

# Heterarchical Control in Sensorimotor Processing

Spencer R. Wilson

March 6, 2021

## 1 Introduction & Aims

*Movement is nothing but the quality of our being.*

— Sunryu Suzuki, *Zen Mind, Beginner’s Mind*

The purpose of this thesis is to build a high-dimensional electromyography recording setup, and use this to produce a platform for closed-loop motor control and learning experiments on which a suite of tasks and perturbations can be tested. With this relatively novel setup, we hope to advance our understanding of trial-to-trial motor learning by testing new model predictions through mixture of experiment and theory.

Hans Moravec’s eponymous paradox states that it is easier to generate artificially intelligent performance on tasks we think of as intellectually challenging, such as chess, than to provide a machine with faculties we take for granted, such as movement. Moravec’s Paradox, for example, encourages us not to look past the complex computations generated by the human motor system. Following Moravec, this work focuses on what is arguably the most advanced control apparatus in the known universe: the human movement machine.

A recent review provides a clear call to action for work this direction:

The processes by which biological control solutions spanning large and continuous state spaces are constructed remain relatively unexplored. Future investigations may need to embed rich dynamical interactions between object dynamics and task goals in novel and complex movements.<sup>1</sup>

Over the last few decades, there has been considerable amount of work done to untangle the abilities of the motor system to flexibly control the body including through optimal control theory<sup>2</sup>, reinforcement learning in continuous action space<sup>3</sup>, and detailed physiological studies<sup>4</sup>. However, as the quote above suggests, a holistic understanding of the computations underlying the generation of skilled movement remains an open problem. The aim of this thesis is to progress our understanding of skilled movement by studying the solutions produced by human subjects to motor tasks in dynamically rich, yet controlled, virtual environments. Our goal is to reverse-engineer the ability to acquire and perform novel motor skills. First we define what we mean by the terms *skill*, and *task*.

Humans produce a great variety of movements every day, often without conscious thought. For example, movements like bringing a cup of coffee to our lips for a sip are generally out of reach for state-of-the-art robotic systems. We claim that this “motor gap” between biological and artificial motor systems is due to a lack of *dexterity* in the latter. Soviet neuroscientist Nikolai Bernstein defined dexterity as the ability to “find a motor solution in any situation and in any condition.”<sup>5</sup> The crux of this definition is the flexibility of such solutions. This flexibility, or robustness<sup>16</sup>, is the ability to optimize internal parameters in response to external perturbations and adapt to new information to achieve the goals of an ongoing plan.

---

<sup>1</sup>Kitano defines robustness as “the maintenance of specific functionalities of the system against perturbations, and it often requires the system to change its mode of operation in a flexible way.” He claims that robustness requires control, alternative mechanisms, modularity and decoupling between high and low level variability.

While a robot may be able to move a cup of coffee to a precise location in space, its solution is often found to be brittle in a new context, or unable to generalize to the movement of new objects. We define a skill as a behavior that involves dexterity in Bernstein's sense. The use of a tool such as a screwdriver is an example of a motor skill. We define a task as the production of skilled movement in a particular context. Driving a screw in a particular posture using a particular screwdriver is an example of a task. These concepts will be further formalized in later chapters.

Human movement is ultimately the result of the activation and contraction of muscle fibers, and movements lie on a spectrum between reflexive and volitional. The supramuscular circuitry which determines the degree of volition we ascribe to movement, where volitional movement relies on supraspinal (though not necessarily conscious) processes. The human hand is a unique evolutionary invention that underlies our ability to perform various skills in a range of tasks— movements that are decidedly volitional<sup>2</sup>. The hand is the pinnacle of dexterity and, as such, it is a fruitful testbed for studying the computations and circuitry that drive dexterous movement. A detailed physiological review of the hand and its relation to skilled movement is described in Section 2.

We are more interested in the leveraging the hand as a readout of flexible motor behavior than we are in the kinematics of hand movement itself. For this reason, we chose to develop an experimental setup that is capable of recording directly from muscle activations. We achieve this through surface electromyography recordings taken from the forearms controlling subjects' dominant hands. This allows us to track the sequential selection of muscle activations during both skill acquisition and subsequent performance of that skill to achieve desired goal. As we are interested in subjects' abilities to acquire new skills, we design tasks that require subjects to use available, but uncommon, motor activations. We then track the selection and execution of these activation during virtual tasks. The details of how this is achieved are described in Section 4.

Using data from our experimental setup, we hope to gain an understanding of how the structure of muscle activation variability in evolves during skill acquisition and how the motor system constructs skilled movement through the composition of component muscle coactivations. We believe that to make progress on these two lines of enquiry we should work to reconcile the language of the experimental sensorimotor control and learning community with the language of the control theory and reinforcement learning community, as each of these communities shares a common goal of understanding the computation underlying the production of skilled movement. Here we develop several models in this direction, as described in ??.

## 2 Physiology of the Skilled Movement

*Even a simple movement is a global body event.*

— Bizzi & Ajemian, 2020

As we hope to make progress engineering naturalistic artificial movement, it will be beneficial to review what is known about the biological movement system. Beginning with the architecture of the motor system and its relation to dexterity will provide a scaffold on which we can hang our experimental and theoretical investigations detailed in Section 4 and ?? . Specifically, we can use results from prior physiological investigations to ground our perspective on the computations relevant to skilled hand movements. We find that the dexterous solutions produced by the human motor system rely on a incredibly complex architecture, but one in which a spectrum of modularity and redundancy appear to be organizing principles.

### 2.1 Motor Units to Muscles

Muscles are composed of fibers which contract due to chemical gradients produced at the neuromuscular junction by action potentials emanating from alpha-motoneurons (AMN) in the ventral horn of the spinal cord. The quantum of motor output is the motor unit (MU), which is defined as

---

<sup>2</sup>It could be argued that the hand is in fact a crucial aspect of humanness. It is thought that the human cerebellar and neocortices evolved reciprocally to expand and support the computational burden of increasingly complex motor tasks such as tool-making and language production. REF?

a single motoneuron axon and the set of junctions its axon branches form with one or more muscle fibers. The innervation ratio of a particular muscle unit is the number of junctions it innervates. In muscles of the arm, the number of MUs and their innervation ratios each range from tens to hundreds per muscle and per motor unit, respectively, decreasing as muscles become more distal.

The MU thus provides the motor system with spatial redundancy at the muscle level: multiple muscle fibers contract due to a single AMN spike, and multiple AMNs may overlap in their innervations. The forces produced by motor units span several orders of magnitude, though most units produce very small forces. Here we find temporal redundancy: in order to produce movements, MUs combine to generate a range of forces<sup>7</sup>.

Since the innervation ratios of muscles in the forearm and hand are relatively small compared to more proximal muscles (which contain thousands of MUs), the logarithmic recruitment and redundancy of motor units enables the hand to produce movements with very fine spatiotemporal resolution. Paradoxically, however, the well-known signal-dependent noise in models of motor output has been found to be higher for hand muscles than for more proximal muscles, likely due to small numbers of motor units compare to larger muscles<sup>7,8</sup>.

Muscle fibers are contained within muscle compartments, and each muscle may have one or more compartments. The fingers of the hand are extended by the extensor digitorum (ED) which contains four compartments, one for each of the tendons the muscle produces. Each tendon connects to the three metacarpal joints of each digit. The fingers are flexed by two muscles, the flexor digitorum superficialis (FDS) and the flexor digitorum profundus (FDP). Like the ED, these muscles produce four tendons, one to each finger from each of their four compartments. As such, one must coactivate these agonist and antagonist muscles in order to extend or flex a single finger in isolation<sup>7</sup>. Adduction and Abduction of the fingers is produced by the 19 intrinsic muscles of the hand, each of which has their origin and insertion points within the hand itself<sup>9</sup>. The intrinsic muscle tendons form a kind of network around each of the digits.

The human hand, thumb, and forearm system contains more than 30 muscles and at least 20 degrees of freedom are theoretically available for actuation. However, due to biomechanical coupling, the effective degrees of freedom is presumably less than 20. One study found that tendons of the fingers are arranged in such a way as to perform a kind of anatomical computation which expands the mechanical capabilities of the appendage by sharing force across its tendon network<sup>10</sup>. Such computations embedded in the musculoskeletal structure are additional complexity when theorizing about neural control of the hand.

We believe this structure exists in order to facilitate the acquisition of new skills and the generalization of existing skills to new contexts. While the anatomy of the hand and forearm presents constraints on movement, the system remains capable of producing a incredible variety of movement patterns<sup>11,123</sup>. The structure of the neuromuscular system that underlies this variety offers many clues as to the relevant computations required for dexterous movement.

## FIGURE OF MUSCLES, MOTOR UNITS, CORD, DIFFERENT ARRANGEMENTS – this can be referenced later for coordinative/fractionative

### 2.2 Coordinative Structures

*We have some idea as to the intricate design of the puppet and the puppet strings, but we lack insight into the mind of the puppeteer.*

— Bizzi & Ajemian, 2020

Nikolai Bernstein coined the phrase “the degrees-of-freedom problem” to describe the challenge the motor system faces in coordinating its many dimension to achieve a goal. Solving this problem requires dexterity.<sup>5</sup> As we have seen, redundancy is present from joints and muscles to motor units and their upstream synaptic partners. However, rather than asking how the motor control system deals with this “problem” overwhelming complexity, we might instead question why this complexity

---

<sup>3</sup>In a classic study, Basmajian and colleagues showed that it is possible to activate single motor units in the thumb abductor.

evolved at all. What does the availability of this redundancy afford the motor system? How does this redundancy enable dexterous movement?

A considerable amount of discussion has focused on the existence of synergies as a simplifying structure which allows the motor system to “solve” the redundancy “problem”. Working to avoid the trap of pure semantics, we might ask whether “synergy” is a descriptive or normative concept. The term motor synergy can be used descriptively to describe the spatiotemporal coactivation of muscles necessary for an ongoing task. Synergetic control implies control in the space of a low-dimensional set of synergy weights rather than independent control over the actuator dimensions themselves. The control dimensions are functionally coupled as a result of synergetic action, which both simplifies the control task and constrains behavior to the low-dimensional subspace defined by the synergy weights<sup>13</sup>. This is what Bizzi and colleagues refer to as “the puppet’s strings”. The term can also be used as a normative model of motor coordination which implies a constraint in the dimensionality of the descending supraspinal control signal, the simplifying movements of the puppeteer.

Many studies have contributed to the concept of synergies as a hard-wired organizing feature of the motor system<sup>14,4</sup>. However, these works tend to extrapolate from non-primate preparations, particularly in the frog, and use tasks which are inherently low-dimensional to explain covariance structure in primate and human kinematic and electromyography data<sup>15,Gao2017?</sup>. That said, it would be foolish to deny the existence of synergistic muscle coactivation even at the structural level. Careful studies of force control by the fingertips present a complex story of dimensionality of control in this regime.<sup>16</sup> Constraints exist in the architecture of the hand as well as its control system, though we maintain that concept of synergies, especially in the context of dexterous movement, is often presented as an oversimplification rather than a mere simplification. We believe the story of the hand is more complex.

Studies have attempted to quantify the number of effective degrees of freedom of the hand with various methods. This has primarily been taken to be the number of linear features which contain a desired level of the original signal variance, where the signal is the joint angles of the hand engaged in various behaviors<sup>17,18</sup>. These methods have resulted in roughly 8 linear features of hand kinematics to solve a variety of tasks, with subtleties found in inter-task and inter-subject variations.<sup>^</sup>[It has been argued that the motor repertoire is hardly high-dimensional when compared to the dimensionality of the visual feature extraction system<sup>11</sup>. A recent study found that low-variance linear, kinematic components displayed significantly higher correlation within condition (e.g. grasp of a specific object) than across condition. This suggests that these components carry task-dependent information rather than condition-independent, task-irrelevant noise<sup>11</sup>. This suggests that the control of the hand is more nuanced than a set of fixed synergies.

What Bizzi and colleagues call “the problem of supraspinal pattern formation”—how synergies are activated through time— we argue, in the context of hand control, is not simplified by the existence of hard-wired or soft-wired synergies<sup>19</sup>. Rather, the CNS produces control signals in a range of contexts and in response to continually changing task demands. Rather than the CNS “simplifying movement” through synergetic action, it is more likely that hand synergies fall out of a optimization strategy which trades off effort and accuracy where effort may, in part, correspond to independent control of individual control dimensions. In this view, synergies, hard-wired or not, reflect the statistics of the environment in which movement is constructed<sup>20</sup>. If we limit ourselves to synergetic control, then we have simply passed the problem to a lower-dimensional one of the same fundamental nature. Neural control of the hand likely contains a spectrum of modularity in order to maintain its role as a flexible instrument. Synergetic action is one end of this spectrum resulting from the computations inherent to, along with the structures of the human movement machine.

### 2.3 Fractionating Structures

At the other end of the spectrum, years of research has contributed to a more complex picture of hand function which embraces non-synergistic movement<sup>21–23</sup>. The key insight of the work is that while “the organization of the spinal cord is based on relatively rigid muscular modes, a mechanism

---

<sup>4</sup>I should really have more studies here, or a really nice review.

to fractionate this is of particular importance for the muscles of the hands and digits which may need to be employed in a variety of flexible associations during voluntary movements.” Careful anatomical work has shown how monosynaptic corticospinal, or corticomotoneuronal (CM), connections provide such fractionation in primates which use tools requiring dexterity<sup>24</sup>. M connections are specific to the primate corticospinal tract and specific to distal muscles of the hands and arm. It appears that the rodent CST contains CM connections until they recede around P10 at which point they recede<sup>25,26</sup>.

Just as many muscle fibers may be innervated by a single AMN, up to thousands of neurons contact single AMNs through CM connections or a variety of spinal interneuron circuits. The hallmark of CM connections is their influence over multiple muscle compartments as well as multiple muscles, though typically agonist or antagonist sets<sup>27</sup>. This may seem counter-intuitive as a means to produce individuated movement, but experimental evidence in primates has shown that the convergence of many CM collateral fibers onto single AMNs driving the distal muscles in particular can produce a fine grading of activity over motor units driving the distal joints. CM cells also appear to play a role in the inhibition of antagonist muscles prior to contractions required for movement.<sup>28</sup> These findings confirm theories about the excitatory and inhibitory role of these connection dating back decades, and combine to suggest that variables encoded in cortical ensembles are more complex than kinematics or dynamics alone.<sup>27</sup>.

