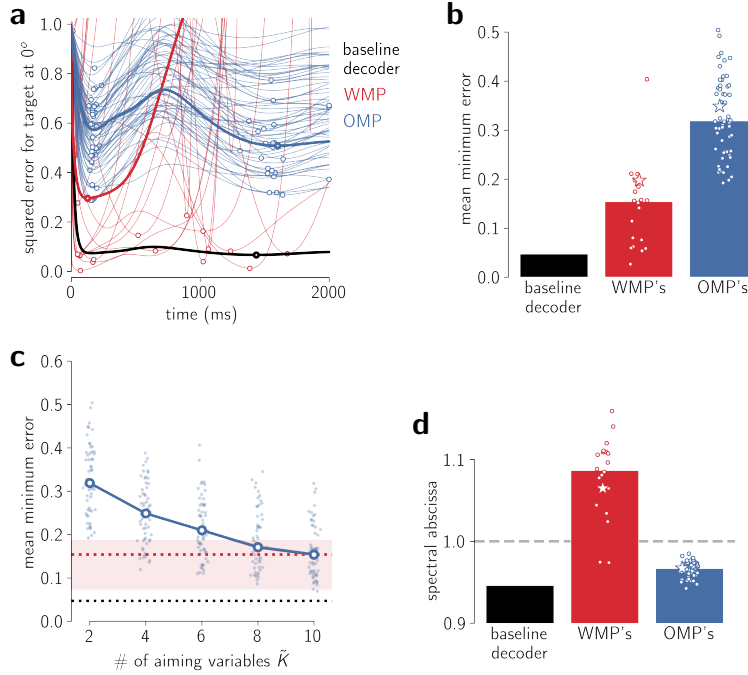
Figure 15: Simulation of the myopic closed-loop re-aiming strategy.
(15a) Squared error achieved by myopic closed-loop re-aiming under each decoder, for a single target readout. The black line corresponds to the baseline decoder, whereas the red and blue lines correspond to each sampled WMP and OMP, respectively. These are exactly the same decoders used in the simulations presented in figure 13. The thicker red and blue lines correspond to the example WMP and OMP analyzed in that figure. The open circles overlaid on each trajectory show the minimum squared error acheived within this time window.
(15b) Minimum squared error achieved over this time window, averaged over all eight radial reach targets. Each point corresponds to a different decoder, with medians over all decoders in each class marked by the height of the bars. The stars correspond to the example decoders corresponding to the highlighted trajectories in the previous panel.
(15c) Mean minimum error under myopic re-aiming for each OMP, with 2, 4, 6, 8, and 10 aiming variables. Light blue points denote this quantity for individual OMP's, larger open circles on top show the median. For reference, dotted horizontal lines show the mean minimum error achieved by myopic re-aiming with $\widetilde{K} = 2$ aiming variables for the baseline decoder (black) and WMP's (red); the red dotted line shows the median over all sampled WMP's with shading marking the upper and lower quartiles.
(15d) Spectral abscissa – that is, the maximum of the real parts of its eigenvalues – of the closed-loop dynamics matrix, $\mathbf{W}^{\mathrm{rec}} - \mathbf{W}^{\mathrm{in}}\widetilde{\mathbf{U}}\mathbf{G}\mathbf{D}$, for each decoder $\mathbf{D}$ (note that $\mathbf{G}$ implicitly depends on $\mathbf{D}$). As in the previous panel, bars mark medians and the stars mark the example decoders highlighted in previous panels. Values above the dashed horizontal line at 1.0 imply that the closed-loop dynamics under myopic re-aiming will be unstable.

mean minimum error is plotted in figure 15b for each decoder. In general, the mean minimum error is substantially lower for WMP's than for OMP's, revealing that the myopic closed-loop re-aiming strategy generally succeeds in producing readouts near the target under WMP's and the baseline decoder, but not under OMP's.

In chapter 1, we found that open-loop re-aiming could succeed for OMP's if we increased the number of motor variables used for re-aiming, $\widetilde{K}$. Might such a generalized re-aiming strategy also succeed in this closed-loop setting? The answer is yes: figure 15c shows that, as we increase $\widetilde{K}$, myopic re-aiming can achieve low mean minimum error even under OMP's. For $\widetilde{K} = 10$ this quantity reaches the same level that can be generally achieved under WMP's using $\widetilde{K} = 2$ aiming variables.

It is important to note, however, that this metric only tells us whether the readout reaches the target *at a single point in time*; it tells us nothing about whether the readout can then be kept near the target for longer than that instant. Figure 15a suggests that this is actually impossible under WMP's, as the

error typically remains low only transiently under these decoders. This phenomenon can be understood by analyzing the closed-loop dynamics under myopic closed-loop re-aiming. When the aiming variables are set according to equation 62, the network dynamics reduce to the following (obtained by plugging in equation 62 into equations 41 and 42)

$$\tau\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = -\mathbf{r} + (\mathbf{W}^{\mathrm{rec}} - \mathbf{W}^{\mathrm{in}}\widetilde{\mathbf{U}}\mathbf{G}\mathbf{D})\mathbf{r} + \mathbf{W}^{\mathrm{in}}\widetilde{\mathbf{U}}\mathbf{G}(\mathbf{y}^* + \mathbf{D}\boldsymbol{\mu}).$$

Because these dynamics are linear, the stability of this system is completely determined by the eigenvalues of the matrix $\mathbf{W}^{\mathrm{rec}} - \mathbf{W}^{\mathrm{in}}\widetilde{\mathbf{U}}\mathbf{G}\mathbf{D}$. If any of the eigenvalues of this matrix have a real part larger than 1, then the system will be unstable, and it won't be possible to maintain a fixed readout for an extended period of time. In figure 15d we plot the maximum of the real parts of the eigenvalues of this matrix (i.e. its spectral abscissa), for each decoder $\mathbf{D}$. The fact that this quantity is greater than 1.0 for most WMP's reveals that the closed-loop dynamics under myopic re-aiming are typically unstable for these decoders. This explains why the error does not remain low in figure 15a for WMP's.

The reason why this happens can be understood in terms of our derivation of this closed-loop control strategy, in which took the limit of $t_{\mathrm{end}} \to 0$ to avoid having to "look ahead". In this limit, the controller's objective is narrowed down to getting the readout closer to the target only at this current instant in time, putting aside any concerns about where the readout will be in the future. Indeed, it is straightforward to verify that the myopic closed-loop controller of equation 62 can be equivalently derived from the objective

$$\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*) = \arg\min_{\widetilde{\boldsymbol{\theta}}} \ \frac{\mathrm{d}}{\mathrm{d}t}\big\|\mathbf{y}\big(t_{\mathrm{end}}; \widetilde{\boldsymbol{\theta}}\big) - \mathbf{y}^*\big\|^2 + \frac{\gamma}{M}\sum_{i=1}^{M} u_i(\widetilde{\boldsymbol{\theta}})^2. \tag{63}$$

This strategy risks putting the network in a state that, while greedily reducing the error right now, might make it more difficult to keep the error low later. It is this short-sightedness that leads to unstable dynamics under certain decoders, prompting us to call this control strategy "myopic".

The behavioral implication of this result is that, under certain decoders, the myopic re-aiming strategy can't maintain a given target readout for a long period of time. This is reminiscent of the observation that non-human primates are typically unable to hold a BMI cursor still (see [Golub et al., 2014] for a brief review of such observations). It is possible that, on short timescales, they learn by myopic re-aiming, and thus have a hard time producing constant BMI readouts. Over time, however, the sensorimotor transform $\mathbf{G}$ might be further optimized to stabilize control, both by directly modifying the parameters of this transformation and by increasing the number of motor variables used for re-aiming.

III.1.2 Optimized closed-loop re-aiming

We next ask whether these instabilities can be avoided by simply modifying the weights of the error feedback loop, $\mathbf{G}$. In the myopic closed-loop re-aiming model, these were derived from taking the limit

of $t_{\text{end}} \to 0$. This leads to instabilities because it implicitly optimizes a greedy short-sighted objective (equation 63). Moreover, under non-linear dynamics, this approach leads to a more complex architecture that requires learning a high-dimensional $\mathcal{O}N$ mapping from errors to population activity[7], defeating the point of using this strategy to reduce the dimensionality of learning. So here we take inspiration from the simple error feedback control architecture derived in the previous section to construct closed-loop controllers that (1) are stable and (2) work under non-linear dynamics.

We begin by simply assuming the error feedback architecture of equatiion 62 with the minor addition of a constant drive term $\mathbf{g} \in \mathbb{R}^{\widetilde{K}}$,

$$\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*) = \mathbf{G}(\mathbf{y}^* - \mathbf{y}(t)) + \mathbf{g}. \tag{64}$$

For simplicity, we assume linear encoding in inputs,

$$\mathbf{u}(\widetilde{\boldsymbol{\theta}}) = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}.$$

Plugging these two equations (along with the readout equation 3) into the RNN dynamics (equation 1), we arrive at the following closed-loop dynamics,

$$\tau \frac{d\mathbf{x}}{dt} = -\mathbf{x} + (\mathbf{W}^{\text{rec}} - \mathbf{W}^{\text{in}}\widetilde{\mathbf{U}}\mathbf{G}\mathbf{D})\mathbf{r} + \mathbf{W}^{\text{in}}\widetilde{\mathbf{U}}\left(\mathbf{G}(\mathbf{y}^* + \mathbf{D}\boldsymbol{\mu}) + \mathbf{g}\right).$$

Rather than prescribing the feedback weights $\mathbf{G}, \mathbf{g}$ ourselves as we did in chapter 1, we will then numerically optimize them by minimizing the objective function

$$\hat{\mathbf{G}}, \hat{\mathbf{g}} = \underset{\mathbf{G},\mathbf{g}}{\arg\min} \; \frac{1}{T}\int_0^T \left\|\mathbf{y}\big(t; \mathbf{G}, \mathbf{g}\big) - \mathbf{y}^*\right\|^2 dt + \gamma\left(\|\mathbf{g}\|^2 + \|\mathbf{G}\|_{\mathcal{F}}^2\right), \tag{65}$$

---

[7]Plugging in equation 1 into equation 63, with $\phi(\cdot)$ set to the rectified linear activation function used in chapter 1, we get the following equation for $\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*)$,

$$\mathbf{0} = \frac{1}{\tau}\widetilde{\mathbf{U}}^T \boldsymbol{\Phi}'\left(\widetilde{\mathbf{U}}\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*)\right)\mathbf{W}^{\text{in}\,T}\boldsymbol{\Phi}'\left(\mathbf{x}(t)\right)\mathbf{D}^T\left(\mathbf{y}(t) - \mathbf{y}^*\right) + \frac{\gamma}{M}\widetilde{\mathbf{U}}^T\boldsymbol{\Phi}'\left(\widetilde{\mathbf{U}}\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*)\right)\phi\left(\widetilde{\mathbf{U}}\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*)\right),$$

where $\boldsymbol{\Phi}'(\mathbf{x})$ is a diagonal matrix with the derivative of the activation function $\phi(\cdot)$ evaluated at each coordinate of $\mathbf{x}$. For a rectified linear activation function, this matrix has only 1's and 0's on its diagonal, effectively acting as a multi-dimensional gate. This equation is non-linear in $\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*)$ and is unclear how to solve it. If we instead assume the upstream inputs to be linear in the motor variables, $\mathbf{u}(\widetilde{\boldsymbol{\theta}}) = \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\theta}}$, this simplifies substantially (namely, $\boldsymbol{\Phi}'\left(\widetilde{\mathbf{U}}\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*)\right) = \mathbf{I}$) and the following analytical solution can be derived,

$$\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*) = \mathbf{G}(t)(\mathbf{y}^* - \mathbf{y}(t))$$
$$\mathbf{G}(t) = \frac{1}{\gamma\tau}\left(\frac{1}{M}\widetilde{\mathbf{U}}^T\widetilde{\mathbf{U}}\right)^{-1}\widetilde{\mathbf{U}}^T\mathbf{W}^{\text{in}\,T}\boldsymbol{\Phi}'\left(\mathbf{x}(t)\right)\mathbf{D}^T.$$

Note that, unlike the case of linear dynamics (equation 62), in this case the optimal feedback weights $\mathbf{G}(t)$ change in time. Importantly, they change in time as a function of the full network state $\mathbf{x}(t)$. From a circuit perspective, this creates the difficulty of looping back this state to gate the pathway that maps sensory inputs $(\mathbf{y}^* - \mathbf{y}(t))$ to internal motor variables $(\dot{\widetilde{\boldsymbol{\theta}}}(t; \mathbf{y}^*))$. From a learning perspective, note that this gating operates on an N-dimensional intermediate sensory representation, $\mathbf{D}^T(\mathbf{y}^* - \mathbf{y}(t))$. Learning the low-dimensional $\widetilde{K} \times 2$ optimal feedback loop, $\mathbf{G}(t)$, thus requires learning the high-dimensional $N \times 2$ sensory mapping $\mathbf{D}^T$.
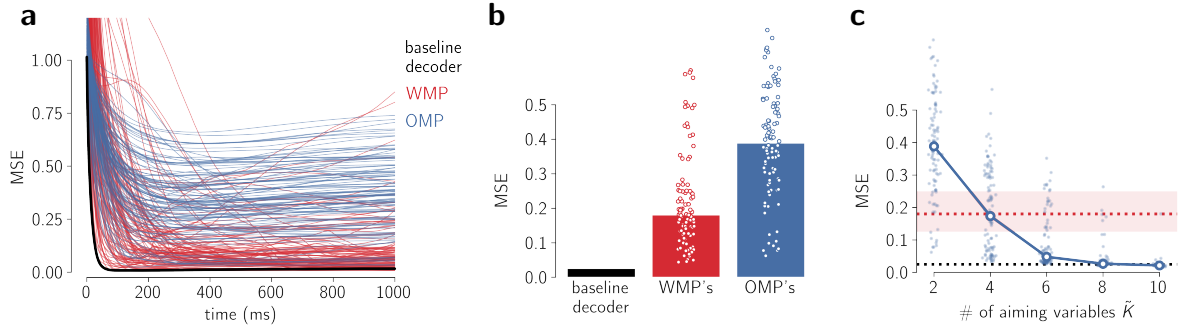
Figure 16: Simulation of optimized error feedback control under non-linear dynamics.
(16a) Mean squared error (mean over all eight radial target readouts) achieved by error feedback controllers with $\widetilde{K} = 2$ aiming variables optimized for each decoder, plotted over a 1000ms time window. Each line corresponds to a different set of feedback weights, optimized for a given decoder (equation 65).
(16b) Mean squared error (mean over target readouts and over time) achieved by error feedback controllers with $\widetilde{K} = 2$ aiming variables optimized for each decoder. As in previous figures, each point corresponds to a different decoder, with medians over all decoders in each class marked by the height of the bars.
(16c) Mean squared error (mean over target readouts and over time) achieved by error feedback controllers optimized for each OMP, with $\widetilde{K} = 2, 4, 6, 8$ and 10 aiming variables. Light blue points denote this quantity for individual OMP's, larger open circles on top show the median. For reference, dotted horizontal lines show the mean squared error achieved by optimized error feedback with $\widetilde{K} = 2$ aiming variables for the baseline decoder (black) and WMP's (red); the red dotted line shows the median over all sampled WMP's with shading marking the upper and lower quartiles.

where $\|\mathbf{G}\|_{\mathcal{F}}^2 = \sum_{i=1}^{\widetilde{K}} \sum_{j=1}^{2} G_{ij}^2$ is the squared Frobenius norm, and $\mathbf{y}(t; \mathbf{G}, \mathbf{g})$ denotes the readout produced at time t under the closed-loop dynamics.

We performed this optimization for the non-linear network analyzed in chapter 1, with a rectified linear activation function $r_i = \phi(x_i)$ (equation 1) and deterministic dynamics (no noise in the dynamics, initial conditions fixed to 0). We used the Adam optimization algorithm with an initial learning rate of .01 [Kingma and Ba, 2017]. A separate optimization was done for each of the baseline/WMP/OMP decoders fit to this network (the same decoders used in chapter 1, cf. section I.3.8). Following an analagous procedure to that used in chapter 1 (cf. section I.3.3), for the baseline decoder we performed this optimization over multiple values of the metabolic cost weight γ so as to identify the largest value of γ that permitted a time-averaged squared error of less than .05 for all eight target readouts under this decoder. We then used this value of γ to optimize the feedback weights for each decoder perturbation.

Results from simulating center-out reaching with each of the optimal feedback weights for $\widetilde{K} = 2$ aiming variables are shown in figure 16a, where we plot the mean squared error over all eight target readouts over time. We find that for almost all decoders, the mean squared error decreases to a certain level and remains low for the rest of this time window of 1000ms. Notably, this is the case for most WMP's, for which the myopic re-aiming model struggled to keep the error low. Because we optimized the error feedback loop to minimize the error over the full 1000ms window, it is able to effectively take into account the future dynamics of the network to maintain a low error level, overcoming the instabilities inherent to the myopic re-aiming model.

Importantly, the mean squared error to which these decoder-specific closed-loop controllers converge to is typically higher for OMP's than for WMP's (fig. 16c), extending our results on open-loop control

from chapter 1 (fig. 3g) to the closed-loop setting. This again reflects the limitations of re-aiming with $\widetilde{K} = 2$ motor variables. In this case, this manifests itself in restricting how the error can be fed back into the network: the error gets mapped to a $\widetilde{K}$-dimensional vector through equation 64 before being fed back into the network. As we saw occurs for the open-loop controller, this results in a restriction of how population activity can be modulated, making it difficult to generate the patterns of activity required to produce the target readouts under OMP's. Figure 16c show that these restrictions can again be relaxed by increasing the number of motor variables used for re-aiming, $\widetilde{K}$. In this case, re-aiming with only $\widetilde{K} = 6$ motor variables suffices to obtain a mean squared error less than 0.1 with OMP's (as opposed to $\widetilde{K} > 10$ under open-loop re-aiming, cf. fig. 6a), reflecting the benefits of closed-loop control.