The CM tract thus acts in coordination with synergistic muscle activations of the hand to achieve control that is balanced between modularity and flexibility. Findings suggest that there is a bipartite structure in human motor cortex driving dexterous control of the distal part of the upper limb which, it has been suggested, evolved under pressure to quickly generalize between tasks. This work argues that these two streams of hand control, namely “fractionated” and “synergistic” control, may interact to produce versatility, and balancing these subsystems may be a key part of the optimization function when learning new skills<sup>29–31</sup>. This dualism is likely not rigidly dichotomous, but rather a spectrum of overriding fractionation (so-called “New M1”) atop a phylogenetically older system of synergistic action<sup>32</sup>. Griffin and colleagues found that CM cells are functionally tuned to a muscle’s mode of activity (agonist, antagonist, fixator) to “bypass spinal cord mechanisms and sculpt novel patterns of motor output that are essential for highly skilled movements”.<sup>30</sup>

The degree to which fractionation of movement is learned is unknown. Skilled piano performers have been found to exhibit a higher degree of independent movement among the fingers compared to control participants. Control groups displayed a hierarchical, presumably lower dimensional, organization of finger movement patterns while pianists showed distinct but individuated movement patterns<sup>33</sup>. These results imply that with skilled practice humans can produce finer and more independent movements of the fingers, and construct bespoke coactivations to solve specific goals. Similarly, studies have found that coherence between the index finger and thumb is greater on the dominant hand. This might imply a developmental lateralization, but use-dependent plasticity due to greater precision grip behavior of the dominant hand is also a viable explanation<sup>7</sup>.

The concept of a balanced control may prove to be a fruitful direction for theoretical work on dexterous motor control, the goal being to construct a model which takes into account this spectrum of individuation into account. The experimental challenge is to identify tasks which ostensibly require the direct descending connections to fractionate learned synergies. There is work suggesting that CM connections synapse primarily on low threshold, low force motor units that are recruited first. This would imply a difference in synergy fractionation at lower force as opposed to higher force. This can be tested easily by including a force parameter in a hand control task. The hypothesis stemming from the previously described work is that CM connections override the “consolidated” patterns putatively generated via spinal interneuron circuitry.

**FIGURE – CARTOON HERE DEPICTING ANATOMY DISCUSSED SO FAR – spinal circuits, motor cortex, BG, cerebellum**

## 2.4 Supraspinal Motor Maps

It is known from recent work that primary motor cortex (M1) is not an isolated movement-generating dynamical system, but rather a node in the network of a feedback-modulated, distributed movement

machine. This is reflected in recent work in the rodent which suggests that task-relevant movement depends on these network connections.<sup>4</sup> This finding is relevant for our purposes as it demonstrates a fundamental function for cortical input as opposed to a specific substructure of motor cortex as detailed above in the primate literature.

Thinking structural architecture of M1 as an input-driven system with outputs along a spectrum of modularity from synergistic to fractionated, we can then ask what kind of functional architecture might have evolved in the neuromuscular controller? Graziano and colleagues found that 500ms electrical stimulation to M1 reliably produced stereotyped movements in primates, shown in Fig. 1<sup>grazianoORGANIZATIONBEHAVIORALREPERTOIRE2006?</sup>. These movements appeared to produce goal-oriented actions pulled out of other contexts such as bringing food to the mouth, and seemed to be arranged on the cortical sheet topographically in terms of spatial endpoints rather than as a homunculus. Graziano refers to this as the cortical “action map”, that these stimulations tapped into the control mechanisms of the primate’s motor system<sup>34</sup>. These results have recently been confirmed by optogenetics work in marmosets and macaques.<sup>35,36</sup>

Recent work identifying movement syllables on a behaviorally relevant timescale has a similar flavor. Along with behavioral syllables, the motor map concept posits the idea that M1 might be looked at like a field of feedback control microcircuits, integrating and transforming inputs, both internal and external, to sculpt ongoing movement.<sup>37</sup> This is in accordance with the idea that there is a structural hierarchy in M1 covering a spectrum of movement modularity. These ideas together form a picture of the motor system as a structural scaffold upon which behaviorally relevant feedback mappings from cortex to the spinal cord are continuously activated and modulated based on information and estimates about the periphery. In this view, the encoded variables of interest depend on the goals, context, and perturbations of the intended movement. Graziano writes:

“The usefulness of a feedback-dependent mapping from cortex to muscles is that it can in principle allow neurons in motor cortex to control a diversity of movement variables, such as direction, speed, hand position, or posture that transcend a fixed pattern of muscle activation. If the network receives feedback information about a specific movement variable, then it can learn to control that variable.”

Muscle activity is, in a sense, a readout from a network transforming state-dependent inputs into movement goals. Rather than playing chords, the motor system is improvisational jazzmaster. The movement machine yields its complexity to construct a movement fit to purpose, to suit its context and the information it receives. Rather than choosing muscle patterns in reconfigurable blocks, it creatively constructs and sculpts movements. That is, the hierarchy of the motor system is not rigidly organized around a particular set of variables. Many loops exist connecting cortex with the spinal cord, the cerebellum, the basal ganglia, and the sensorimotor periphery. Each of these loops contributes information to flexibly activate the relevant action maps. Prevailing theory suggests that cerebellar loops provide predictive state information while basal ganglia loops provide state and/or action value information. Along these lines, recent work studying patients with cerebellar ataxia suggests that the cerebellum plays a role in the temporal recruitment of behavioral syllables, while motor cortex may be implicated in the spatial structure of synergistic action, though this study focused on 13 proximal muscles of the shoulder and arm rather than the distal muscles driving the hand<sup>38</sup>.

**I added this on Friday:** Fig. 1 depicts the hierarchical nature of the motor system that enables its dexterity. The motor system is tuned to produce varying levels of modularity, and this is shown in Rastello’s work at a structural level: CM cells evolved to provide modifications to coarse, synergistic action. This is reflected in Graziano’s work where, loosely, more dexterous behaviors are produced when stimulation is applied to the caudal regions of motor cortex. These dexterous behaviors are driven by a hierarchical stack of cellular machinery, each level of which is modulated by estimated state, goals, uncertainty, and value.

The movement machine reasons in the space of feedback control systems and their ensuing trajectories. The phenomenal thing about the motor system is that it is able to tune itself rapidly with both high-dimensional sensory inputs and sparse reward signals<sup>39,40</sup>. This has some precedence in the literature and will be discussed further in ???. This section has attempted to illustrate the complexity

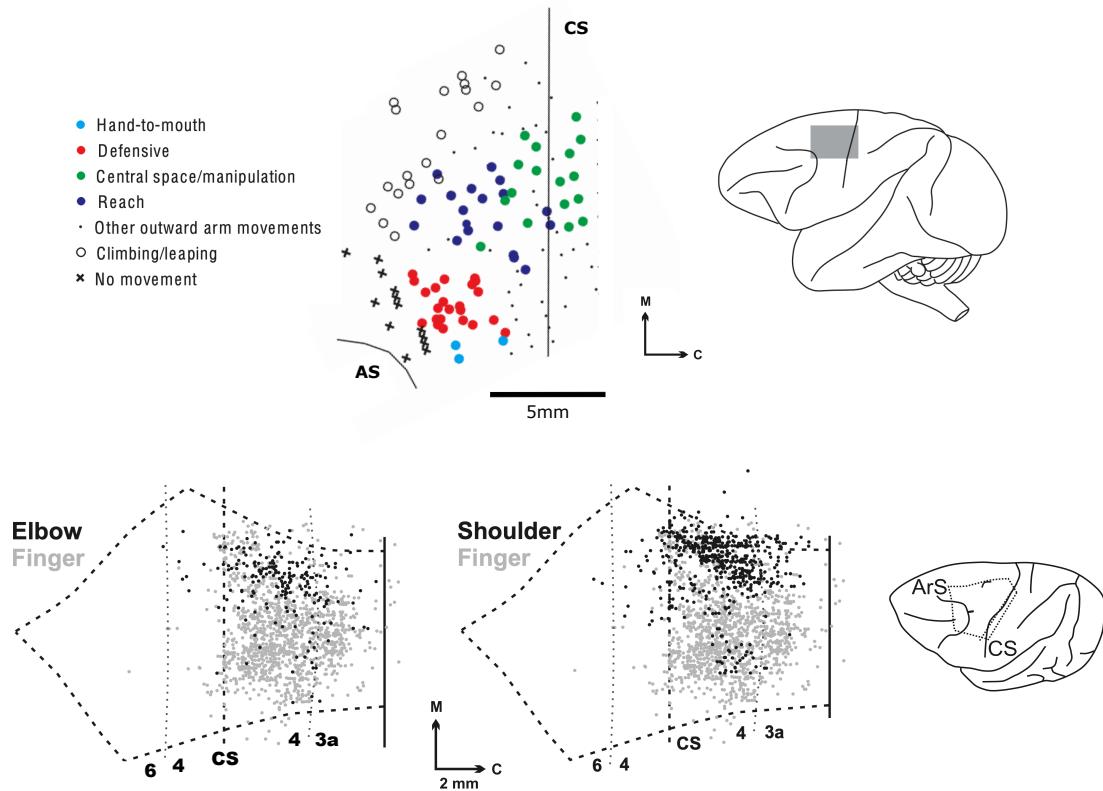


Figure 1: Similarities between electrical stimulation on behavioral timescales and rabies tracing identification of CM cells. CM cells are largely confined to the caudal half of M1, while this region tends to evoke complex manipulatory movements when electrically stimulated. (Top Left) Corticomotoneuronal (CM) cells traced using rabies from muscles of the elbow and finger. (Top Right) CM cells traced using rabies from muscles of the shoulder and finger. (Bottom) Complex movements evoked by 500ms electrical stimulation pulse trains. *Adapted from Graziano 2005 and Rathelot 2009*

of the motor control system specifically with regard to dexterous control of the hand, with an eye toward experimental and theoretical avenues for exploration. The goal is to build and test a theoretical scheme for aspects of the compositional nature of the neural hand controller.

## 2.5 Flexibilty in Spinal Circuits

Renshaw cells – synergist inhibition, maybe to synchronize synergistic activations

### 2.5.1 I added this section on Friday!\*\*

Looking at rapid visuomotor responses, this work found that these reflexive movement were modulated by the value of multiple goals, just as in cognitive tasks. This supports the idea that there exists flexibility at all “levels” of the hierarchy, all receiving similar feedbacks and all similarly modulated by context:

If low-level sensorimotor circuits can contribute to value- based decisions through continuous feedback control, rather than merely executing the outcome of discrete action decisions taken in higher-order brain areas, it would support for the hypothesis that value-based decision algorithms are distributed throughout multiple levels of sensorimotor and cognitive processing hierarchies (Hunt et al., 2014; Hunt and Hayden, 2017). This notion differs from the traditional view that decisions arise from a serial process with modular units for choice evaluation, value comparison and action selection. According to the alternative view, the basis for decisions is mutual inhibition between neural representations of alternative options, and these computations occur simultaneously in multiple brain areas along both motor and abstract-value dimensions of tasks (Wang, 2012). Our current evidence that value-based decisions can be implemented through sensorimotor feedback control supports the alternative view, and the general notion that behavior emerges via a distributed consensus between circuits engaged nominally in decision and sensorimotor processes (Cisek, 2012).<sup>41</sup>

When we move, the brain specifies a set of feedback control gains that enable low-level motor areas not only to generate efficient and accurate movement, but also to rapidly and adaptively respond to evolving sensory information in a manner consistent with value-based decision-making.<sup>41</sup>

Spinal stretch reflexes may also be modulated by posture, like in Graziano’s work:

We found that changing the arm’s orientation diametrically altered how spinal reflexes in the elbow muscles were evoked, and in such a way that were again efficiently scaled to the hand’s distance from the target. These findings demonstrate that spinal circuits can help efficiently control the hand during dynamic reaching actions, and show that efficient and flexible motor control is not exclusively dependent on processing that occurs within supraspinal regions of the nervous system.<sup>42</sup>

This is supported by another paper:

Our results reveal complex goal-dependent modulation of fast feedback responses in M1 that are present early enough to account for goal-dependent stretch responses in arm muscles<sup>43</sup>

Feedback and internal dynamics play a role, and models reflect either (in this model they forgo sensory feedback, which we see as integral to modulating feedback control):

Sensory feedback takes at least 25 ms to influence cortical responses and >50 ms to reflect the current goal. Thus, during this ~200-ms interval, the neural dynamics are not yet affected by sensory feedback and should presumably be explained via internal dynamics. This is true even of optimal feed- back control architectures, which employ a dynamically varying control policy and internal ‘efference-copy’ recurrence to generate time-varying output patterns before the arrival of feedback. Given the practical choice to use a model without sensory feedback, we verified with additional simulations that

the solutions found by the model were robust to the addition of reasonable forms of feedback<sup>44</sup>

Andy has written that sensory and spinal systems work “in parallel”, but do we agree with this?

In terrestrial mammals, the rhythmic and coordinated leg movements during locomotion are controlled by a combination of interconnected neurons in the spinal cord, referred to as the central pattern generator, and sensory feedback from the segmental somatosensory system and supraspinal centers such as the vestibular system. How segmental somatosensory and the vestibular systems work in parallel to enable terrestrial mammals to locomote in a natural environment is still relatively obscure.<sup>45</sup>

## 2.6 The Heterarchical Motor System

This distributed view is crucial; a view that sensorimotor processing is a, perhaps, “complex hierarchy”, or even a heterarchy...

4647 48

This term was first used by McCulloch to describe the way networks give rise to multiple competing values:

Circularities in preference instead of indicating inconsistencies, actually demonstrate consistency of a higher order than had been dreamed of in our philosophy. An organism possessed of this nervous system—six neurons—is sufficiently endowed to be unpredictable from any theory founded on a scale of values. It has a heterarchy of values, and is thus interneatively too rich to submit to a summum bonum.<sup>49</sup> <sup>5</sup>

He's saying that networks without a hierarchy of values, networks that inherently loop, give rise to “unpredictability”, or perhaps flexibility – implies that if the system is optimizing, there is no Supreme Good, but rather a composite Good comprised of component values.

Summum bonum is a Latin expression meaning the highest or ultimate good, which was introduced by the Roman philosopher Cicero to denote the fundamental principle on which some system of ethics is based — that is, the aim of actions, which, if consistently pursued, will lead to the best possible life. (wiki)

used in social sciences to describe power relations between groups that aren't strictly hierarchical, but exist in a more complex arrangement:

when a given production mechanism is regulated by multiple control mechanisms without these control mechanisms being themselves subsumed under a higher-level controller. To the degree one can distinguish levels of control, there may be more controllers at higher levels than at lower levels<sup>50</sup>

The addition of the term heterarchy to the vocabulary of power relations reminds us that forms of order exist that are not exclusively hierarchical and that interactive elements in complex systems need not be permanently ranked relative to one another.<sup>50</sup>

many structures, both biological and social, are not organized hierarchically. There is nothing intrinsically hierarchical about an oak tree or a symphony, yet each has undeniable structure and constitutes an orderly representation of the relations among elements. Nonetheless, few terms identify other kinds of order. Hierarchy—inasmuch as it is often a reductionist metaphor for order—has disproportionately influenced theory building in both social and natural scientific contexts.<sup>50</sup>

control hierarchy: decisions at higher levels affect the operation of lower levels.<sup>50</sup>

is this really a good definition? i suppose it's something about the agency of the decisionmaking, it's more about control – does the upstream control only the downstream?