It is worth noting, however, that for $\widetilde{K} = 2$, performance is highly variable across different decoder perturbations of the same class. In particular, it's substantially more variable than was the case for the open-loop controller analyzed in chapter 1, particularly for WMP's (compare fig. 16b to fig. 3g). It is unclear if this is actually a property of the error feedback architecture used here or, rather, a property of our optimization procedure, which we indeed observed often converged to bad local optima especially for WMP's (a problem we tried to avoid by repeating the optimization from five different random initializations and keeping only the feedback weights from the best of these five runs). For what it's worth, substantial variability was also observed in the results of Sadtler et al. (2014).

Discussion

In this chapter we constructed and analyzed closed-loop variants of the re-aiming strategy. Concretely, we analyzed two variants based on an error feedback architecture, whereby the displacement between the current readout $\mathbf{y}(t)$ and the target readout $\mathbf{y}^*$ is fed back to the network at each timestep. Critically, it is fed back by first being mapped to a set of $\widetilde{K}$ motor variables. This means that all that needs to be learned to implement this closed-loop strategy is a low-dimensional mapping from the two-dimensional workspace in which the readouts are observed to the $\widetilde{K}$-dimensional space of motor variables used for re-aiming. The low dimensionality of this mapping facilitates efficient learning but restricts the flexibility of this control strategy, manifested through differences in its effectiveness at solving a center-out reaching task with within- vs. outside- manifold decoder perturbations. This extends our results from chapter 1, where we considered only models of open-loop control, to the context of closed-loop control.

While we found that the myopic closed-loop re-aiming model explored in the first part suffered from problematic instabilities, we noted an intriguing connection to the observation that non-human primates typically struggle to keep the BMI cursor still in traditional 2D cursor reaching BMI tasks [Golub et al., 2014]. The myopic closed-loop control strategy was derived by taking the limit of $t_{\mathrm{end}} \to 0$, raising the intriguing possibility that non-human primates may take a similarly myopic approach to BMI

control early in learning. This connection certainly warrants further exploration.

Omitted from this chapter is an exploration into the biases that might exist under the closed-loop re-aiming strategies analyzed here. Given our insights in chapter 2 into the structure of the reachable repertoire under linear dynamics, there is little reason to expect the linear myopic re-aiming model to demonstrate unimodal biases. But it is certainly possible that such biases might arise under the non-linear dynamics endowed by the rectified linear activation function considered in the second part of this chapter. This remains an important line of future research on closed-loop re-aiming.

Another important line of future research is to extend these models to address some glaring oversimplifications implicit in their construction. The most obvious one is that sensory feedback is instantaneously available to the motor system. In both of the models considered in this chapter, we assumed that the aiming variables at time t, $\hat{\bar{\boldsymbol{\theta}}}(t; \mathbf{y}^*)$, had access to the readout at that time, $\mathbf{y}(t)$. In reality, this is impossible, as sensory information takes time to be communicated from the sensory periphery to the central motor system. This is particularly true for visual information – which is the sensory information relevant to the BMI tasks considered here –, which is known to require at least about 80-100ms to influence motor cortical activity [Golub et al., 2015, Stavisky et al., 2017].

These latencies are thought to be addressed by the motor system through internal predictive models of how its outputs influence the environment (often referred to as "forward models") [Wolpert et al., 1995, Shadmehr et al., 2010]. Such internal models are thought to be learned and exploited by the motor system to estimate the current state of the environment (e.g. the current BMI readout or cursor position) based on past observations of the environment together with afferent copies of its motor outputs [Shadmehr and Krakauer, 2008]. Indeed, Golub et al. (2015) found that non-human primates' motor cortical activity during a BMI task was highly consistent with predictions from a simple predictive model of the BMI effector's dynamics [Golub et al., 2015]. These authors found that, despite clear sensory delays evident in the subjects' reaction times to target presentation, their motor cortical activity was able to compensate for these delays during the reach by producing activity patterns that would move the cursor in the right direction under an imperfect yet adequate predictive model of the motor cortical activity's influence on the BMI cursor position. Such an internal predictive model is entirely missing from the theoretical models of closed-loop control presented here. Closed-loop re-aiming should be seen as a theory of how the predictions of such an internal predictive model (i.e. of the current BMI readout, $\mathbf{y}(t)$) would be used during BMI control.

That said, the parcellation of the BMI learning problem into learning a control strategy and learning an internal predictive model is completely artificial – in reality, both the control strategy as well as the internal predictive model are learned simultaneously. An important and surely fruitful direction of future research will be to investigate how internal sensory prediction models might also be learned in a low-dimensional manner to enable efficient learning. Understanding how such constraints might influence

motor cortical activity and behavior during BMI control could lead to complementary insights on the limits of BMI learning and the related phenomena discussed and studied in chapter 1.

One important improvement (relative to our open-loop models) implicit in our approach to modeling re-aiming in a closed-loop setting is that the re-aiming objective takes into account the whole timecourse of readouts, rather than just the readout at a single endpoint time (compare equation 4 to equation 65). This is important when considering situations in which the BMI readout does not directly specify the BMI effector's movement. In fact, this is the case in the experiments of Sadtler et al. (2014), where the BMI readout in fact specifies the velocity of the BMI cursor rather than its position. In this case, just optimizing the readout at a single future time does not suffice to move the cursor to the right place – you need to ensure that all the readouts produced at each point in time integrate to the desired target cursor position. The true re-aiming objective should thus take into account all readouts produced at each point in time. It should also permit the target readouts to change in time during the course of a movement or trial, as fluctuations in the cursor position would require changes in the cursor velocities needed to reach to the same target. While this is an important limitation of all the models presented here, it is worth noting that there is little reason not to expect that most of the intuitions provided by the simple models explored in this study remain true when considering time-varying target readouts and alternative BMI effector dynamics.

General Discussion

"Since all models are wrong the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad."

George Box, Science and Statistics (1976)

In this thesis, we have outlined a theory of brain-machine interface learning based on what we have called the re-aiming learning strategy. We demonstrated in chapter 1 that a simplified open-loop realization of this theory can provide a unifying explanation for disparate phenomena across three different BMI learning tasks: (i) two-dimensional cursor reaching under within- and outside- manifold decoder perturbations [Sadtler et al., 2014, Oby et al., 2019], (ii) two-dimensional cursor reaching under credit assignment rotation perturbations [Jarosiewicz et al., 2008, Chase et al., 2012, Zhou et al., 2019], and (iii) operant conditioning of individual neurons [Fetz, 1969, Fetz and Baker, 1973, Koralek et al., 2012, Clancy et al., 2014, Athalye et al., 2018]. We moreover argued that this explanation is a normative one, as learning by re-aiming is a potentially data-efficient learning strategy whereby the learning process is constrained to a low-dimensional optimization space; this stands in stark contrast to learning strategies based solely on synaptic plasticity. In chapter 2 we further grounded the arguments and intuitions derived from our simulations in chapter 1 through mathematical analysis of a simple case (namely, linear dynamics). This revealed deep connections between our analysis of the re-aiming learning strategy and standard concepts from classical linear control theory. Finally, in chapter 3 we extended a subset of these results to a more complex but biologically realistic setting, in which re-aiming happens continuously in real time via closed-loop interaction with the environment.

We highlight here what we believe to be the main contributions of this work:

A novel perspective on the low dimensionality of population activity in motor cortex and its implications for BMI control. Our model of re-aiming can be seen as a mechanistic realization of the statistical theory of Gao and Ganguli (2017), whereby the dimensionality of population activity is limited by behavior [Gao et al., 2017]. In our theory, these limits are imposed by the dimensionality of the motor variables represented in the upstream populations driving motor cortex.

Theoretical tools for analyzing the structure of low-dimensional population activity during motor control. In particular, we introduced the concept of the reachable repertoire. As discussed at

length in chapters 1 and 2, this theoretical construct bears a complementary relationship to the similar theoretical constructs in the literature, such as the so-called intrinsic manifold [Sadtler et al., 2014, Wärnberg and Kumar, 2019, Feulner and Clopath, 2021] and the controllability Gramian [Kao and Hennequin, 2019].

Analytical and numerical tools for analyzing the properties of recurrent neural networks with a rectified linear activation function. Of particular note is our analysis of the reachable repertoire's geometry in section I.4.4 and our algorithm for efficiently computing re-aiming solutions (section I.3.2).

The demonstration that classical observations about changes to individual neurons can be attributed to global changes in the motor system. We show that apparent so-called credit assignment effects [Jarosiewicz et al., 2008, Chase et al., 2012, Zhou et al., 2019] can arise naturally without solving the credit assignment problem [Legenstein et al., 2010]. Our theory also reveals that the ability of primates [Fetz, 1969, Fetz and Baker, 1973, Engelhard et al., 2019] and mice [Koralek et al., 2012, Clancy et al., 2014, Athalye et al., 2018] to modulate the firing rates of specific neurons does not necessarily indicate their motor systems are capable of identifying and directly manipulating the properties of individual neurons in isolation.