---

<sup>5</sup>Supreme Good

in philosophy: > when a given production mechanism is regulated by multiple control mechanisms without these control mechanisms being themselves subsumed under a higher-level controller. To the degree one can distinguish levels of control, there may be more controllers at higher levels than at lower levels.<sup>51</sup>

...the formation of a voluntary movement is much more complicated. To think that a voluntary action is formed in the narrow field of the motor cortex would be a mistake similar to an assumption that all the goods exported through a terminal are produced in the terminal. The system of cortical zones participating in the creation of a voluntary movement includes a complex of subcortical and cortical zones, each playing a highly specific role in the whole functional system.

shift in vernacular can be a shift in knowledge point of science is to describe the world concretely, so our words using to do this description matter

### 3 Background Experimental Methods

- what are the most important concepts / results that inform our experiments?
- What do we think we know from experiments about motor control? motor adaptation? motor learning?

#### 3.1 Control

Is LQR (as it's claimed to be) a reasonable model for feedback control and error reduction + variability prediction for dimensionality reduction-based motor interface

(task reads out from D muscles, find modes of that data; do PCA to get  $K < D$  dimensions, controller only responds to motion in those K directions)—does behavior + motor activity follow LQR? this question has already been asked, but it hasn't been asked for this kind of high-to-low dim mapping. It's been asked in tasks where muscles haven't been directly in control (Bolero 2009). Todorov: do a task, look at muscle signal. Muscles that aren't necessary for task have higher variability b/c they're not being optimized for task (but doesn't introduce perturbations). Also see Loeb (2012) for a negative result saying that muscle coordination is habitual rather than optimal, but it has issues (low # muscles). Can we replicate previous reaching optimality results in our set-up? What's unique about our set-up is the PCA/dimensionality reduction in muscle activity space. This is important because you can create arbitrary muscle-cursor mappings, so you have to learn a new skill/mapping. This is different than perturbing a fundamental movement and forcing adaptation, which is what has been previously done. For our task, the participants actually have to learn a new task/mapping, rather than just do what they already know and be robust to perturbations. We test the LQR hypothesis once they've learned the task, because LQR isn't a learning theory, it's a theory about optimal control. We can see if, once people learn a new skill, their behavior is optimal wrt LQR theory. If we establish this, then we can think about how this LQR model is actually learned (enter RL).

Our results are consistent with a recently described model in which an optimal feedback control policy is calculated independently for each potential target and a weighted average of these policies (that is, feedback gains) is computed at each point in time based on the relative desirability of each target<sup>50</sup>. Notably, this model, which predicts averaging of feedback gains, can also account for spatial (that is, trajectory) averaging in go-before-you-know tasks. We submit that our result showing feedback gain averaging, coupled with previous work demonstrating trajectory averaging, provides strong support for the compelling idea that the CNS, under cases of target uncertainty, encodes in parallel multiple motor plans, along with their associated control policies, for competing action options. (Wolpert Nature 2018 competing control policies)

Nashed 2014 – short-latency R1 and long-latency R2 responses (60ms; 45–75ms; 75–105 ms) stretch responses R1 show dexterity (Andrew Pru 2019,2020) in holding and in reaching

### 3.1.1 Adaptation of Reaching

- prisms
- rotations
- forcefield
- nothing – van Beers variability

The vast majority of research in motor learning studies this capacity through adaptation para- digms in which a systematic perturbation is introduced to disrupt a well-practiced behavior, such as point-to-point reaching.<sup>52</sup>

- classic reaching adaptation -> this is a different goal
  - shadmehr
  - krakauer
- unperturbed movements
  - van beers

53,54

There exist a handful of prior studies mapping EMG activity and finger joint angles directly to virtual stimuli, though few are focused on the learning process and none have the input dimensionality we aim to achieve in work proposed here.

57

### 3.1.2 Arbitrary Visuomotor Mappings

58

There are several studies using non-EMG-driven sensorimotor mappings to study human motor control and learning.

- Remapping Hand Movements in a Novel Geometrical Environment <https://www.ncbi.nlm.nih.gov/pubmed/1614827>

59

vocoder machine bell labs

Hinton, Fells

palsy study

takehome: humans are really good at learning tasks like these, especially with their hands. this type of dexterity is specific to primates if not humans. let's use this ability to understand and try to model how this learning process unfolds.

*What does this give us that a force-field reaching task can't?*

60

### 3.1.3 Skill Learning Tasks

- skill learning tasks
  - ball and cup
  - dart throwing tasks

### 3.1.4 Learning in Cortical Interfaces

- cortical BMI work
  - Batista papers, lee miller papers
- speech learning – analogy to speech
- bird vocal learning

- we're doing the same experiment, at the muscle level
- try to convince why this is useful, but not too hard

### 3.2 Properties of Electromyographic Signals

- not actually gaussian, but super-gaussian<sup>61</sup>
- some work using bayesian filtering methods to infer the signal envelope<sup>62</sup>
- Farina paper on EMG as convolution

A window of EMG of length  $T$  samples can be modeled like

$$\mathbf{z} = \sum_t^T \mathbf{h} * \mathbf{s}$$

where  $\mathbf{h}$  is a motor unit activation template, which itself is a particular neural spike waveform, and  $\mathbf{s}$  is the incidence of a spike, which might be modeled as a point process.

EMG records were rectified, smoothed and averaged before further analysis.<sup>63</sup>

EMG activity was recorded using hook-wire electrodes (44 gauge with a 27 gauge cannula; Nicolet Biomedical, Madison, WI) placed in the muscle for the duration of single recording sessions. [...] Electrode voltages were amplified, bandpass filtered (150–500 Hz, four pole, 24 db/octave), sampled at 1000 Hz, and digitized. Off-line, raw traces were differentiated (to remove any remaining baseline), rectified, smoothed with a Gaussian (SD of 15 ms) and averaged.<sup>64</sup>

inter-subject variability is high, but seems to show individual strategies for movement<sup>65</sup>

The surface EMG signal can be modeled as the convolution of

Can muscle coordination be precisely studied by surface electromyography?<sup>66</sup>

This requires first mapping the intrinsic available dynamics of the hand per user.

#### 3.2.1 Skill Learning in Myoelectric Interfaces

“performance levels and rates of improvement were significantly higher for intrinsic hand muscles relative to muscles of the forearm.”<sup>67</sup>

#### 3.2.2 Berger et al. 2013

<sup>68</sup>

Berger et al.’s preprocessing - lowpassed butterworth at 5Hz - normalized to MVC calibration - periodic baseline noise subtraction - choosing synergies by uniformity in force direction (pretty arbitrary)

Using EMG in a learning experiment is not unheard of. Berger et al. 2013 use EMG with 13 muscles to test whether learning new synergy combinations for a task is more difficult than recombining existing synergies@Berger2013a. As we would expect, learning new synergy combinations is more difficult. I would argue that the demand in their “incompatible virtual surgeries” is too strict, that we need to more carefully design synergy perturbations to develop a model of learning in such a task.

Berger et al. fit a muscle-space to force-space mapping  $H$  using a force-driven calibration task, and a synergy-space to muscle-space mapping  $W$  using NMF.

$$f = Hm$$

$$m = Wc$$

$\dim(m) = M$  muscles,  $\dim(c) = N$  synergies, and  $\dim(f) = D$  dimensions of task space where  $M > N > D$ . Because  $H$  and  $W$  are rectangular, they have at most rank  $D$  and  $N$ , and we constrain these matrices to be full rank. There are three key subspaces: the nullspace of  $H$  mapping muscle activations to 0, the column space or range of  $W$  mapping synergy activations to muscle activations, and the common subspace between these two. That is, there are synergy activations which generate muscle activations which lie in the null space of  $H$ . The paper uses this fact to develop mappings that specifically rotate muscle activations produced by synergies into the null space of  $H$  which were not there prior to rotation. The dimensionalities of these subspaces are defined:

$$\begin{array}{ll} \dim(\text{null}(H)) = M - D & \text{muscle vectors } \rightarrow 0 \\ \dim(\text{col}(W)) = N & \text{synergy activations } \rightarrow \text{muscle subspace} \\ \dim(\text{null}(H) \cap \text{col}(W)) = N - D & \text{synergy activations } \rightarrow 0 \end{array}$$

In the paper, the authors find an orthonormal basis  $W_o$  for the range (column space) of the synergy weight matrix  $W$  (presumably using a QR factorization) and find the nullspace  $H_{\text{null}}$  of  $H$ . These computations are done presumably through QR factorizations (an orthonormal basis multiplied by a rotation and scaling) by finding  $Q$  in the first case and finding the latter  $M - D$  columns  $Q_2$  of  $Q = [Q_1 Q_2]$  which are  $H_{\text{null}}$  in the second case:

$$\begin{aligned} W &= Q_W^{M \times M} R_W^{M \times N} \\ &= \left[ Q_{W,1}^{M \times N} Q_{W,2}^{M \times M-N} \right] \begin{bmatrix} R_{W,1} \\ 0 \end{bmatrix} \\ W_o &= Q_{W,1}^{M \times N} \\ W_o^T &= Q_{W,1}^{T, N \times M} \\ H_{M \times D}^T &= Q_H^{M \times M} R_H = \left[ Q_{H,1}^{M \times D} Q_{H,2}^{M \times M-D} \right] \begin{bmatrix} R_{H,1} \\ 0 \end{bmatrix} \\ H_{\text{null}}^T &= Q_{H,2}^{M \times M-D} \end{aligned}$$

To find each of the three subspaces, we take the SVD of the composition  $W^T H^T$

$$\begin{aligned} W_o^T H_{\text{null}}^T &= Q_{W,1}^{T, N \times M} Q_{H,2}^{M \times M-D} \\ &= Q_W^T Q_H \\ &= U_{N \times N} \Sigma_{N \times M-D} V_{M-D \times M-D}^T \\ H_{\text{null}}^T V &= W_o U \Sigma \end{aligned}$$

Now we can pick out the three subspaces using the SVD

$$\begin{array}{ll} W_c = W_o U[1 : N - D] & \text{synergy activations } \rightarrow \text{muscle activations in task null space} \\ H_c = H_{\text{null}}[1 : N - D] & \text{synergetic muscle activations } \rightarrow 0 \\ W_{nc} = W_o U[N - D + 1 : N] & \text{synergy activations } \rightarrow \text{nonzero muscle activations} \\ H_{nc} = H_{\text{null}} V[N - D + 1 : M - D] & \text{non-synergetic muscle activations } \rightarrow 0 \end{array}$$

To construct new mappings, the authors construct rotations to alter muscle activation vectors by rotating them from  $W_n c$  and remaining in  $W_n c$  and from  $W_n c$  into  $H_n c$ . In the first case this alters the mapping by changing the effective muscle activations produced by the existing (learned) synergistic actions. That is, muscle activations putatively produced by synergistic action will be altered to produce different forces in task space (compatible rotations). In the second case, muscle activations putatively produced by existing synergistic action (via  $W$ ) will be mapped into the null space of  $H$  and produce zero force in task space (incompatible rotations).

A key critique of this paper is that such a transformation is too harsh. The compatible rotation allows you to recombine the same muscle patterns, the incompatible doesn't allow you to use existing coactivation patterns at all. The authors do see new synergies emerging even after their training session, consisting of:

- 16 trials of maximum voluntary contraction in 8 directions (calibration)
- 72 trials using force control (calibration)
- 24 trials familiarization
- 144 trials baseline
- 288 trials surgery
- 144 trials washout
- 144 trials baseline

After 288 trials subjects aren't able to complete the task for some movement directions.

### **3.2.3 Nazarpour 2012 J.Neuro**

60

x Flexible Cortical Control of Task-Specific Muscle Synergies <https://www.jneurosci.org/content/32/36/12349>

Fig. 4A – cursor controlled muscles begin to dissociate from non cursor controlled muscles.

Feedforward processing to muscle fields / tunings in the presence of signal dependent noise

Feedback processing based on visual errors

### **3.2.4 Radhakrishnan 2008**

69

x Learning a Novel Myoelectric-Controlled Interface Task — Radhakrishnan, 2008 <https://www.ncbi.nlm.nih.gov/pmc/ar>

proprioception is not required to learn nonintuitive MCI mappings

several hundred trials subjects learned pointing with six muscles

prism adaption requires active movement; efference copy implicated if proprioception doesn't seem to be required

control models Fig 10

### **3.2.5 de Rugy 2012 - Habitual not Optimal**

70

just because it's harder to adapt to incompatible surgeries doesn't mean that there are fixed synergies, it just means there are multiple timescales of adaptation available in the neural control hierarchy – diversity in the neural controller depending on context

learning inverse model may be separate from learning to optimize trajectories on top of that model – some type of “fine tuning”

skill acquisition (slow, constructing novel synergies) vs. motor adaptation (less slow, adapting existing synergy activations)

It's a good test, but it pushes the optimal control framework too hard? perhaps we need a model for what "good enough" is? If we penalize moving to a new controller from a previously optimized movement, the findings make sense. An optimal control model would predict the exactly optimal coordination patterns for the new scenario, it wouldn't say anything about adaptation from an old solution to a new one. This is why we need to develop a model of adaptation that formalizes this scenario not of kinematic perturbations (noise during movement), but to a drastic change in the plant itself (e.g. muscle failure).

### 3.2.6 Mussa-Ivaldi 2019

Earlier theoretical work by Jordan and Rumelhart [14] considered how the learning of actions can be viewed as the concurrent learning of forward and inverse models of actions. They introduced the concept of distal learning, where the learner has to find a mapping from desired outcomes to actions in order to achieve a desired outcome. To do so, the subject begins by forming a predictive forward model of the transformation from actions to distal outcomes. Such transformations are often not known a priori, thus the forward model must generally be learned by exploring the outcomes associated with particular choices of action. Once the forward model has been at least partially learned, it can be used to guide the learning of an inverse model that predicts the action needed to achieve the distal outcome. Mussa-Ivaldi2019

Our findings are consistent with the hypothesis that learning proceeds through the concurrent evolution of coupled forward and inverse models of the body-to-object mapping established by the BoMI. Mussa-Ivaldi2019

Not being square, the matrix  $H$  does not have a unique inverse. But there exist infinite "right inverses" that combined with  $H$  yield the  $K \times K$  identity matrix in the task space of external control signals. Each such right inverse transforms a desired position of the controlled object into one particular set of values for the body signals. We consider users to be competent when they are able to move their body successfully in response to a presented target for the controlled object. Mathematically, we consider this as finding one right inverse  $G$  of the mapping  $H$ , out of a multitude of possible and equally valid choices. Mussa-Ivaldi2019

Gradient learning of a forward and inverse model (mapping):

$$\begin{aligned}\hat{H}_{n+1} &= \hat{H}_n + \epsilon(p_n - H_n q_n) q_n^T \\ G_{n+1} &= G_n - \eta \hat{H}_n^T e_n u_n^T \\ e_n &= p_n - u_n\end{aligned}$$

The comparison between model predictions and actual data in Fig 3 indicates that our proposed model of learning is sufficient to explain the data. However, the mechanism we propose is not necessary; we cannot rule out other possibilities, such as reinforcement learning. [...] This agreement between model and experimental results does not exclude the possibility of alternative learning mechanisms, such as a direct learning of the inverse model [24] or the use of reinforcement learning [25] to acquire an action policy that would play the role of the inverse model. Mussa-Ivaldi2019

How do we break a simple gradient model? On a task that is more difficult? will learning take longer? - savings phenomenon -

Although the interface forward map is linear (Methods, Eq (5)), this is a many-to-one map admitting a multitude of inverses. This "redundancy" opens the possibility of successful linear and nonlinear inverse maps. Redundancy also leads to an important consideration about gradient descent learning. The reaching error surface in the space of the inverse model elements does not have a unique minimum, but a continuously

connected set of minima corresponding to the null space of the forward map. In the metaphor of a skier descending from a mountain following the gradient, this space of equivalent inverse models corresponds to a flat elongated valley at the bottom of the mountain. Anywhere along the valley is a valid end to the ride, as it corresponds to a valid inverse model. The inverse model on which the steepest descent ends depends on the initial conditions, as predicted by the dynamical model (see Fig 3b—evolution of the norm of the inverse model error), as well as on the realization of the noise employed in any given simulation of the learning model.

Although the two-dimensional subspace formed by the first two PCs captured a large fraction of the total variance of body motions, it did not necessarily reflect the natural up-down/left-right orientation of the display monitor. Therefore, following calibration and PC extraction, there was a customization phase in which users were allowed to set the origin, orientation, and scaling of the coordinates in task space, based on their preference.

Subjects have prior knowledge of their directions in task space?

x 90% isn't enough – Follow-up on the previous paper – critiques “direct evidence”  
<https://www.biorxiv.org/content/10.1101/634758v1>

- Structured variability of muscle activations supports the minimal intervention principle of motor control <https://www.ncbi.nlm.nih.gov/pubmed/19369362>

## 4 Experimental Contributions

- shape of the data
- structure of variability in the EMG compared to the structure of variability in the behavior?

we know how to design and interpret experiments that involve many repetitions of the same movement however there is limited role for online optimization in that context. instead we need experiments where subjects are required to come up with new movements all the time. how can we get experimenters to do such experiments? show cool movies of robots doing cool things, and hopefully get the experimenters excited. (todorov online optimization slides)

### 4.1 Experimental Setup

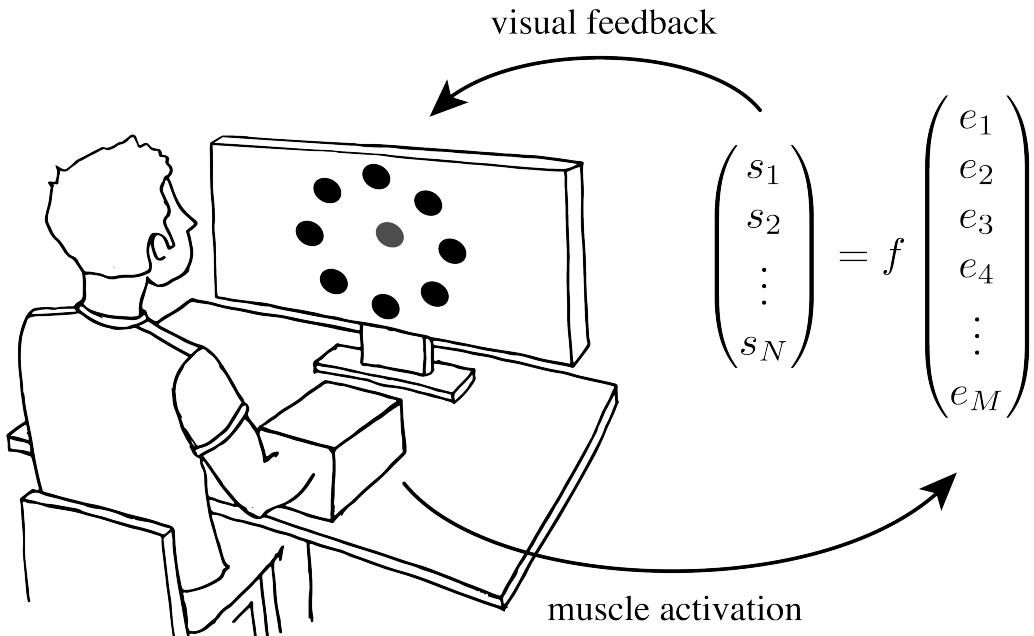
My setup records EMG from the forearm.

The goal of the project’s first phase is to develop a high-dimensional surface EMG recording rig to generate datasets with high signal-to-noise ratio and dense coverage over superficial muscles of the arm and hand. The first question of this phase is: what are the limitations of a closed-loop myocontrol experiment, and how can such constraints be avoided or leveraged? To answer this question we will develop a signal processing pipeline and diagnostics suite to identify constraints in the setup and aim to overcome, as much as possible, the limitations inherent in surface EMG recording such as muscle crosstalk and rigorous electrode placement.

The concept of the experimental setup is shown in Figure 1, where 64 monopolar electrodes are attached to a subject’s arm and hand to record muscle activity. The arm and hand are kinematically constrained in a custom fixture and motor activity is recorded during isometric muscle contractions at levels less than 20% maximum voluntary contraction to lessen the risk of involuntary co-contractions. The setup circumvents the limb biomechanics by mapping muscle output directly to virtual stimuli shown on a computer monitor. Additionally, our study focuses on low-force, isometric contractions to avoid complications due to artifacts in dynamic, high- force movements.

We chose 64 channels in order to have at least two electrodes per muscle implicated in control of the hand in the event that we require differential recording. This choice limits our analysis to the motor pool level. If our questions require recording at the motor unit level, we will need to move

to a higher channel count system. Literature in this field typically use a much lower number of channels. We believe that using 64 electrodes will help develop a more complete picture of the superficial muscle activity of the arm and hand across learning. A diagram of muscles relevant to the control of the hand and wrist is shown in Figure 2 on page 4. We are not aware of a rigorous study testing which muscles of the arm and hand can be accurately captured using surface EMG.



Sketch of the experimental concept

We aim to extend this prior work using learning algorithms that take into account time-varying dynamics of the signal in addition to common tools like components analysis and matrix factorization. This analysis will help generate an understanding of intersubject, intersession, and intertask variability. Both an analysis of dynamic correlations and a validation of dimensionality using EMG would be a novel contribution to the literature.

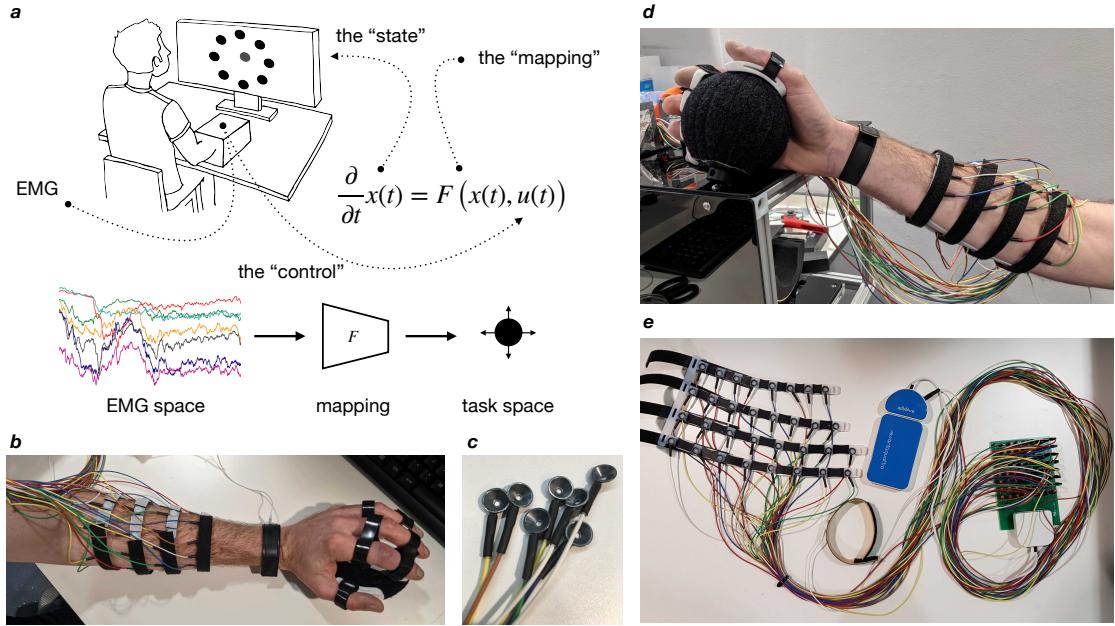
We anticipate that quantifying electrode placement and calibrating across sessions will be a major challenge. We aim to develop a mechanical fixture for recording as well as alignment tools to aid in placing electrodes in precisely the same location each session. Properly separating variability due to electrode placement from behavioral and physiological variability will be paramount to establish repeatability in our results. Once we have collected a naturalistic activity dataset, we can begin to design bespoke feedback mappings and perturbations, as discussed in Section .

Goal here is to use the linear dynamics environment to isolate the control strategies of the CNS under these constraints—how does the CNS adapt to this environment? How does it construct solutions to control problems of various dimensionalities? How does it produce dexterous responses to perturbations of these solutions?

#### 4.1.1 Hardware

#### 4.1.2 etc

what experiments do we need to do? experimental setup i made a thing, it works like this, here's the data - detail how this works - what are the constraints? - what perturbations can be achieved?  
 - prelim data from the rig - figures of this data - thoughts about how versions of the task - hardware  
 - recording 64 channels of EMG from multiple muscles the arm and hand with realtime feedback - in an isometric learning task - software - pictures n stuff



Prototype of recording hardware. Monopolar recording with reference electrode at the wrist.

## 4.2 Feature Extraction

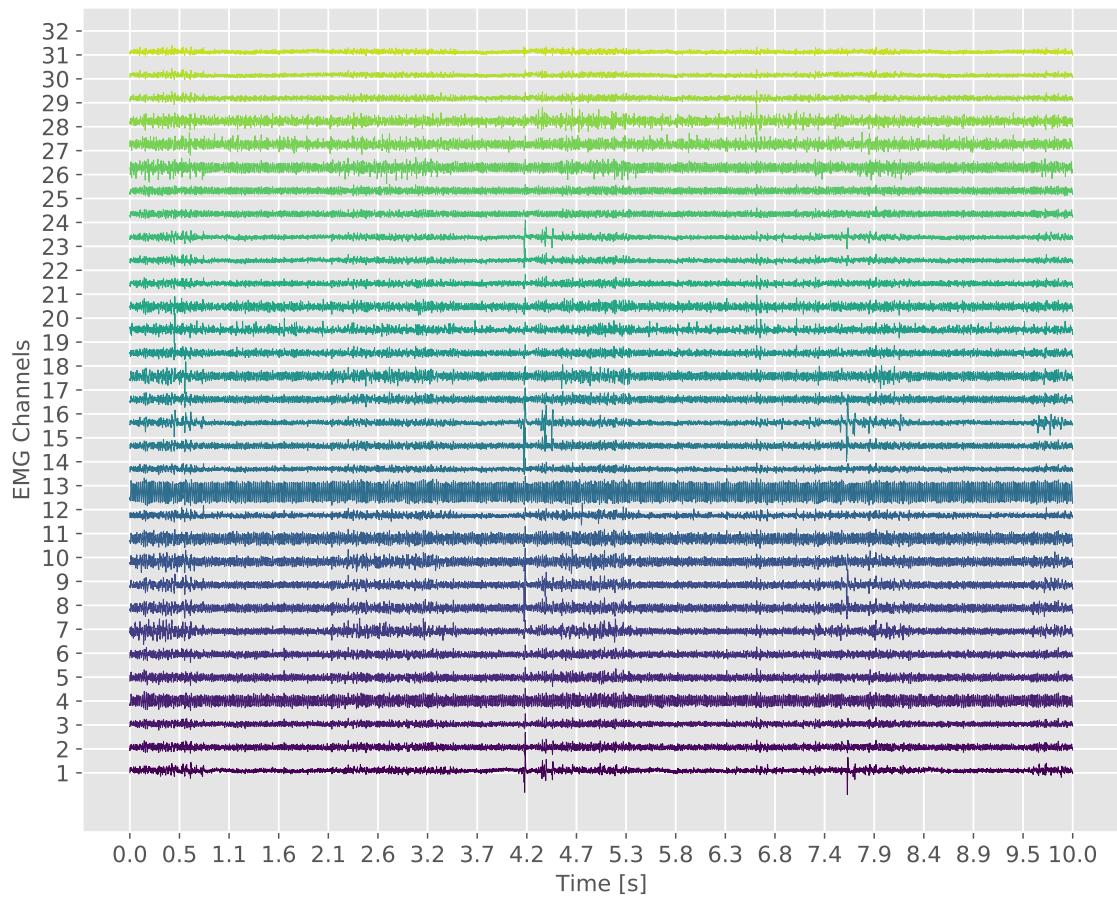
We want features that are: - smooth; having low spatial frequency - equal in the variance that they capture; equally likely to exist in - future data - Perhaps use CCA to find N features that are maximally dissimilar? - Use ICA to find independent features?

The second question of this phase is: what is the manifold of activity in electrode space during natural hand use? To answer this question, we will record naturalistic activity by subjects completing a set protocol that covers the naturalistic space of electrode covariance. For comparison, we will record a dataset of naturalistic tasks using a separate, mobile setup with the same electrode placement pattern but without the isometric constraint. These datasets could be collected from a range subjects going throughout their daily tasks, or using a specific set of tasks in the laboratory such as handwriting and the use of various tools. Encouragingly, a recent review noted that “Similarly to the breakthroughs in understanding vision that followed the quantification of statistics of natural scenes, a clear description of the statistics of natural tasks might revolutionize our understanding of the neural basis of high-level learning and decision-making”[18].