The construction of circuit models of closed-loop BMI control. While we only scratch the surface in chapter 3, these provide the groundwork for future elaborations, such as including a predictive model for state estimation.

In striving for transparent explanation and intuitive understanding, we focused on developing simple models of re-aiming. It is this simplicity that lead us to obtain strong intuitions in chapter 1 that carried over to more complex settings like time-varying inputs (section II.2.2) and closed-loop control (sections III.1.1, III.1.2). However, a number of important limitations remain that should be addressed in future work. Three particular future lines of work along these lines that are worth mentioning are:

Extending these results to take into account the BMI effector dynamics. For example, in the experiments of Sadtler et al. (2014), the BMI readout determines the velocity of the cursor rather than its position. That means that simply producing a target readout $\mathbf{y}^*$ matching a given reach target isn't enough to solve the task – a whole timecourse of readouts needs to be produced that will integrate over time to move the cursor to the reach target. It is straight-forward to extend the theoretical results from section II.2.1 to this setting, by simply replacing the reachable repertoire matrix, $\mathbf{M}(t_{\mathrm{end}})$, with its integral over time, $\int_0^{t_{\mathrm{end}}} \mathbf{M}(t)\mathrm{d}t$. Accordingly adapting the remaining simulations and derivations for the non-linear network has not been attempted. However, there is little reason not to expect the results and intuitions elaborated in this thesis to extend to this mathematically less tractable setting.

Analyzing the case of non-silent initial conditions. It is straight-forward to incorporate the case of non-zero initial conditions under linear dynamics – this boils down to simply adding an extra term (cf. equation 44) that requires extra algebra to deal with. Under non-linear dynamics, on the other hand, non-zero initial conditions can change things in non-trivial ways. In particular, the scale-invariance property of the dynamics under a rectified linear activation function (equation 10) no longer holds when the initial conditions are non-zero. It will be interesting to investigate how non-zero initial conditions could potentially alter the goemetry of the reachable repertoire, although again there is little reason to expect this would fundamentally alter our conclusions regarding the limitations of the re-aiming strategy.

Incorporating delayed sensory feedback into closed-loop models of BMI control. A future line of research inspired by the closed-loop models considered in chapter 3 is to consider how predictive models for state estimation could be used to account for latencies in sensory feedback for error feedback control. Three fascinating questions along these lines are

▷ How can such predictive models be parametrized in low-dimensions so that learning can be efficient?

▷ How would concurrent state estimation interact with low-dimensional error feedback control?

▷ How might both of these components be learned simultaneously? And how would these two learning processes interact with each other?

Finally, like any worthwhile piece of research should, this thesis raises more questions than it answers. We highlight here what we believe to be the most promising aveneus of future research going beyond the scope of the theory and empirical phenomena discussed in this thesis:

How are re-aiming solutions learned without synaptic plasticity? A significant shortcoming of the theory presented here is that it only concerns hypotheses about *what* subjects learn during BMI learning, rather than *how* they learn it. Assuming subjects do engage in re-aiming, a crucial avenue of future research will be to understand the algorithms and circuitry underlying how re-aiming solutions are computed by the motor system. Critically, for the advantages of re-aiming to be maintained, these must operate within the low-dimensional space of the motor variables. As discussed at the end of section I.2.2, one intriguing approach is for the motor variables to be stored and updated in the internal state of an upstream recurrent circuit, which, through its dynamics, implicitly implements a gradient-free optimization scheme [Wang et al., 2018]. Testing this simple idea, and to what extent it may be a good model of BMI learning and/or classical motor adaptation, is a natural next step from this work.

How do re-aiming solutions differ from and interact with synaptic plasticity solutions? Our analysis of the re-aiming learning strategy does not rule out the possibility that BMI learning might instead proceed solely on the basis of changes in neural connectivity within or outside of motor cortex.

In chapter 1, we argued that the pitfalls of high-dimensional optimization would ensure that such a learning strategy would be slow and brittle, unlike what we see in (at least short-term) BMI learning. That said, we did not prove or concretely demonstrate that this is true, and it is worth noting that a few previous theoretical studies have argued that this intuition might be wrong [Raman et al., 2019]. A useful avenue of future research along these lines will be to develop a theory of how synaptic plasticity within motor cortex during BMI learning might reshape the reachable repertoire [Wärnberg and Kumar, 2019, Feulner and Clopath, 2021], and how such signatures might enable us to disambiguate between re-aiming-based learning strategies and plasticity-based learning strategies. A second related line of research is to build models in which synaptic plasticity within motor cortex occurs concurrently with re-aiming, to understand how these two learning processes might interact.

Low-dimensional learning strategies in more complex tasks. Finally, we note here that the idea of optimizing a low-dimensional set of internal variables to drive a downstream network could be applied to learning any other task outside of the context of BMI control. In preliminary work [Zhou et al., 2021], we recently applied this idea successfully to the ready-set-go task of Jazayeri & Shadlen (2010) [Jazayeri and Shadlen, 2010]. By learning an appropriate low-dimensional representation of the sensory cues in the task, we were able to drive a chaotic reservoir [Jaeger, 2001, Sussillo and Abbott, 2009] to produce the correct responses in this task. The resulting solution moreoever produced rational response biases resembling primates' behavior and population dynamics in the reservoir resembling neural activity recorded in primate frontal cortex [Sohn et al., 2019]. This raises the possibility that such low-dimensional learning strategies might be employed by subjects across a wide variety of neuroscience tasks, and during natural learning in the real world. Such a modular approach to learning might support successful transfer across tasks, whereby the same downstream motor network is shared across tasks and the "motor variables" stored upstream change. This provides a novel perspective on the teleology of central motor circuits directly driving movement (such as motor cortex) [Lopes et al., 2017] and the macrostructure of mammalian motor systems [Dum and Strick, 2002, Shadmehr and Krakauer, 2008].

References

[Aflalo et al., 2015] Aflalo, T., Kellis, S., Klaes, C., Lee, B., Shi, Y., Pejsa, K., Shanfield, K., Hayes-Jackson, S., Aisen, M., Heck, C., Liu, C., and Andersen, R. A. (2015). Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. *Science*, 348(6237):906–910.

[Andersen et al., 2019] Andersen, R. A., Aflalo, T., and Kellis, S. (2019). From thought to action: The brain–machine interface in posterior parietal cortex. *Proceedings of the National Academy of Sciences*, 116(52):26274–26279.

[Andersen and Buneo, 2002] Andersen, R. A. and Buneo, C. A. (2002). Intentional Maps in Posterior Parietal Cortex. *Annual Review of Neuroscience*, 25(1):189–220.

[Athalye et al., 2020] Athalye, V. R., Carmena, J. M., and Costa, R. M. (2020). Neural reinforcement: Re-entering and refining neural dynamics leading to desirable outcomes. *Current Opinion in Neurobiology*, 60:145–154.

[Athalye et al., 2017] Athalye, V. R., Ganguly, K., Costa, R. M., and Carmena, J. M. (2017). Emergence of Coordinated Neural Dynamics Underlies Neuroprosthetic Learning and Skillful Control. *Neuron*, 93(4):955–970.e5.

[Athalye et al., 2018] Athalye, V. R., Santos, F. J., Carmena, J. M., and Costa, R. M. (2018). Evidence for a neural law of effect. *Science*, 359(6379):1024–1029.

[Bartunov et al., 2018] Bartunov, S., Santoro, A., Richards, B., Marris, L., Hinton, G. E., and Lillicrap, T. (2018). Assessing the Scalability of Biologically-Motivated Deep Learning Algorithms and Architectures. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems 31*, pages 9368–9378. Curran Associates, Inc.

[Batista et al., 1999] Batista, A. P., Buneo, C. A., Snyder, L. H., and Andersen, R. A. (1999). Reach Plans in Eye-Centered Coordinates. *Science*, 285(5425):257–260.

[Björck and Golub, 1973] Björck, À. and Golub, G. H. (1973). Numerical methods for computing angles between linear subspaces. *Mathematics of Computation*, 27(123):579–594.

[Box, 1976] Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356):791–799.

[Brownstone et al., 2015] Brownstone, R. M., Bui, T. V., and Stifani, N. (2015). Spinal circuits for motor learning. *Current Opinion in Neurobiology*, 33:166–173.

[Buneo et al., 2002] Buneo, C. A., Jarvis, M. R., Batista, A. P., and Andersen, R. A. (2002). Direct visuomotor transformations for reaching. *Nature*, 416(6881):632–636.

[Byrd et al., 1995] Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208.

[Carmena et al., 2003] Carmena, J. M., Lebedev, M. A., Crist, R. E., O'Doherty, J. E., Santucci, D. M., Dimitrov, D. F., Patil, P. G., Henriquez, C. S., and Nicolelis, M. A. L. (2003). Learning to Control a Brain–Machine Interface for Reaching and Grasping by Primates. *PLOS Biology*, 1(2):e42.