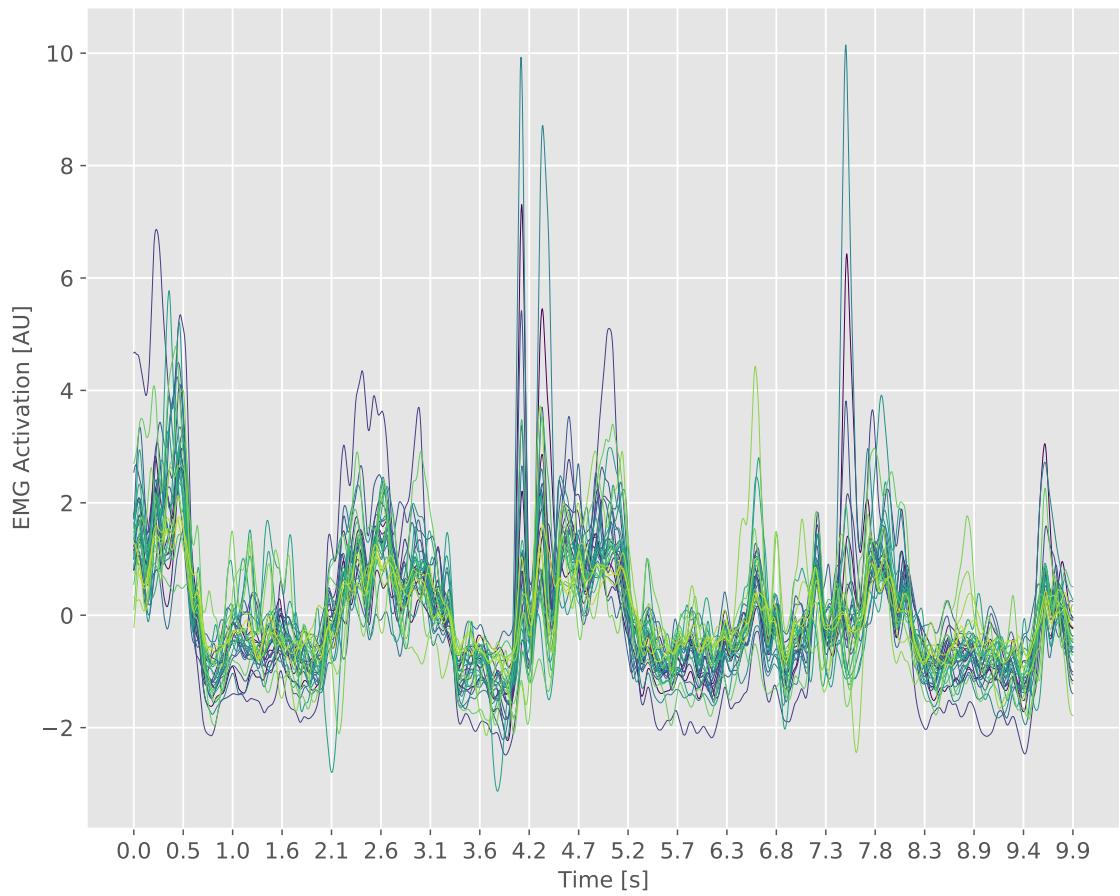
By analyzing the structure of these naturalistic datasets, we can compute the dimensionality of naturalistic movement as a subspace within our electrode space, similar to work done using joint angles of the hand[24, 22, 11]. From this work we know that while the hand has 29 joints and is controlled by 34 muscles, the dimensionality of natural hand movements is closer to 8 in joint angle dimension space based on principle components analysis. This analysis will also help us determine the biomechanical constraints on hand output dimensionality. We hypothesize that this will be higher than 8 and lower than 23, which gives us a large task space to work with for generating learning tasks.

We want to determine a redundant control space from data taken during natural activity. The difficulty with this is that such a natural activity manifold may display spatial (channel-wise) correlations that are possibly physiologically separable. Thus, there are two aims which must be addressed separately:

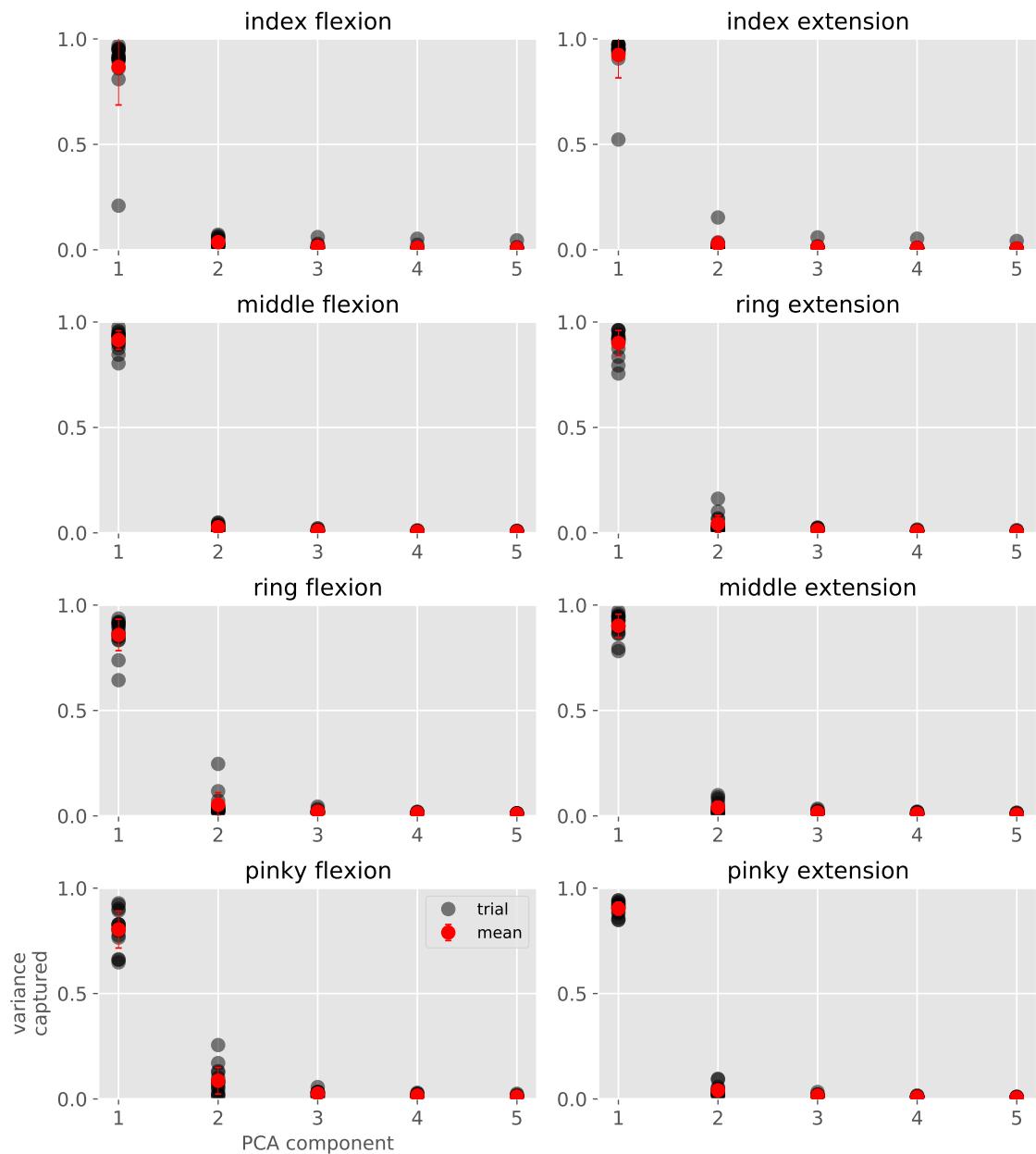
1. Explore subjects’ ability to decorrelate descending output to the muscles which have been shown to be correlated in a natural activity dataset.
  - Such a structured exploration might provide support for the hypothesis that “synergies” are flexible correlations between muscles driven by task demands rather than (or in



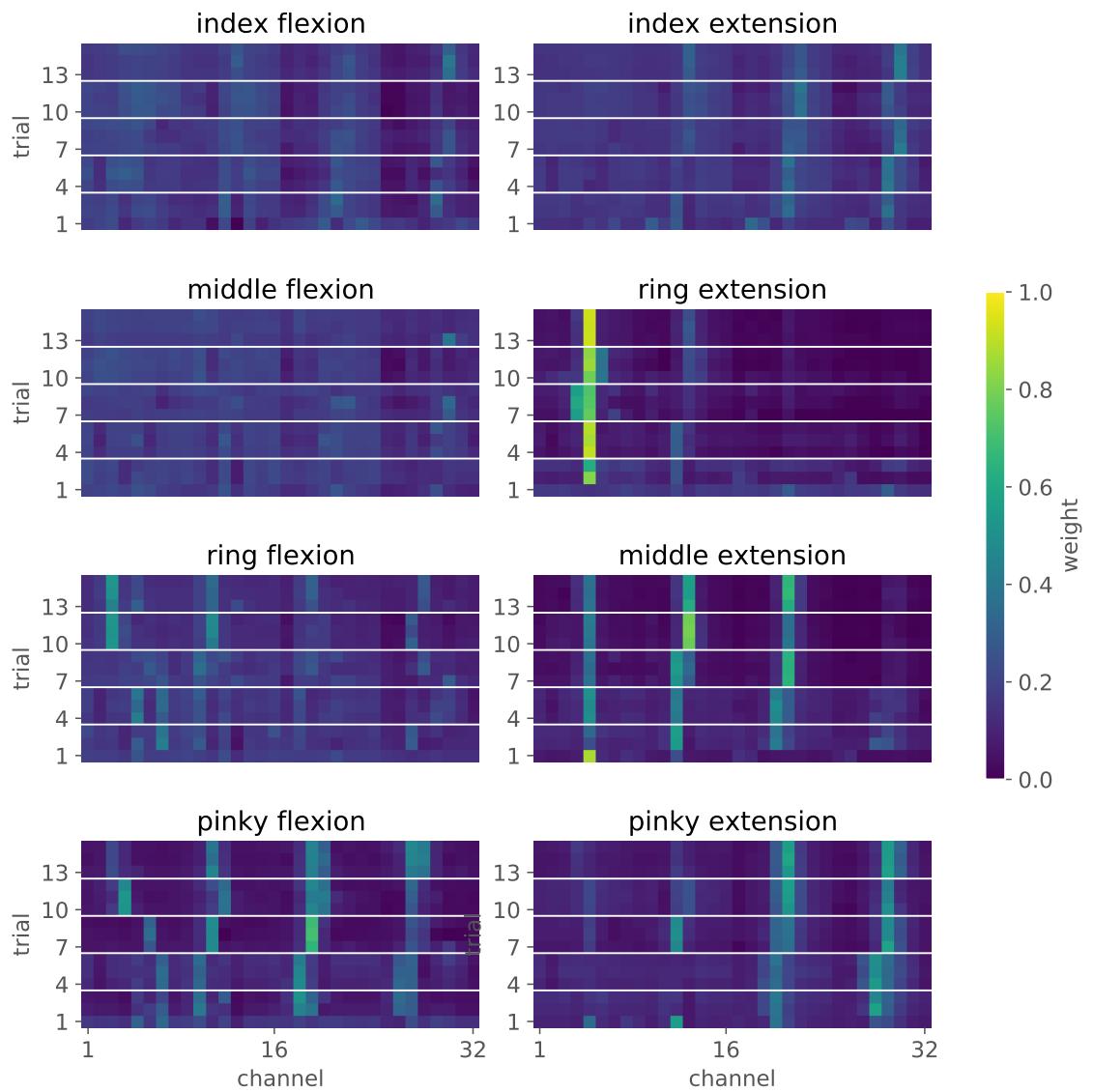
Raw EMG data from a minimal finger flexion before preprocessing.



Raw EMG data from a minimal finger flexion before preprocessing.



Raw EMG data from a minimal finger flexion before preprocessing.



Raw EMG data from a minimal finger flexion before preprocessing.

addition to) physiological structure. This needs to be done incredibly carefully to escape criticism of hard-wired synergy enthusiasts.

- See *de Ruyg 2012* for a critique of OFC and hard-wired synergies
2. Use common correlated outputs to develop a family of BMI-type learning tasks as a proxy for a “novel skill”, then track motor planning of this new skill to compare with motor planning algorithms.
    - We might be able to get #1 for free by going after this goal if we’re careful in the setup
    - This is arguably a more impactful focus as it connects low-level motor hierarchy data (EMG) to high-level planning with a normative hypothesis.

Electrode data from a single trial of a single session is held in a data matrix  $X$  (n\_electrodes, n\_samples), and we wish to find a latent weight matrix  $W$  (n\_electrodes, n\_components) which reconstructs  $X$  by projecting latent trajectories  $H$  (n\_components, n\_samples) into electrode space:

$$X = W \cdot H$$

$H$  is the activity of the latent processes, and  $W$  is there mixing matrix. The columns of  $W$  are the principal vectors spanning the latent subspace in electrode space. If we have new samples, we can project these new points onto this subspace:

$$h_{new} = W^T \cdot w_{new}$$

To justify this decomposition, we have to make some assumptions about the nature of the EMG signal, namely that the signal is linear instantaneous (each EMG sample can be instantly mapped to control space). The other assumption is that the basis  $W$  should be orthonormal, that the columns of  $W$  are orthogonal with unity norm. This ensures that the left inverse  $W^{-1}$  is equal to the transpose  $W^T$  such that:

$$\begin{aligned} X &= W \cdot H \\ W^{-1} \cdot X &= H \\ W^T \cdot X &= H \end{aligned}$$

See *Muceli 2014* for use of the Moore-Penrose pseudoinverse in place of the transpose when the columns of  $W$  do not form an orthonormal basis. This would be the case for NMF. Is there a factorization that produces nonnegative, orthogonal coordinates? Or is the pseudoinverse okay? I will need to test this.

Stated in an information theoretic way, we want to minimize the reconstruction loss  $\mathcal{L}$  for our derived encoder-decoder pair  $(E,D)$ . We’re decoding high dimensional activity into its latent dimensions, and encoding back into the high dimensional space. :

$$\min_{E,D} \mathcal{L}[X - EDX]$$

This way, forget about orthonormality and solve for an encoder and decoder directly. That is,  $E \neq D$  is perfectly acceptable.

Each row of  $D$  might be called a **spatial filter**, a linear combination of electrode activities into a surrogate, hopefully more intuitive space.

In general to find such a basis we must :

- Extract “natural activity manifold” from freeform data
- Use features of this natural subspace to derive control mapping

- Linear iid features:
  - \* PCA
  - \* dPCA
  - \* NMF
  - \* ICA
- Linear time-dependent features:
  - \* SSA
  - \* LDS model / PGM
- Nonlinear
  - \* autoencoders
  - \* networks

The behaviors present in our calibration dataset are crucial, as they determine the spatial correlations used to generate the mapping. If only complex, multi-muscle movements are present in the calibration, it will be impossible to decode subtle movements involving few muscles. Additionally, because extraction is unsupervised, it will be impossible to know how to alter the control basis directions (if we wish to do so) such that they involve single muscles or the smallest sets of muscles.

Ultimately, we want to find reproducible features in our data that are due to muscle coordination alone, rather than volitional movements. We want the lowest level covariance that reflects physiology rather than a task-level behavioral description (see *Todorov, Ghahramani 2005* and *Ingram, Wolpert 2009*). The idea is that if we collect data from enough tasks, we can extract the common modes of muscle activity. This is true only if we are sampling uniformly from the space of tasks. Otherwise one task, and therefore one coordination pattern, will be overrepresented.

### 4.3 Center Hold, Reach Out Task

In this task, the muscles of a subject’s arm are recorded using 32 channels of surface EMG. This EMG vector is mapped to a 2D force acting on a point mass shown on the screen. The mapping  $M \in \mathbb{R}^{2 \times 32}$  maps 8 “columns” each consisting of 4 electrodes placed in a line down the length of the forearm each to one of 2D root of unity. Each column of electrodes is thus mapped to one of 8 two-dimensional forces. Additionally, the point mass has zero inertia and zero friction and as such displays a direct, though redundant, readout of the EMG signal. The task asks the subject to reach one of 32 equally spaced targets on each trial.

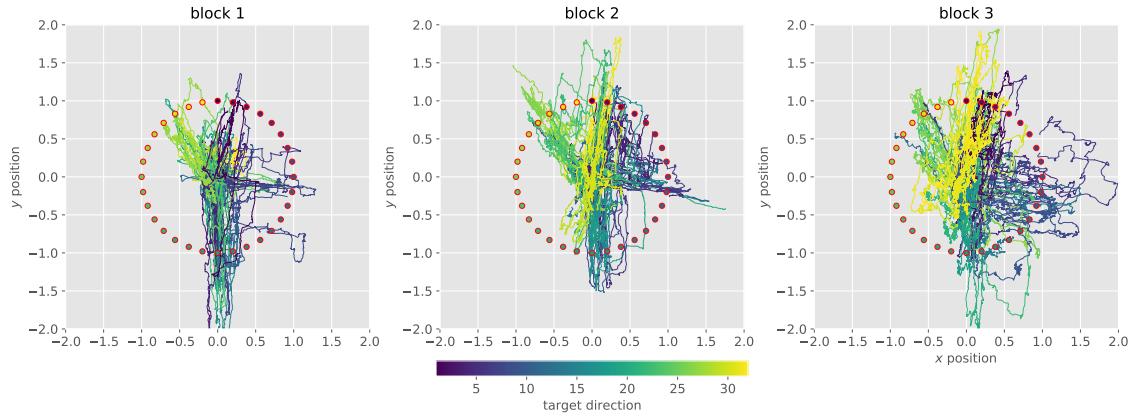
While there are 8 possible force vectors the subject can modulate by controlled the electrode activity on each of her 8 columns, the EMG mapping is ultimately a projection onto the 2D plane. Since the EMG signal is nonnegative, the subject could technically modulate just four modes of electrode activity, the minimum number needed to span the task space, to reach all 32 targets.

We can model the subject as selecting an EMG signal  $x$  which minimizes the distance between a target position  $b$  and the projection of the EMG signal through the mapping  $M$  as well as minimizes the norm of  $x$  in order to conserve metabolic energy. This optimization can be written as a regularized least squares problem:

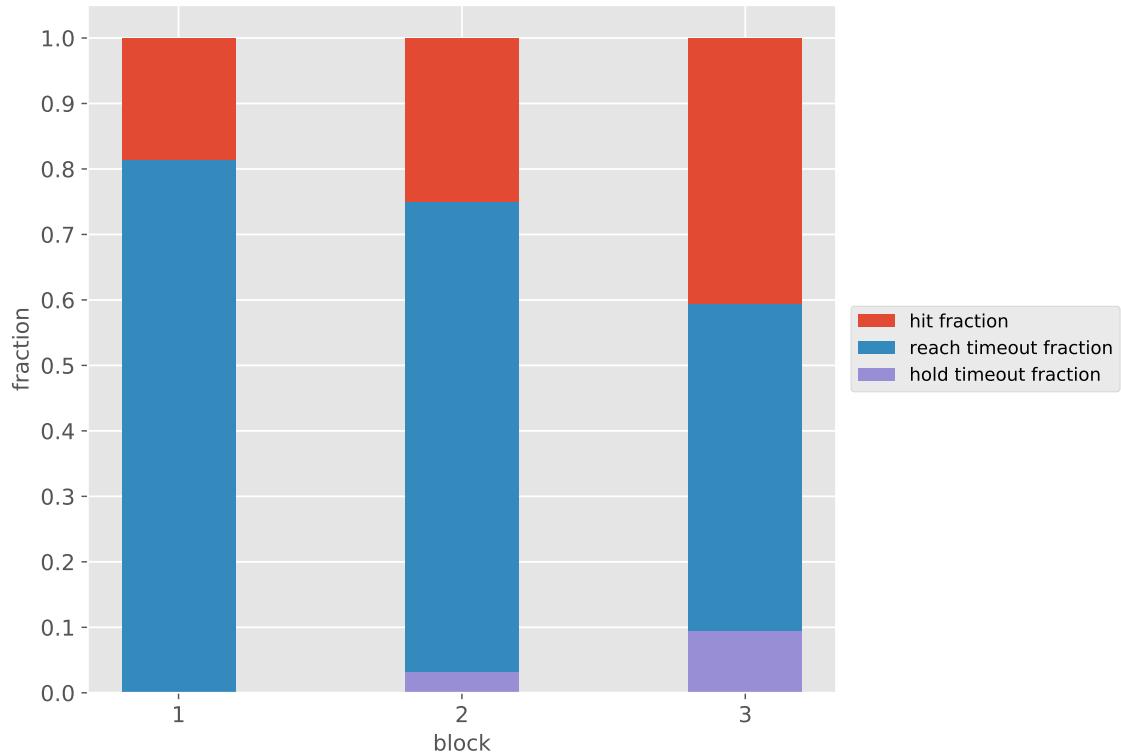
$$\min_x \frac{1}{2} \|Mx - b\|_2^2 + \frac{\lambda}{2} \|x\|_2^2.$$

This problem is known to have a unique minimum for  $\lambda > 0$  which is an approximation  $Mx \approx b$  regardless of the shape or rank of  $M$ . This implies that the subject, if they are biophysically capable to do so, will learn distinct motor outputs for each target rather than reusing modes for multiple targets with different activation levels. That is the subject will, over time, learn to fractionate their muscle output to reach their goal in order to minimize effort. For instance, to reach the the target at position  $(1, 0)$  in Cartesian coordinates, the subject could activate a bespoke activity mode or activity the combination of two modes for targets at  $\pm 45^\circ$  from this central target. If this is the case, the model predicts that the dimensionality of the EMG signal will increase over the course of training as the subject learns to construct bespoke activity modes for each of the eight targets.

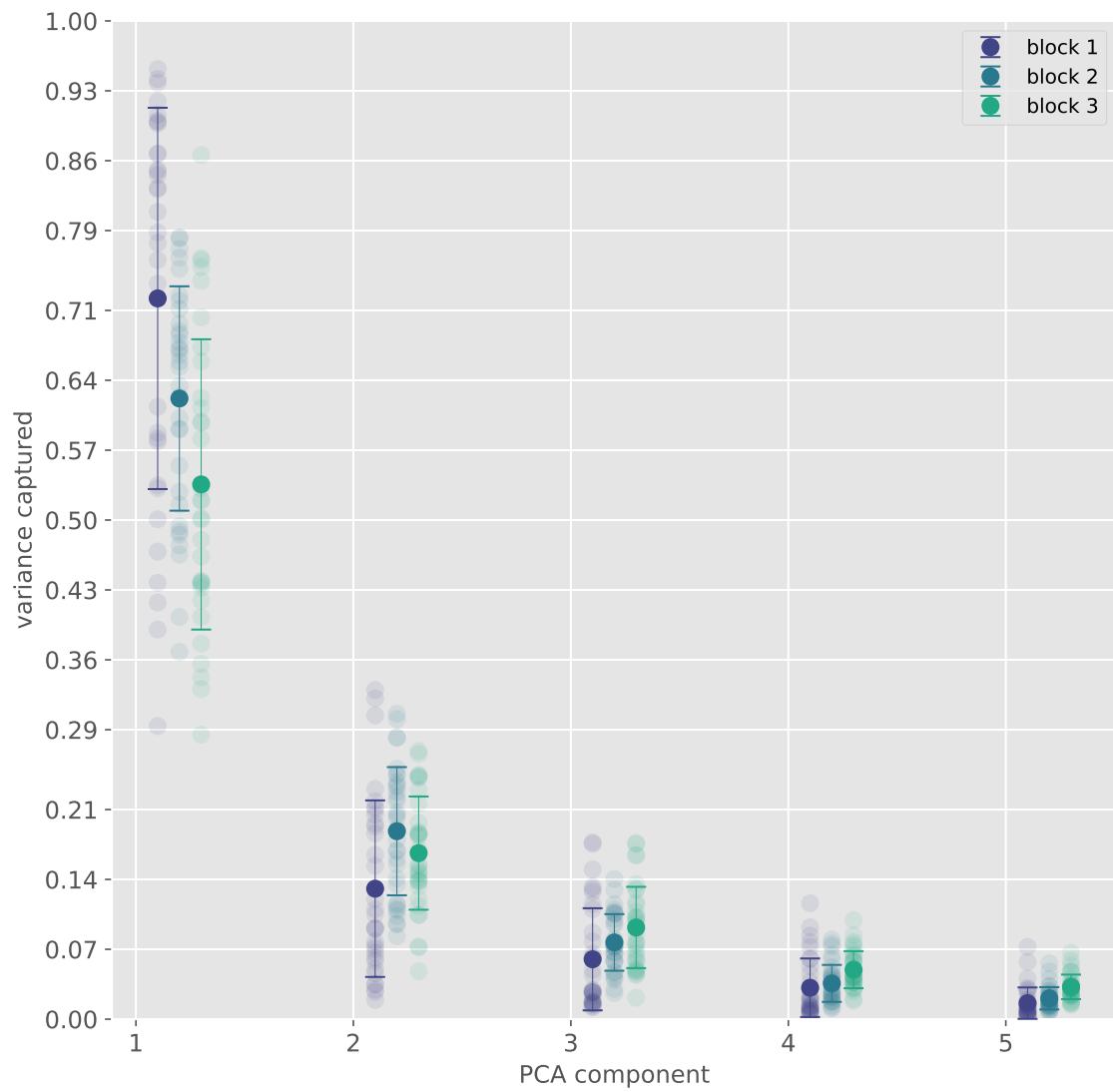
Preliminary data for this task, through the mapping:



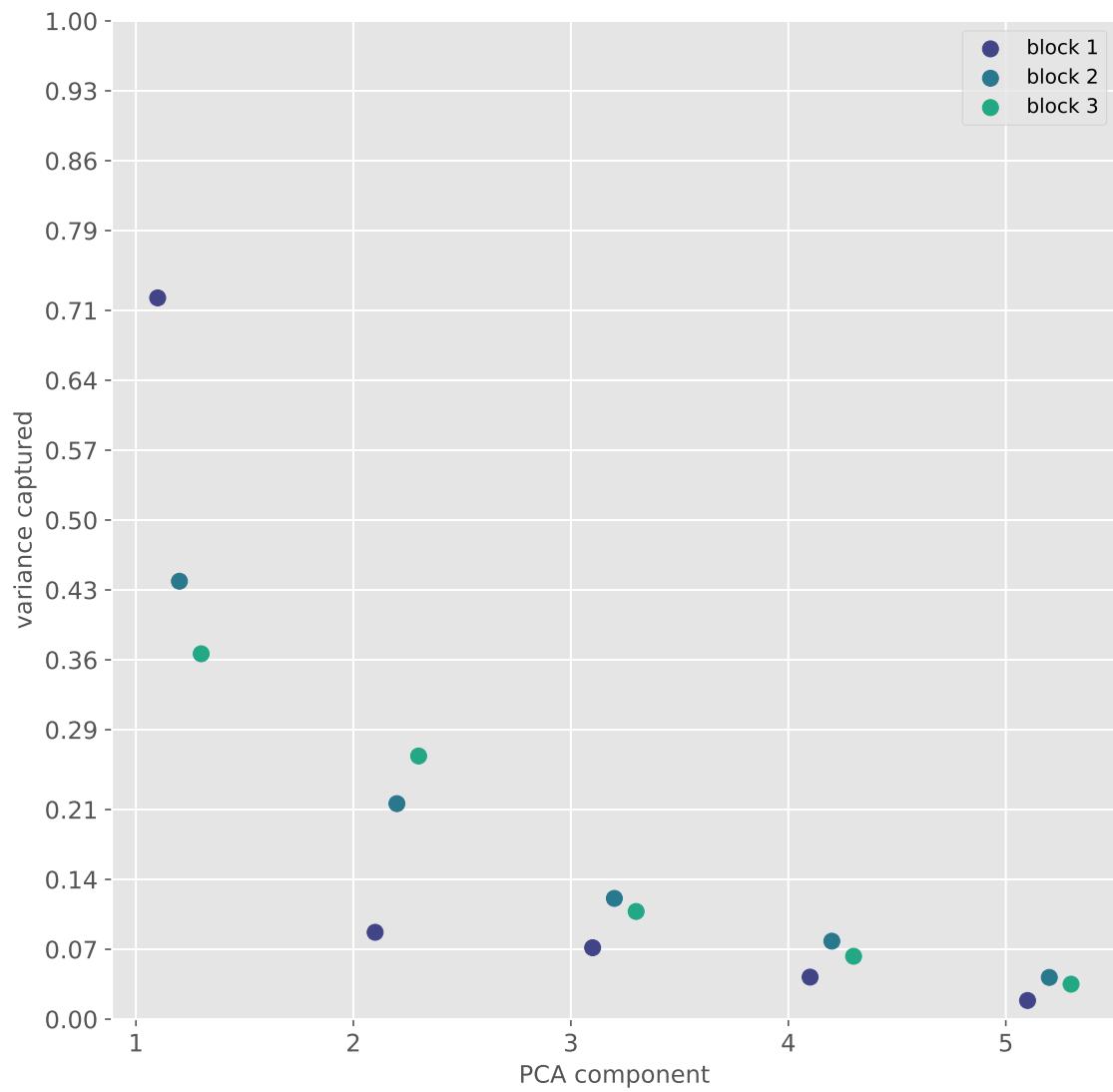
Point mass position trajectories in two-dimensional task space during the center-hold, reach-out task with 8 targets spaced evenly around the unit circle. Training was conducted over 3 blocks each with 32 trials, 4 trials per target. The first block shows roughly four modes, the second block shows four modes more clearly, and the third block may show the beginnings of fractionation.



Raw EMG data from a minimal finger flexion before preprocessing.



Raw EMG data from a minimal finger flexion before preprocessing.



Raw EMG data from a minimal finger flexion before preprocessing.

$$\tilde{M} = [M \ M \ M \ M] M = \begin{bmatrix} 0 & 0.71 & 1 & 0.71 & 0 & -0.71 & -1 & -0.71 \\ 1 & 0.71 & 0 & -0.71 & 1 & -0.71 & 0 & 0.71 \end{bmatrix}$$

In this task, the subject's first goal is to interact through an unknown visuomotor mapping and internalize this model. The second problem is to use this model to solve a control problem.

1. System Identification – learning a transition function  $p(y_t|x_t, u_t)$ 
  - How do you learn the unknown observation model from data?
2. Policy Optimization
  - Once dynamics are learned (or at least stable?), how do we form a policy that is generalizable to new tasks under these dynamics?
  - This is the control problem.

It's safe to assume that these processes are happening in parallel. Because we have complete and arbitrary control over the observation mapping, we can ask the subject to interact through a dynamic that is intuitive (informative prior) or unintuitive (uninformative or inhibitive prior). Each scenario, we hypothesize, will elicit different strategies for learning and control.

The unknown mapping  $M$  from muscle to task space looks like the observation matrix  $H$  in the LQG problem:

$$\begin{aligned} y_t &= Hx_t + v_t \quad (\text{LQG}) \\ y_t &= Mx_t + v_t. \quad (\text{experiment}) \end{aligned}$$

The state dynamics in the task are of the form:

$$\begin{aligned} x_t &= Ax_{t-1} + Bu_{t-1} + w_{t-1} \quad (\text{LQG}) \\ x_t &= Dx_{t-1} + Iu_{t-1} + w_{t-1} \quad (\text{experiment}) \end{aligned}$$

where  $D$  is a diagonal decay matrix of with terms  $e^{-\lambda}$  and  $I$  is the identity. The subject produces muscle contractions which add to the current latent (unobserved) state. In the absence of control signals, the state decays back to 0 in line with the physics of your arm returning to a passive state in the absence of muscle contractions. The terms  $w$  and  $v$  are gaussian noise vectors with distributions  $\mathcal{N}(0, Q)$  and  $\mathcal{N}(0, R)$ . If we combine the transition and observation models:

$$\begin{aligned} y_t &= MDx_{t-1} + Mu_{t-1} + Mw_{t-1} + v_t \\ &= A'x_{t-1} + B'u_{t-1} + Mw_{t-1} + v_t. \end{aligned}$$

We can think of this as the combined system identification problem where  $A' = MD$  and  $B' = M$  are unknown and must be estimated. The noise covariances of this altered system are now non-trivial, however. We could also assume that the transition dynamic  $D$  is known and that the identification problem is learning the mapping  $M$  only. This might not be a poor assumption since the exponential decay is meant to serve as an intuitive passive dynamic.