[Chase et al., 2012] Chase, S. M., Kass, R. E., and Schwartz, A. B. (2012). Behavioral and neural correlates of visuomotor adaptation observed through a brain-computer interface in primary motor cortex. *Journal of Neurophysiology*, 108(2):624–644.

[Chase et al., 2010] Chase, S. M., Schwartz, A. B., and Kass, R. E. (2010). Latent Inputs Improve Estimates of Neural Encoding in Motor Cortex. *Journal of Neuroscience*, 30(41):13873–13882.

[Chen et al., 2018] Chen, R. T. Q., Rubanova, Y., Bettencourt, J., and Duvenaud, D. K. (2018). Neural Ordinary Differential Equations. In *Advances in Neural Information Processing Systems 31*, pages 6571–6583. Curran Associates, Inc.

[Chen et al., 2015] Chen, S. X., Kim, A. N., Peters, A. J., and Komiyama, T. (2015). Subtype-specific plasticity of inhibitory circuits in motor cortex during motor learning. *Nature Neuroscience*, 18(8):1109–1115.

[Churchland et al., 2012] Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyu-jukian, P., Ryu, S. I., and Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature*, 487(7405):51–56.

[Churchland et al., 2010] Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I., and Shenoy, K. V. (2010). Cortical Preparatory Activity: Representation of Movement or First Cog in a Dynamical Machine? *Neuron*, 68(3):387–400.

[Churchland and Shenoy, 2007] Churchland, M. M. and Shenoy, K. V. (2007). Temporal Complexity and Heterogeneity of Single-Neuron Activity in Premotor and Motor Cortex. *Journal of Neurophysiology*, 97(6):4235–4257.

[Cichon and Gan, 2015] Cichon, J. and Gan, W.-B. (2015). Branch-specific dendritic Ca 2+ spikes cause persistent synaptic plasticity. *Nature*, 520(7546):180–185.

[Cisek and Kalaska, 2004] Cisek, P. and Kalaska, J. F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature*, 431(7011):993–996.

[Cisek and Kalaska, 2005] Cisek, P. and Kalaska, J. F. (2005). Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. *Neuron*, 45(5):801–814.

[Cisek and Kalaska, 2010] Cisek, P. and Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience*, 33(1):269–298.

[Clancy et al., 2014] Clancy, K. B., Koralek, A. C., Costa, R. M., Feldman, D. E., and Carmena, J. M. (2014). Volitional modulation of optically recorded calcium signals during neuroprosthetic learning. *Nature Neuroscience*, 17(6):807–809.

[Costa, 2011] Costa, R. M. (2011). A selectionist account of de novo action learning. *Current Opinion in Neurobiology*, 21(4):579–586.

[Crick, 1989] Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337(6203):129–132.

[d'Avella and Bizzi, 1998] d'Avella, A. and Bizzi, E. (1998). Low dimensionality of supraspinally induced force fields. *Proceedings of the National Academy of Sciences*, 95(13):7711–7714.

[d'Avella et al., 2003] d'Avella, A., Saltiel, P., and Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nature Neuroscience*, 6(3):300–308.

[Dayan and Abbott, 2001a] Dayan, P. and Abbott, L. F. (2001a). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*.

[Dayan and Abbott, 2001b] Dayan, P. and Abbott, L. F. (2001b). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*.

[Dum and Strick, 2002] Dum, R. P. and Strick, P. L. (2002). Motor areas in the frontal lobe of the primate. *Physiology & Behavior*, 77(4):677–682.

[Dushanova and Donoghue, 2010] Dushanova, J. and Donoghue, J. (2010). Neurons in primary motor cortex engaged during action observation. *European Journal of Neuroscience*, 31(2):386–398.

[Engelhard et al., 2019] Engelhard, B., Darshan, R., Ozeri-Engelhard, N., Israel, Z., Werner-Reiss, U., Hansel, D., Bergman, H., and Vaadia, E. (2019). Neuronal activity and learning in local cortical networks are modulated by the action-perception state. *bioRxiv*, page 537613.

[Fadiga et al., 1995] Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, 73(6):2608–2611.

[Fetz, 1969] Fetz, E. E. (1969). Operant Conditioning of Cortical Unit Activity. *Science*, 163(3870):955–958.

[Fetz, 1992] Fetz, E. E. (1992). Are movement parameters recognizably coded in the activity of single neurons? *Behavioral and Brain Sciences*, 15(4):679–690.

[Fetz, 2007] Fetz, E. E. (2007). Volitional control of neural activity: Implications for brain–computer interfaces. *The Journal of Physiology*, 579(3):571–579.

[Fetz and Baker, 1973] Fetz, E. E. and Baker, M. A. (1973). Operantly conditioned patterns on precentral unit activity and correlated responses in adjacent cells and contralateral muscles. *Journal of Neurophysiology*, 36(2):179–204.

[Feulner and Clopath, 2020] Feulner, B. and Clopath, C. (2020). Neural manifold under plasticity in a goal driven learning behaviour. *bioRxiv*, page 2020.02.21.959163.

[Feulner and Clopath, 2021] Feulner, B. and Clopath, C. (2021). Neural manifold under plasticity in a goal driven learning behaviour. *PLOS Computational Biology*, 17(2):e1008621.

[Gallego et al., 2017] Gallego, J. A., Perich, M. G., Miller, L. E., and Solla, S. A. (2017). Neural Manifolds for the Control of Movement. *Neuron*, 94(5):978–984.

[Gallego et al., 2018] Gallego, J. A., Perich, M. G., Naufel, S. N., Ethier, C., Solla, S. A., and Miller, L. E. (2018). Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature Communications*, 9(1):1–13.

[Ganguli et al., 2008a] Ganguli, S., Bisley, J. W., Roitman, J. D., Shadlen, M. N., Goldberg, M. E., and Miller, K. D. (2008a). One-Dimensional Dynamics of Attention and Decision Making in LIP. *Neuron*, 58(1):15–25.

[Ganguli et al., 2008b] Ganguli, S., Huh, D., and Sompolinsky, H. (2008b). Memory traces in dynamical systems. *Proceedings of the National Academy of Sciences*, 105(48):18970–18975.

[Ganguli and Latham, 2009] Ganguli, S. and Latham, P. (2009). Feedforward to the Past: The Relation between Neuronal Connectivity, Amplification, and Short-Term Memory. *Neuron*, 61(4):499–501.

[Ganguly and Carmena, 2009] Ganguly, K. and Carmena, J. M. (2009). Emergence of a Stable Cortical Map for Neuroprosthetic Control. *PLOS Biology*, 7(7):e1000153.

[Ganguly et al., 2011] Ganguly, K., Dimitrov, D. F., Wallis, J. D., and Carmena, J. M. (2011). Reversible large-scale modification of cortical networks during neuroprosthetic control. *Nature Neuroscience*, 14(5):662–667.

[Gao and Ganguli, 2015] Gao, P. and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, 32:148–155.

[Gao et al., 2017] Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., and Ganguli, S. (2017). A theory of multineuronal dimensionality, dynamics and measurement. *bioRxiv*, page 214262.

[Gawthrop et al., 2011] Gawthrop, P., Loram, I., Lakie, M., and Gollee, H. (2011). Intermittent control: A computational theory of human control. *Biological Cybernetics*, 104(1):31–51.

[Georgopoulos et al., 1986] Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419.

[Gilja et al., 2012] Gilja, V., Nuyujukian, P., Chestek, C. A., Cunningham, J. P., Yu, B. M., Fan, J. M., Churchland, M. M., Kaufman, M. T., Kao, J. C., Ryu, S. I., and Shenoy, K. V. (2012). A high-performance neural prosthesis enabled by control algorithm design. *Nature Neuroscience*, 15(12):1752–1757.

[Goldman, 2009] Goldman, M. S. (2009). Memory without Feedback in a Neural Network. *Neuron*, 61(4):621–634.

[Golub et al., 2016] Golub, M. D., Chase, S. M., Batista, A. P., and Yu, B. M. (2016). Brain–computer interfaces for dissecting cognitive processes underlying sensorimotor control. *Current Opinion in Neurobiology*, 37:53–58.

[Golub et al., 2018] Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., and Yu, B. M. (2018). Learning by neural reassociation. *Nature Neuroscience*, 21(4):607–616.

[Golub et al., 2015] Golub, M. D., Yu, B. M., and Chase, S. M. (2015). Internal models for interpreting neural population activity during sensorimotor control. *eLife*, 4:e10015.

[Golub et al., 2014] Golub, M. D., Yu, B. M., Schwartz, A. B., and Chase, S. M. (2014). Motor cortical control of movement speed with implications for brain-machine interface control. *Journal of Neurophysiology*, 112(2):411–429.

[Gross et al., 2002] Gross, J., Timmermann, L., Kujala, J., Dirks, M., Schmitz, F., Salmelin, R., and Schnitzler, A. (2002). The neural basis of intermittent motor control in humans. *Proceedings of the National Academy of Sciences*, 99(4):2299–2302.

[Hari et al., 1998] Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences*, 95(25):15061–15065.

[Hassabis et al., 2017] Hassabis, D., Kumaran, D., Summerfield, C., and Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. *Neuron*, 95(2):245–258.

[Hayashi-Takagi et al., 2015] Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y. I., Loshbaugh, A. L., Kuhlman, B., Hahn, K. M., and Kasai, H. (2015). Labelling and optical erasure of synaptic memory traces in the motor cortex. *Nature*, 525(7569):333–338.

[Heald et al., 2020] Heald, J. B., Lengyel, M., and Wolpert, D. M. (2020). Contextual inference underlies the learning of sensorimotor repertoires. *bioRxiv*, page 2020.11.23.394320.

[Héliot et al., 2010] Héliot, R., Ganguly, K., Jimenez, J., and Carmena, J. M. (2010). Learning in Closed-Loop Brain–Machine Interfaces: Modeling and Experimental Validation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(5):1387–1397.

[Hennequin, 2019] Hennequin, G. (2019). Neuroscience out of control: Control-theoretic methods for understanding neural circuits. In *Cosyne Workshop "Data, Dynamics and Computation: Using Data-Driven Methods to Ground Mechanistic Theory"*, Lisbon, Portugal.

[Hennequin et al., 2012] Hennequin, G., Vogels, T. P., and Gerstner, W. (2012). Non-normal amplification in random balanced neuronal networks. *Physical Review E*, 86(1):011909.

[Hennequin et al., 2014] Hennequin, G., Vogels, T. P., and Gerstner, W. (2014). Optimal Control of Transient Dynamics in Balanced Networks Supports Generation of Complex Movements. *Neuron*, 82(6):1394–1406.

[Hennig et al., 2018] Hennig, J. A., Golub, M. D., Lund, P. J., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Yu, B. M., and Chase, S. M. (2018). Constraints on neural redundancy. *eLife*, 7:e36774.

[Hennig et al., 2020] Hennig, J. A., Oby, E. R., Golub, M. D., Bahureksa, L. A., Sadtler, P. T., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., and Yu, B. M. (2020). Learning is shaped by abrupt changes in neural engagement. *bioRxiv*, page 2020.05.24.112714.

[Hochberg et al., 2006] Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., Branner, A., Chen, D., Penn, R. D., and Donoghue, J. P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442(7099):164–171.

[Hwang et al., 2013] Hwang, E. J., Bailey, P. M., and Andersen, R. A. (2013). Volitional Control of Neural Activity Relies on the Natural Motor Repertoire. *Current Biology*, 23(5):353–361.

[Ivanenko et al., 2003] Ivanenko, Y. P., Grasso, R., Zago, M., Molinari, M., Scivoletto, G., Castellano, V., Macellari, V., and Lacquaniti, F. (2003). Temporal Components of the Motor Patterns Expressed by the Human Spinal Cord Reflect Foot Kinematics. *Journal of Neurophysiology*, 90(5):3555–3565.

[Jaeger, 2001] Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13.

[Jarosiewicz et al., 2008] Jarosiewicz, B., Chase, S. M., Fraser, G. W., Velliste, M., Kass, R. E., and Schwartz, A. B. (2008). Functional network reorganization during learning in a brain-computer interface paradigm. *Proceedings of the National Academy of Sciences*, 105(49):19486–19491.

[Jazayeri and Shadlen, 2010] Jazayeri, M. and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, 13(8):1020–1026.

[Jeannerod and Frak, 1999] Jeannerod, M. and Frak, V. (1999). Mental imaging of motor activity in humans. *Current Opinion in Neurobiology*, 9(6):735–739.

[Jiang et al., 2020] Jiang, X., Saggar, H., Ryu, S. I., Shenoy, K. V., and Kao, J. C. (2020). Structure in Neural Activity during Observed and Executed Movements Is Shared at the Neural Population Level, Not in Single Neurons. *Cell Reports*, 32(6):108006.

[Johnson et al., 1996] Johnson, P. B., Ferraina, S., Bianchi, L., and Caminiti, R. (1996). Cortical Networks for Visual Reaching: Physiological and Anatomical Organization of Frontal and Parietal Lobe Arm Regions. *Cerebral Cortex*, 6(2):102–119.

[Kao and Hennequin, 2018] Kao, T.-C. and Hennequin, G. (2018). Null Ain't Dull: New Perspectives on Motor Cortex. *Trends in Cognitive Sciences*, 22(12):1069–1071.

[Kao and Hennequin, 2019] Kao, T.-C. and Hennequin, G. (2019). Neuroscience out of control: Control-theoretic perspectives on neural circuit dynamics. *Current Opinion in Neurobiology*, 58:122–129.

[Kao et al., 2020] Kao, T.-C., Sadabadi, M. S., and Hennequin, G. (2020). Optimal anticipatory control as a theory of motor preparation: A thalamo-cortical circuit model. *bioRxiv*, page 2020.02.02.931246.

[Kaufman et al., 2014] Kaufman, M. T., Churchland, M. M., Ryu, S. I., and Shenoy, K. V. (2014). Cortical activity in the null space: Permitting preparation without movement. *Nature Neuroscience*, 17(3):440–448.

[Kaufman et al., 2016] Kaufman, M. T., Seely, J. S., Sussillo, D., Ryu, S. I., Shenoy, K. V., and Churchland, M. M. (2016). The Largest Response Component in the Motor Cortex Reflects Movement Timing but Not Movement Type. *eNeuro*, 3(4).

[Kawai et al., 2015] Kawai, R., Markman, T., Poddar, R., Ko, R., Fantana, A. L., Dhawale, A. K., Kampff, A. R., and Ölveczky, B. P. (2015). Motor Cortex Is Required for Learning but Not for Executing a Motor Skill. *Neuron*, 86(3):800–812.

[Keele and Posner, 1968] Keele, S. W. and Posner, M. I. (1968). Processing of visual feedback in rapid movements. *Journal of experimental psychology*, 77(1):155.

[Kingma and Ba, 2017] Kingma, D. P. and Ba, J. (2017). Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*.

[Kleim et al., 2004] Kleim, J. A., Hogg, T. M., VandenBerg, P. M., Cooper, N. R., Bruneau, R., and Remple, M. (2004). Cortical Synaptogenesis and Motor Map Reorganization Occur during Late, But Not Early, Phase of Motor Skill Learning. *Journal of Neuroscience*, 24(3):628–633.

[Koralek et al., 2012] Koralek, A. C., Jin, X., Long Ii, J. D., Costa, R. M., and Carmena, J. M. (2012). Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. *Nature*, 483(7389):331–335.

[Krakauer et al., 2019] Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., and Haith, A. M. (2019). Motor Learning. In *Comprehensive Physiology*, pages 613–663. American Cancer Society.

[Krakauer and Mazzoni, 2011] Krakauer, J. W. and Mazzoni, P. (2011). Human sensorimotor learning: Adaptation, skill, and beyond. *Current Opinion in Neurobiology*, 21(4):636–644.

[Kuppuswamy and Harris, 2014] Kuppuswamy, N. and Harris, C. M. (2014). Do muscle synergies reduce the dimensionality of behavior? *Frontiers in Computational Neuroscience*, 8.

[Kutch and Valero-Cuevas, 2012] Kutch, J. J. and Valero-Cuevas, F. J. (2012). Challenges and New Approaches to Proving the Existence of Muscle Synergies of Neural Origin. *PLOS Computational Biology*, 8(5):e1002434.

[Kwakernaak and Sivan, 1972] Kwakernaak, H. and Sivan, R. (1972). *Linear Optimal Control Systems*, volume 1. Wiley-interscience New York.

[Lee, 2012] Lee, J. M. (2012). Smooth Manifolds. In Lee, J. M., editor, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, pages 1–31. Springer, New York, NY.

[Legenstein et al., 2010] Legenstein, R., Chase, S. M., Schwartz, A. B., and Maass, W. (2010). A Reward-Modulated Hebbian Learning Rule Can Explain Experimentally Observed Network Reorganization in a Brain Control Task. *Journal of Neuroscience*, 30(25):8400–8410.

[Legenstein et al., 2008] Legenstein, R., Pecevski, D., and Maass, W. (2008). A Learning Theory for Reward-Modulated Spike-Timing-Dependent Plasticity with Application to Biofeedback. *PLOS Computational Biology*, 4(10):e1000180.

[Lemon, 2008] Lemon, R. N. (2008). Descending Pathways in Motor Control. *Annual Review of Neuroscience*, 31(1):195–218.

[Li et al., 2001] Li, C.-S. R., Padoa-Schioppa, C., and Bizzi, E. (2001). Neuronal Correlates of Motor Performance and Motor Learning in the Primary Motor Cortex of Monkeys Adapting to an External Force Field. *Neuron*, 30(2):593–607.

[Lillicrap et al., 2020] Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6):335–346.

[Logiaco et al., 2019] Logiaco, L., Abbott, L. F., and Escola, S. (2019). A model of flexible motor sequencing through thalamic control of cortical dynamics. *bioRxiv*, page 2019.12.17.880153.

[Logiaco and Escola, 2020] Logiaco, L. and Escola, G. S. (2020). Thalamocortical motor circuit insights for more robust hierarchical control of complex sequences. *arXiv:2006.13332 [cs, q-bio, stat]*.