In each trial of the task, a subject will have some internal representation of the observation dynamic  $M$  which may or may not be accurate. In order to make accurate predictions,  $M$  must be estimated.

Learning linear dynamical systems from data is a hot topic of research, most of which seems to focus on learning in the context of complete state observation ( $M = I$ ,  $y = x$ ). Algorithms to determine parameters of  $A$  and  $B$  are proposed (see Dean, Recht 2018).

From LQG theory we know that the control law is a linear function of the state:

$$u_t = -L_t x_t$$

and thus our combined system dynamic is:

$$y_t = M(D - L_t)x_{t-1} + Mw_{t-1} + v_t.$$

The noise covariance due to the observation  $Q$  is unchanged, but the new noise covariance for the latent process is now  $M R M^T$ . This may make things difficult.

## 5 Theoretical Contributions

- how do we adapt LQR controllers trial-to-trial?
- how do we use existing controllers to construct movements?
- how do we construct controllers under dynamical and goal uncertainty?

### 5.1 Internal Model Adaptation for Linear Quadratic Control

Here I investigate the effects of approximating internal dynamical models for movement and using the resulting endpoint error to update this approximation over trials.

Our state space is denoted  $x$  and our control space  $u$  where  $\dim(x) < \dim(u)$ . Each trial, we move from state  $x(0)$  to  $x(N)$  in  $N$  timesteps. Each trial, we have a goal state  $x^*$  and a resulting endpoint error  $e(N) = |x(N) - x^*|^2$ .

We use a deterministic linear dynamical system to model our within-trial state dynamics:

$$x(t) = Ax(t-1) + Bu(t-1).$$

For this system, we assume there exists a linear feedback control law optimal under a given quadratic state and control cost:

$$u(t) = Kx(t).$$

We can write the controlled, closed-loop system dynamics for the final time step  $N$ :

$$\begin{aligned} x(N) &= (A - BK)x(N-1) = Cx(N-1) \\ x(N) &= Cx(N-1) = C(Cx(N-2)) \\ x(N) &= C^N x(0). \end{aligned}$$

where  $C^N$  might be called the trajectory dynamic. If the trajectory dynamic  $C^N$  is an approximation to the true trajectory dynamic  $C^{N*}$ , we can use the error of a given trajectory to find an incremental update. The error at the final time step  $N$  for trial  $r$  is

$$e(r) = |C^N(r)x(0) - x^*|^2.$$

This error may be due to several sources. Our internal dynamics model  $A$  might have error relative to the true dynamic  $A^*$ . Our control gain  $K$  may be optimal relative to our internal model  $A$  but not with respect to the true dynamic  $A^*$ . Finally, we might have an approximate model  $A$  and a suboptimal control gain  $K$ . Note that since this is still deterministic system, we have yet to include any source of variability in state or control.

If we assume that our computation of the control gain  $K$  is optimal for our approximate internal model  $A$  (we can compute a controller given only our internal representation of the system dynamic being controlled), we can use our endpoint error to derive a gradient descent update for  $A$  on trial  $r$ :

$$A(r+1) = A(r) - \eta \frac{\partial e(r)}{\partial A}$$

We might think about this as an internal simulation of trial  $r$ 's trajectory, and a subsequent post hoc evaluation of the movement. To compute  $\delta$ , we must take the gradient with respect to  $A$  of the error:

$$\frac{\partial e(r)}{\partial A} = \frac{\partial}{\partial A} |C^N(r)x(0) - x^*|^2$$

Since the gradient with respect to  $A$  is the same as the gradient with respect to  $C$ , we can compute the gradient with respect to  $C$  to find:

$$\frac{\partial e}{\partial A_{ij}} = 2 \sum_{k=1}^N [(C^N x(0) - x^*)^T C^{k-1}]_i [C^{N-k} x(0)]_j$$

Below is a figure showing LQR simulations across gradient descent updates to the  $A$  matrix after it is corrupted by Gaussian noise. Each trajectory is a single run of the LQR controlled for 200 time steps. The star shows the target state, the colored circles show the endpoints of the trajectories. The red circle is the initial state.

The descent is converging in endpoint error in position, velocity, and force space. Unfortunately, this optimization is causing the dynamics to change. The routine is also very fragile to parameter changes. Next steps:

- Gain a better understanding of the loss landscape, including it's degeneracy. It may be possible to compute the optimum analytically.
- Corrupt the  $A$  matrix in a more principled way, working to alter the passive dynamics in a physically realistic manner.
- Explore the action of the resulting gradient through it's eigenvalues and vectors. This can be done in two dimensions as a starting point.
- Compute second-order derivatives and work towards a Newton's method.
- Compute derivatives with respect to the control law  $K$  as a comparison.
- Analyze results of the routine in comparison with the reaching adaptation literature.

Think more about subspaces - preparatory activity in one subspace, online control in another? - learning in one subspace but not another? - compression of model to a subspace?

## 5.2 Policy Selection

each timestep you combine actions from component policies to choose an action

Here we'll review and discuss models of action selection and policy composition as a means of theorizing about how subjects learn novel skills.

In a sense, we're setting up several different directions for our understanding of composition and action selection which can be experimentally tested.

We have a direct selection algorithm, composition through policy addition, and composition through policy multiplication.

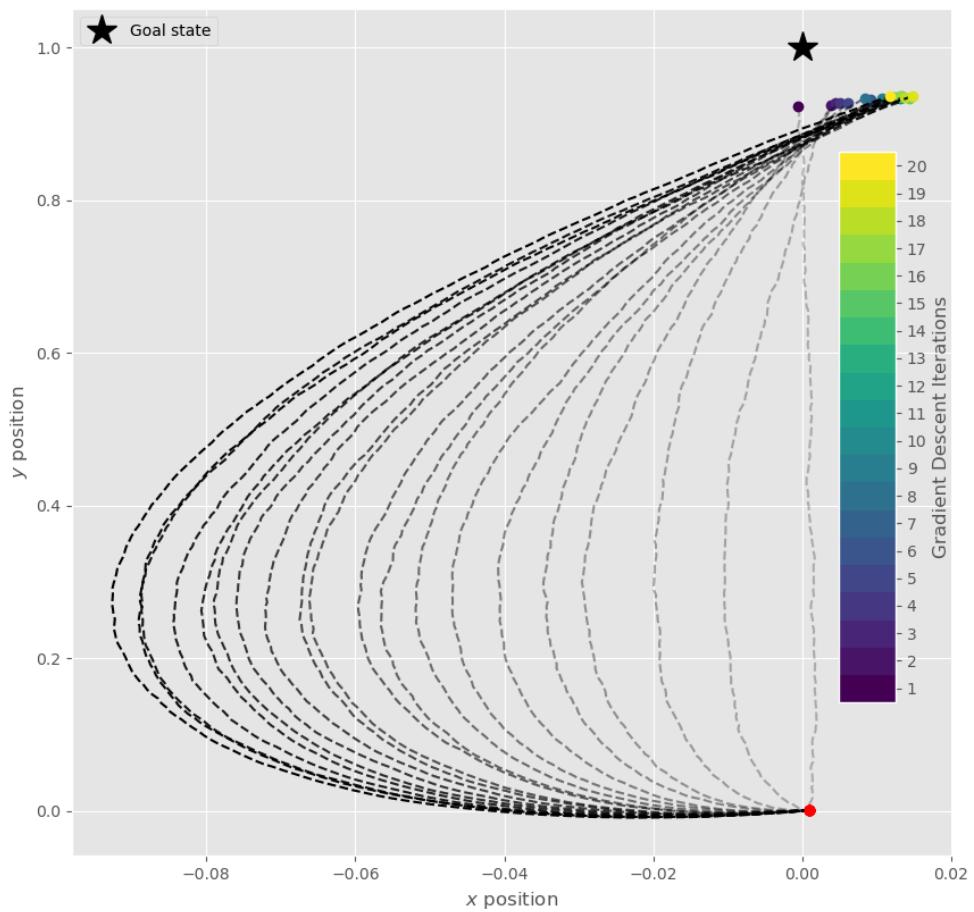


Figure 2: Iterations of gradient descent on the  $A$  matrix of an infinite-horizon LQR. Each dotted line in a trajectory with a different  $A$  matrix. The red circle denotes the initial state, the star denotes the goal state, and the colored circles denote the endpoints of each iteration.

### 5.2.1 KL-control Composition (1 day)

This setup is particular subset of OFC problems.

Dynamics Cost

Composable policies

PLOT OF INTUITIVE EXAMPLE

### 5.2.2 Temporal Composition

there is a spectrum of latency in the feedback response

can different controllers be used for different latencies, and adjusted accordingly?

### 5.2.3 Generalized Policy Selection (1 day)

This is in the MDP case

Learning happens in several ways— reward regression, Q-learning

What are rewards? What are tasks? What are actions?

Is GPI with LQRs / LQR-RL a good model for motor learning? Define a model and see if it recapitulates known motor learning phenomena on existing experiments + accounts for things that previous models don't. (Similar in spirit to Geerts et al. (2020)). Can this model track the higher-order statistics of trajectories during motor learning?

### 5.2.4 Model-based Reinforcement Learning

Since we only have an approximate model of the system dynamic, we could simply work towards an optimal policy directly using gradient derivative-free optimization methods in a model-free approach. Since we have good evidence that humans leverage internal models to make decisions (at least in a motor problem domain), we need to define an algorithm which uses past observations and controls to update our approximation for the system dynamic. Here is a very general algorithm:

0. Define a base policy/controller and base system model ( $L_0$  and  $\hat{M}_0$ )
1. Collect samples (by interacting with the true environment  $M_{true}$ ) using the current policy/controller (collect  $y_t, u_t, y_{t+1}$  triples using  $L_i$  for  $i \in \{0 \dots N\}$ )
2. Use sample(s) / trajectories to update current system dynamical model  $\hat{M}_i$
3. Update current policy/controller  $L_i$  (using the system dynamics or using a direct policy method)

If the true system dynamics were known, we could solve the Algebraic Riccati Equation with a backwards pass, and compute our controls in a forward pass. This general algorithm structure highlights how the (unknown) system identification and controller design are intertwined: identifying a system appropriately must rely on sampling and fitting regions of the state space pertinent to adequate control in terms of cost (Ross ICML 2012). Otherwise, our approximation to the true system dynamic will only produce a valid controller in regions we have previously explored. The question is how we can effectively (sample and time efficiently) utilize new state transitions we encounter either online as feedback or between trials to update our model and policy. That is, the number of trials and/or trajectories to use before updating either the system model and/or policy is an important parameter.

In the LQG setting, this might be called “adaptive LQG”.

## 5.3 Internal Model Adaptation for Linear Quadratic Control

Here I investigate the effects of approximating internal dynamical models for movement and using the resulting endpoint error to update this approximation over trials.

Our state space is denoted  $x$  and our control space  $u$  where  $\dim(x) < \dim(u)$ . Each trial, we move from state  $x(0)$  to  $x(N)$  in  $N$  timesteps. Each trial, we have a goal state  $x^*$  and a resulting endpoint error  $e(N) = |x(N) - x^*|^2$ .

We use a deterministic linear dynamical system to model our within-trial state dynamics:

$$x(t) = Ax(t-1) + Bu(t-1).$$

For this system, we assume there exists a linear feedback control law optimal under a given quadratic state and control cost:

$$u(t) = Kx(t).$$

We can write the controlled, closed-loop system dynamics for the final time step  $N$ :

$$\begin{aligned} x(N) &= (A - BK)x(N-1) = Cx(N-1) \\ x(N) &= Cx(N-1) = C(Cx(N-2)) \\ x(N) &= C^N x(0). \end{aligned}$$

where  $C^N$  might be called the trajectory dynamic. If the trajectory dynamic  $C^N$  is an approximation to the true trajectory dynamic  $C^{N*}$ , we can use the error of a given trajectory to find an incremental update. The error at the final time step  $N$  for trial  $r$  is

$$e(r) = |C^N(r)x(0) - x^*|^2.$$

This error may be due to several sources. Our internal dynamics model  $A$  might have error relative to the true dynamic  $A^*$ . Our control gain  $K$  may be optimal relative to our internal model  $A$  but not with respect to the true dynamic  $A^*$ . Finally, we might have an approximate model  $A$  and a suboptimal control gain  $K$ . Note that since this is still deterministic system, we have yet to include any source of variability in state or control.

If we assume that our computation of the control gain  $K$  is optimal for our approximate internal model  $A$  (we can compute a controller given only our internal representation of the system dynamic being controlled), we can use our endpoint error to derive a gradient descent update for  $A$  on trial  $r$ :

$$A(r+1) = A(r) - \eta \frac{\partial e(r)}{\partial A}$$

We might think about this as an internal simulation of trial  $r$ 's trajectory, and a subsequent post hoc evaluation of the movement. To compute  $\delta$ , we must take the gradient with respect to  $A$  of the error:

$$\frac{\partial e(r)}{\partial A} = \frac{\partial}{\partial A} |C^N(r)x(0) - x^*|^2$$

Since the gradient with respect to  $A$  is the same as the gradient with respect to  $C$ , we can compute the gradient with respect to  $C$  to find:

$$\frac{\partial e}{\partial A_{ij}} = 2 \sum_{k=1}^N [(C^N x(0) - x^*)^T C^{k-1}]_i [C^{N-k} x(0)]_j$$

Below is a figure showing LQR simulations across gradient descent updates to the  $A$  matrix after it is corrupted by Gaussian noise. Each trajectory is a single run of the LQR controlled for 200 time

steps. The star shows the target state, the colored circles show the endpoints of the trajectories. The red circle is the initial state.