[Lopes et al., 2017] Lopes, G., Nogueira, J., Dimitriadis, G., Menendez, J. A., Paton, J. J., and Kampff, A. R. (2017). A robust role for motor cortex. *bioRxiv*, page 058917.

[Losey et al., 2020] Losey, D., Hennig, J., Oby, E., Golub, M., Sadtler, P., Quick, K., Ryu, S., Tyler-Kabara, E., Batista, A., Yu, B., and Chase, S. (2020). Evidence of a memory trace in motor cortex after short-term learning. In *Cosyne Abstracts*, Denver, CO.

[Mastrogiuseppe and Ostojic, 2018] Mastrogiuseppe, F. and Ostojic, S. (2018). Linking Connectivity, Dynamics, and Computations in Low-Rank Recurrent Neural Networks. *Neuron*, 99(3):609–623.e29.

[Mathis et al., 2017] Mathis, M. W., Mathis, A., and Uchida, N. (2017). Somatosensory Cortex Plays an Essential Role in Forelimb Motor Adaptation in Mice. *Neuron*, 93(6):1493–1503.e6.

[Mazzoni and Krakauer, 2006] Mazzoni, P. and Krakauer, J. W. (2006). An Implicit Plan Overrides an Explicit Strategy during Visuomotor Adaptation. *Journal of Neuroscience*, 26(14):3642–3645.

[McDougle et al., 2015] McDougle, S. D., Bond, K. M., and Taylor, J. A. (2015). Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *Journal of Neuroscience*, 35(26):9568–9579.

[McDougle et al., 2016] McDougle, S. D., Ivry, R. B., and Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences*, 20(7):535–544.

[McDougle and Taylor, 2019] McDougle, S. D. and Taylor, J. A. (2019). Dissociable cognitive strategies for sensorimotor learning. *Nature Communications*, 10(1):1–13.

[Merel et al., 2015] Merel, J., Pianto, D. M., Cunningham, J. P., and Paninski, L. (2015). Encoder-Decoder Optimization for Brain-Computer Interfaces. *PLOS Computational Biology*, 11(6):e1004288.

[Minsky, 1961] Minsky, M. (1961). Steps toward Artificial Intelligence. *Proceedings of the IRE*, 49(1):8–30.

[Moran and Schwartz, 1999] Moran, D. W. and Schwartz, A. B. (1999). Motor Cortical Representation of Speed and Direction During Reaching. *Journal of Neurophysiology*, 82(5):2676–2692.

[Mountcastle et al., 1975] Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H., and Acuna, C. (1975). Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. *Journal of Neurophysiology*, 38(4):871–908.

[Mulliken et al., 2008] Mulliken, G. H., Musallam, S., and Andersen, R. A. (2008). Decoding Trajectories from Posterior Parietal Cortex Ensembles. *Journal of Neuroscience*, 28(48):12913–12926.

[Munzert et al., 2009] Munzert, J., Lorey, B., and Zentgraf, K. (2009). Cognitive motor processes: The role of motor imagery in the study of motor representations. *Brain Research Reviews*, 60(2):306–326.

[Murphy and Miller, 2009] Murphy, B. K. and Miller, K. D. (2009). Balanced Amplification: A New Mechanism of Selective Amplification of Neural Activity Patterns. *Neuron*, 61(4):635–648.

[Murray and Escola, 2020] Murray, J. M. and Escola, G. S. (2020). Remembrance of things practiced: Fast and slow learning in cortical and subcortical pathways. *bioRxiv*, page 797548.

[Musallam et al., 2004] Musallam, S., Corneil, B. D., Greger, B., Scherberger, H., and Andersen, R. A. (2004). Cognitive Control Signals for Neural Prosthetics. *Science*, 305(5681):258–262.

[Neely et al., 2018] Neely, R. M., Koralek, A. C., Athalye, V. R., Costa, R. M., and Carmena, J. M. (2018). Volitional Modulation of Primary Visual Cortex Activity Requires the Basal Ganglia. *Neuron*, 97(6):1356–1368.e4.

[Oby et al., 2019] Oby, E. R., Golub, M. D., Hennig, J. A., Degenhart, A. D., Tyler-Kabara, E. C., Byron, M. Y., Chase, S. M., and Batista, A. P. (2019). New neural activity patterns emerge with long-term learning. *Proceedings of the National Academy of Sciences*, 116(30):15210–15215.

[Omrani et al., 2017] Omrani, M., Kaufman, M. T., Hatsopoulos, N. G., and Cheney, P. D. (2017). Perspectives on classical controversies about the motor cortex. *Journal of Neurophysiology*, 118(3):1828–1848.

[Orsborn et al., 2014] Orsborn, A. L., Moorman, H. G., Overduin, S. A., Shanechi, M. M., Dimitrov, D. F., and Carmena, J. M. (2014). Closed-Loop Decoder Adaptation Shapes Neural Plasticity for Skillful Neuroprosthetic Control. *Neuron*, 82(6):1380–1393.

[Orsborn and Pesaran, 2017] Orsborn, A. L. and Pesaran, B. (2017). Parsing learning in networks using brain–machine interfaces. *Current Opinion in Neurobiology*, 46:76–83.

[Paz et al., 2005] Paz, R., Nathan, C., Boraud, T., Bergman, H., and Vaadia, E. (2005). Acquisition and generalization of visuomotor transformations by nonhuman primates. *Experimental Brain Research*, 161(2):209–219.

[Perdikis and Millan, 2020] Perdikis, S. and Millan, J. d. R. (2020). Brain-Machine Interfaces: A Tale of Two Learners. *IEEE Systems, Man, and Cybernetics Magazine*, 6(3):12–19.

[Perich et al., 2018] Perich, M. G., Gallego, J. A., and Miller, L. E. (2018). A Neural Population Mechanism for Rapid Learning. *Neuron*, 100(4):964–976.e7.

[Peters et al., 2017] Peters, A. J., Liu, H., and Komiyama, T. (2017). Learning in the Rodent Motor Cortex. *Annual Review of Neuroscience*, 40(1):77–97.

[Raman et al., 2019] Raman, D. V., Rotondo, A. P., and O'Leary, T. (2019). Fundamental bounds on learning performance in neural circuits. *Proceedings of the National Academy of Sciences*, 116(21):10537–10546.

[Rastogi et al., 2020] Rastogi, A., Vargas-Irwin, C. E., Willett, F. R., Abreu, J., Crowder, D. C., Murphy, B. A., Memberg, W. D., Miller, J. P., Sweet, J. A., Walter, B. L., Cash, S. S., Rezaii, P. G., Franco, B., Saab, J., Stavisky, S. D., Shenoy, K. V., Henderson, J. M., Hochberg, L. R., Kirsch, R. F., and Ajiboye, A. B. (2020). Neural Representation of Observed, Imagined, and Attempted Grasping Force in Motor Cortex of Individuals with Chronic Tetraplegia. *Scientific Reports*, 10(1):1429.

[Rokni et al., 2007] Rokni, U., Richardson, A. G., Bizzi, E., and Seung, H. S. (2007). Motor Learning with Unstable Neural Representations. *Neuron*, 54(4):653–666.

[Russo et al., 2018] Russo, A. A., Bittner, S. R., Perkins, S. M., Seely, J. S., London, B. M., Lara, A. H., Miri, A., Marshall, N. J., Kohn, A., Jessell, T. M., Abbott, L. F., Cunningham, J. P., and Churchland, M. M. (2018). Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response. *Neuron*, 97(4):953–966.e8.

[Sadtler et al., 2014] Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., Yu, B. M., and Batista, A. P. (2014). Neural constraints on learning. *Nature*, 512(7515):423–426.

[Sakellaridi et al., 2019] Sakellaridi, S., Christopoulos, V. N., Aflalo, T., Pejsa, K. W., Rosario, E. R., Ouellette, D., Pouratian, N., and Andersen, R. A. (2019). Intrinsic Variable Learning for Brain-Machine Interface Control by Human Anterior Intraparietal Cortex. *Neuron*, 102(3):694–705.e3.

[Sauerbrei et al., 2020] Sauerbrei, B. A., Guo, J.-Z., Cohen, J. D., Mischiati, M., Guo, W., Kabra, M., Verma, N., Mensh, B., Branson, K., and Hantman, A. W. (2020). Cortical pattern generation during dexterous movement is input-driven. *Nature*, 577(7790):386–391.

[Scharnowski et al., 2015] Scharnowski, F., Veit, R., Zopf, R., Studer, P., Bock, S., Diedrichsen, J., Goebel, R., Mathiak, K., Birbaumer, N., and Weiskopf, N. (2015). Manipulating motor performance and memory through real-time fMRI neurofeedback. *Biological Psychology*, 108:85–97.

[Schuessler et al., 2020] Schuessler, F., Mastrogiuseppe, F., Dubreuil, A., Ostojic, S., and Barak, O. (2020). The interplay between randomness and structure during learning in RNNs. *arXiv:2006.11036 [q-bio]*.

[Scott, 2004] Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532–545.

[Scott, 2008] Scott, S. H. (2008). Inconvenient Truths about neural processing in primary motor cortex. *The Journal of Physiology*, 586(5):1217–1224.

[Serruya et al., 2002] Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R., and Donoghue, J. P. (2002). Instant neural control of a movement signal. *Nature*, 416(6877):141–142.

[Seung, 1996] Seung, H. S. (1996). How the brain keeps the eyes still. *Proceedings of the National Academy of Sciences*, 93(23):13339–13344.

[Shadmehr and Krakauer, 2008] Shadmehr, R. and Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, 185(3):359–381.

[Shadmehr and Mussa-Ivaldi, 1994] Shadmehr, R. and Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience*, 14(5):3208–3224.

[Shadmehr et al., 2010] Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience*, 33(1):89–108.

[Shanechi et al., 2017] Shanechi, M. M., Orsborn, A. L., Moorman, H. G., Gowda, S., Dangi, S., and Carmena, J. M. (2017). Rapid control and feedback rates enhance neuroprosthetic control. *Nature Communications*, 8(1):13825.

[Shenoy and Carmena, 2014] Shenoy, K. V. and Carmena, J. M. (2014). Combining Decoder Design and Neural Adaptation in Brain-Machine Interfaces. *Neuron*, 84(4):665–680.

[Shenoy et al., 2013] Shenoy, K. V., Sahani, M., and Churchland, M. M. (2013). Cortical Control of Arm Movements: A Dynamical Systems Perspective. *Annual Review of Neuroscience*, 36(1):337–359.

[Silversmith et al., 2020] Silversmith, D. B., Abiri, R., Hardy, N. F., Natraj, N., Tu-Chan, A., Chang, E. F., and Ganguly, K. (2020). Plug-and-play control of a brain–computer interface through neural map stabilization. *Nature Biotechnology*, pages 1–10.

[Sitaram et al., 2017] Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M. L., Rana, M., Oblak, E., Birbaumer, N., and Sulzer, J. (2017). Closed-loop brain training: The science of neurofeedback. *Nature Reviews Neuroscience*, 18(2):86–100.

[Smith et al., 2006] Smith, M. A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLOS Biology*, 4(6):e179.

[Snyder et al., 1997] Snyder, L. H., Batista, A. P., and Andersen, R. A. (1997). Coding of intention in the posterior parietal cortex. *Nature*, 386(6621):167–170.

[Sohn et al., 2019] Sohn, H., Narain, D., Meirhaeghe, N., and Jazayeri, M. (2019). Bayesian Computation through Cortical Latent Dynamics. *Neuron*, 103(5):934–947.e5.

[Stavisky et al., 2017] Stavisky, S. D., Kao, J. C., Ryu, S. I., and Shenoy, K. V. (2017). Motor Cortical Visuomotor Feedback Activity Is Initially Isolated from Downstream Targets in Output-Null Neural State Space Dimensions. *Neuron*, 95(1):195–208.e9.

[Sun et al., 2020] Sun, X., O'Shea, D. J., Golub, M. D., Trautmann, E. M., Vyas, S., Ryu, S. I., and Shenoy, K. V. (2020). Skill-specific changes in cortical preparatory activity during motor learning. *bioRxiv*, page 2020.01.30.919894.

[Susman et al., 2021] Susman, L., Mastrogiuseppe, F., Brenner, N., and Barak, O. (2021). Quality of internal representation shapes learning performance in feedback neural networks. *Physical Review Research*, 3(1):013176.

[Sussillo and Abbott, 2009] Sussillo, D. and Abbott, L. F. (2009). Generating Coherent Patterns of Activity from Chaotic Neural Networks. *Neuron*, 63(4):544–557.

[Sussillo et al., 2015] Sussillo, D., Churchland, M. M., Kaufman, M. T., and Shenoy, K. V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nature Neuroscience*, 18(7):1025–1033.

[Taylor et al., 2002] Taylor, D. M., Tillery, S. I. H., and Schwartz, A. B. (2002). Direct Cortical Control of 3D Neuroprosthetic Devices. *Science*, 296(5574):1829–1832.

[Taylor and Ivry, 2012] Taylor, J. A. and Ivry, R. B. (2012). The role of strategies in motor learning. *Annals of the New York Academy of Sciences*, 1251(1):1–12.

[Taylor et al., 2014] Taylor, J. A., Krakauer, J. W., and Ivry, R. B. (2014). Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *Journal of Neuroscience*, 34(8):3023–3032.

[Thach, 1978] Thach, W. T. (1978). Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. *Journal of Neurophysiology*, 41(3):654–676.

[Tipping and Bishop, 1999] Tipping, M. E. and Bishop, C. M. (1999). Probabilistic Principal Component Analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3):611–622.

[Tkach et al., 2007] Tkach, D., Reimer, J., and Hatsopoulos, N. G. (2007). Congruent Activity during Action and Action Observation in Motor Cortex. *Journal of Neuroscience*, 27(48):13241–13250.

[Todorov, 2000] Todorov, E. (2000). Direct cortical control of muscle activation in voluntary arm movements: A model. *Nature Neuroscience*, 3(4):391–398.

[Todorov, 2004] Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9):907–915.

[Todorov and Jordan, 2002] Todorov, E. and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226–1235.

[Todorov and Li, 2003] Todorov, E. and Li, W. (2003). Optimal control methods suitable for biomechanical systems. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, volume 2, pages 1758–1761 Vol.2.

[Ungerleider et al., 2002] Ungerleider, L. G., Doyon, J., and Karni, A. (2002). Imaging Brain Plasticity during Motor Skill Learning. *Neurobiology of Learning and Memory*, 78(3):553–564.

[Velliste et al., 2008] Velliste, M., Perel, S., Spalding, M. C., Whitford, A. S., and Schwartz, A. B. (2008). Cortical control of a prosthetic arm for self-feeding. *Nature*, 453(7198):1098–1101.

[Vendrell-Llopis et al., 2019] Vendrell-Llopis, N., Koralek, A., Costa, R., and Carmena, J. (2019). Ventral striatum uses a temporal difference rule for prediction during neuroprosthetic control. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 562–565.

[Vigneswaran et al., 2013] Vigneswaran, G., Philipp, R., Lemon, R. N., and Kraskov, A. (2013). M1 Corticospinal Mirror Neurons and Their Role in Movement Suppression during Action Observation. *Current Biology*, 23(3):236–243.

[Vyas et al., 2018] Vyas, S., Even-Chen, N., Stavisky, S. D., Ryu, S. I., Nuyujukian, P., and Shenoy, K. V. (2018). Neural Population Dynamics Underlying Motor Learning Transfer. *Neuron*, 97(5):1177–1186.e3.

[Vyas et al., 2020] Vyas, S., O'Shea, D. J., Ryu, S. I., and Shenoy, K. V. (2020). Causal Role of Motor Preparation during Error-Driven Learning. *Neuron*, 106(2):329–339.e4.

[Wang et al., 2018] Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., and Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, 21(6):860–868.

[Wärnberg and Kumar, 2019] Wärnberg, E. and Kumar, A. (2019). Perturbing low dimensional activity manifolds in spiking neuronal networks. *PLOS Computational Biology*, 15(5):e1007074.

[White et al., 2004] White, O. L., Lee, D. D., and Sompolinsky, H. (2004). Short-Term Memory in Orthogonal Neural Networks. *Physical Review Letters*, 92(14):148102.

[Willett et al., 2020] Willett, F. R., Deo, D. R., Avansino, D. T., Rezaii, P., Hochberg, L. R., Henderson, J. M., and Shenoy, K. V. (2020). Hand Knob Area of Premotor Cortex Represents the Whole Body in a Compositional Way. *Cell*, 181(2):396–409.e26.

[Wise et al., 1997] Wise, S. P., Boussaoud, D., Johnson, P. B., and Caminiti, R. (1997). Premotor and Parietal Cortex: Corticocortical Connectivity and Combinatorial Computations. *Annual Review of Neuroscience*, 20(1):25–42.

[Wolpert et al., 1995] Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882.

[Yin et al., 2009] Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., and Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*, 12(3):333–341.

[Yu et al., 2009] Yu, B. M., Cunningham, J. P., Santhanam, G., Ryu, S. I., Shenoy, K. V., and Sahani, M. (2009). Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology*, 102(1):614–635.

[Zhang et al., 2017] Zhang, C. Y., Aflalo, T., Revechkis, B., Rosario, E. R., Ouellette, D., Pouratian, N., and Andersen, R. A. (2017). Partially Mixed Selectivity in Human Posterior Parietal Association Cortex. *Neuron*, 95(3):697–708.e4.

[Zhou et al., 2021] Zhou, L., Latham, P., and Menendez, J. A. (2021). Learning sensory representations for flexible computation with recurrent circuits. In *Cosyne Abstracts*.

[Zhou et al., 2019] Zhou, X., Tien, R. N., Ravikumar, S., and Chase, S. M. (2019). Distinct types of neural reorganization during long-term learning. *Journal of Neurophysiology*, 121(4):1329–1341.