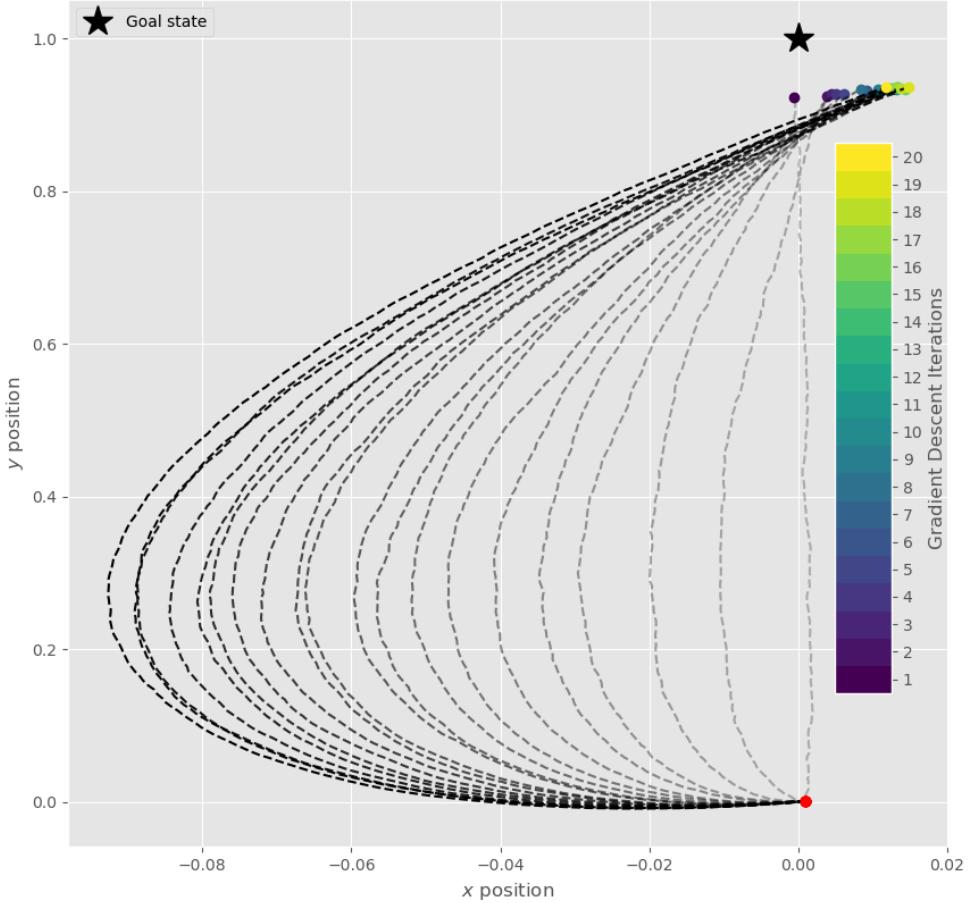


Figure 3: Iterations of gradient descent on the  $A$  matrix of an infinite-horizon LQR. Each dotted line in a trajectory with a different  $A$  matrix. The red circle denotes the initial state, the star denotes the goal state, and the colored circles denote the endpoints of each iteration.

The descent is converging in endpoint error in position, velocity, and force space. Unfortunately, this optimization is causing the dynamics to change. The routine is also very fragile to parameter changes. Next steps:

- Gain a better understanding of the loss landscape, including it's degeneracy. It may be possible to compute the optimum analytically.
- Corrupt the  $A$  matrix in a more principled way, working to alter the passive dynamics in a physically realistic manner.
- Explore the action of the resulting gradient through it's eigenvalues and vectors. This can be done in two dimensions as a starting point.
- Compute second-order derivatives and work towards a Newton's method.
- Compute derivatives with respect to the control law  $K$  as a comparison.
- Analyze results of the routine in comparison with the reaching adaptation literature.

## 6 Next Steps

Big Things

First completing a methodsy style piece of work about the setup, illustrating the generation of bespoke control strategies by careful tracking of the EMG modes relative to task performance.

Next is to work on models of learning by extending the framework of OFC through additions of composition and error-based adaptation.

### 6.1 Data Collection at Scale

- scale up data collection with more subjects across many days
- intersubject differences

#### 6.1.1 Data Analysis

Our preliminary data confirms the working principle of the setup and highlights the next steps for producing a quality dataset.

Address the noise concerns in the data – better hardware, should be very low noise relative to our signal.

- more advanced EMG analyses in the task setting
  - autoencoders (farina paper)<sup>71</sup>
  - bayesian inference techniques bespoke for EMG signals – pull out control modes
    - \* choosing different aspects of the data model for inference
      - what kind of priors for tunings, latent space,
      - we would need some kind of round truth? a different experiment...
- better techniques for developing mappings
- long-range correlations in the data, correlation functions
  - 72
- Dynamical modes in the data using dynamical systems analysis techniques

#### 6.1.2 Task Design

- formalize specific task designs which link with our theoretical interests

#### 6.1.3 Optimization

we want to stay close to models, testing them as we go - optimization models (regularized regression) – pure force learning - perturbations of this + predictions? mapping perturbation, noise perturbations

#### 6.1.4 Optimal Control

- stochastic optimal control model comparisons
  - cost models
  - perturbations in goal
  - go-before you know / goal uncertainty
  - noise perturbations – do responses match the models?
- dynamics model fitting
  - internal model uncertainty
  - modeled with robust optimal control?

stochastic control vs. robust control vs. adaptive control

learning control over trials learning control via reward (RL)<sup>73</sup> constructing control from primitives for transfer (GPS, KL-control)

## 6.2 Eye Tracking

Active learning, perception + action

## 6.3 Open Questions

The following questions need answers, to make progress we must form hypotheses around the most pressing of these questions and design experiments to test these hypotheses.

- how does a subject sample the state space as to efficiently learn? do they sample optimally?  
how does controller/policy optimization proceed based on system identification?
- how does a human subject use error information from each trial and feedback from each time step to update their model and/or policy?
  - how does a subject balance policy updates with model updates?
- On what scale (trials, timesteps) is the model altered? the policy?
  - Replanning at every timestep is a model predictive control algorithm
  - What prediction can we make for ID/learning every trial?
- how does a subject avoid “distribution mismatch” between their base policy and their optimal policy? How do they efficiently explore and use this new data to update their internal model?
  - what exploration strategy does a subject use to avoid mismatch?
  - what
- What is a subject’s baseline/prior model?  $y_t = \hat{f}_0(x_t, u_t)$  or  $y_t \propto p_0(y_t | x_t, u_t)$
- What is the base policy / prior policy?  $u_t = \pi_0(\hat{x}_t)$
- How do we think about learning a distribution over trajectories in control law space, or perhaps equivalently, in covariance/precision space?
- We might hypothesize that a subject will act as randomly as possible while minimizing cost, a maximum entropy solution that converges to an optimal controller?  $\mathcal{H}(p(u_t | x_t))$
- How does a subject penalize changes to their controllers? Do they follow a KL-divergence type of measurement when improving their policy?
- In a behavioral experiment, how can you disentangle system identification/estimation and control? Is suboptimality due to one or the other?
- How does the observation mapping relate to the latent state covariance? The task state covariance?
- How do we formalize this into a probabilistic graphical model? Why would we?
  - Would this make it easier to reason about what the goals are?
  - Would learning  $M$  become an inference problem?
  - Would solving the control problem become an inference problem...?
- What noise assumptions can we make? Can we not make?
  - How can we incorporate signal-dependent noise?

## Bibliography

1. McNamee, D. & Wolpert, D. M. Internal Models in Biological Control. *Annual Review of Control, Robotics, and Autonomous Systems* **2**, 339–364 (2019).
2. Todorov, E. Optimality principles in sensorimotor control. *Nature Neuroscience* **7**, 907–915 (2004).
3. Kober, J., Bagnell, J. A. & Peters, J. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research* **32**, 1238–1274 (2013).
4. Sauerbrei, B. A. *et al.* Cortical pattern generation during dexterous movement is input-driven. *Nature* (2019) doi:10.1038/s41586-019-1869-9.
5. Bernstein, N. *The coordination and regulation of movements*. (Pergamon, 1967).
6. Kitano, H. Biological robustness. *Nature Reviews Genetics* **5**, 826–837 (2004).
7. Fuglevand, A. J. Mechanical properties and neural control of human hand motor units: Control of human hand motor units. *The Journal of Physiology* **589**, 5595–5602 (2011).
8. Harris, C. M. & Wolpert, D. M. Signal-dependent noise determines motor planning. *Nature* **394**, 780–784 (1998).
9. van Duinen, H. & Gandevia, S. C. Constraints for control of the human hand: Control of the hand. *The Journal of Physiology* **589**, 5583–5593 (2011).
10. Valero-Cuevas, F. J. *et al.* The tendon network of the fingers performs anatomical computation at a macroscopic scale. *IEEE Transactions on Biomedical Engineering* **54**, 1161–1166 (2007).
11. Yan, Y., Goodman, J. M., Moore, D. D., Solla, S. A. & Bensmaia, S. J. Unexpected complexity of everyday manual behaviors. *Nature Communications* **11**, 3564 (2020).
12. Basmajian, J. V. Control and Training of Individual Motor Units. *Science* **141**, 440–441 (1963).
13. Merel, J., Botvinick, M. & Wayne, G. Hierarchical motor control in mammals and machines. *Nature Communications* **10**, 5489 (2019).
14. D'Avella, A., Saltiel, P. & Bizzi, E. Combinations of muscle synergies in the construction of a natural motor behavior. *Nature Neuroscience* **6**, 300–308 (2003).
15. Giszter, S. F. Motor primitivesnew data and future questions. *Current Opinion in Neurobiology* **33**, 156–165 (2015).
16. Rácz, K. & Valero-Cuevas, F. J. Spatio-temporal analysis reveals active control of both task-relevant and task-irrelevant variables. *Frontiers in Computational Neuroscience* **7**, (2013).
17. Ingram, J. N. & Wolpert, D. M. The statistics of natural hand movements. *Brain* **188**, 223–236 (2009).
18. Todorov, E. & Ghahramani, Z. Analysis of the synergies underlying complex hand manipulation. in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* vol. 4 4637–4640 (IEEE, 2005).

19. Bizzzi, E. & Ajemian, R. From motor planning to execution: A sensorimotor loop perspective. *Journal of Neurophysiology* **124**, 1815–1823 (2020).
20. Bruton, M. & O'Dwyer, N. Synergies in coordination: A comprehensive overview of neural, computational, and behavioral approaches. *Journal of Neurophysiology* **120**, 2761–2774 (2018).
21. Lemon, R. N. Cortical control of the primate hand. *Experimental Physiology* **78**, 263–301 (1993).
22. Lemon, R. N. Mechanisms of cortical control of hand function. *Neuroscientist* **3**, 389–398 (1997).
23. Lemon, R. N. Descending Pathways in Motor Control. *Annual Review of Neuroscience* **31**, 195–218 (2008).
24. Lemon, R. & Kraskov, A. Starting and stopping movement by the primate brain. *Brain and Neuroscience Advances* **3**, 239821281983714 (2019).
25. Kawasawa, Y. I. *et al.* Control of species-dependent cortico-motoneuronal connections underlying manual dexterity. *Science* **357**, 400–404 (2017).
26. Murabe, N. *et al.* Higher primate-like direct corticomotoneuronal connections are transiently formed in a juvenile subprimate mammal. *Scientific Reports* **8**, 1–10 (2018).
27. Cheney, P. D. & Fetz, E. E. Functional classes of primate corticomotoneuronal cells and their relation to active force. *Journal of Neurophysiology* **44**, 773–791 (1980).
28. Griffin, D. M. & Strick, P. L. The motor cortex uses active suppression to sculpt movement. *Science Advances* **6**, eabb8395 (2020).
29. Rathelot, J.-A. & Strick, P. L. Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proceedings of the National Academy of Sciences* **106**, 918–923 (2009).
30. Griffin, D. M., Hoffman, D. S. & Strick, P. L. Corticomotoneuronal cells are "functionally tuned". *Science* **350**, 667–670 (2015).
31. Takei, T., Confais, J., Tomatsu, S., Oya, T. & Seki, K. Neural basis for hand muscle synergies in the primate spinal cord. *Proceedings of the National Academy of Sciences* **114**, 8643–8648 (2017).
32. Dum, R. P. & Strick, P. L. The Corticospinal System: A Structural Framework for the Central Control of Movement. in *Comprehensive Physiology* (John Wiley & Sons, Inc., 2011). doi:10.1002/cphy.cp120106.
33. Furuya, S. & Altenmüller, E. Flexibility of movement organization in piano performance. *Frontiers in Human Neuroscience* **7**, (2013).
34. Graziano, M. S. A. *The intelligent movement machine: An ethological perspective on the primate motor system.* (Oxford University Press, 2009).
35. Ebina, T. *et al.* Arm movements induced by noninvasive optogenetic stimulation of the motor cortex in the common marmoset. *Proceedings of the National Academy of Sciences* **116**, 22844–22850 (2019).
36. Watanabe, H. *et al.* Forelimb movements evoked by optogenetic stimulation of the macaque motor cortex. *Nature Communications* **11**, 3253 (2020).

37. Wiltschko, A. B. *et al.* Mapping Sub-Second Structure in Mouse Behavior. *Neuron* **88**, 1121–1135 (2015).
38. Berger, D. J., Masciullo, M., Molinari, M., Lacquaniti, F. & d'Avella, A. Does the cerebellum shape the spatiotemporal organization of muscle patterns? Insights from subjects with cerebellar ataxias. *Journal of Neurophysiology* **123**, 1691–1710 (2020).
39. Bahl, S., Mukadam, M., Gupta, A. & Pathak, D. Neural Dynamic Policies for End-to-End Sensorimotor Learning. *5* (2020).
40. Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P. & Schaal, S. Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors. *Neural Computation* **25**, 328–373 (2013).
41. Carroll, T. J., McNamee, D., Ingram, J. N. & Wolpert, D. M. Rapid Visuomotor Responses Reflect Value-Based Decisions. *The Journal of Neuroscience* **39**, 3906–3920 (2019).
42. Weiler, J., Gribble, P. L. & Pruszynski, J. A. *Spinal stretch reflexes support efficient control of reaching.* (2020) doi:10.1101/2020.01.06.896225.
43. Pruszynski, J. A., Omrani, M. & Scott, S. H. Goal-Dependent Modulation of Fast Feedback Responses in Primary Motor Cortex. *Journal of Neuroscience* **34**, 4608–4617 (2014).
44. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that finds a naturalistic solution for the production of muscle activity. *Nature Neuroscience* **18**, 1025–1033 (2015).
45. Akay, T. & Murray, A. J. Relative Contribution of Proprioceptive and Vestibular Sensory Systems to Locomotion: Opportunities for Discovery in the Age of Molecular Science. *Int. J. Mol. Sci.* **18** (2021).
46. Cohen, A. H. The Role of Hierarchical Control in the Evolution of Central Pattern Generators. *Brain, Behavior and Evolution* **40**, 112–124 (1992).
47. Hunt, L. T. & Hayden, B. Y. A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews Neuroscience* **18**, 172–182 (2017).
48. Hunt, L. T., Dolan, R. J. & Behrens, T. E. J. Hierarchical competitions subserving multi-attribute choice. *Nature Neuroscience* **17**, 1613–1622 (2014).
49. McCulloch, W. S. A heterarchy of values determined by the topology of nervous nets. *The Bulletin of Mathematical Biophysics* **7**, 89–93 (1945).
50. Crumley, C. L. Heterarchy and the Analysis of Complex Societies. *Archeological Papers of the American Anthropological Association* **6**, 1–5 (2008).
51. Bechtel, W. & Bich, L. Grounding cognition: Hierarchical control mechanisms in biology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20190751 (2021).
52. Adrian, M. H. & John, W. K. Theoretical Models of Motor Control and Motor Learning. in *Routledge Handbook of Motor Control and Motor Learning* (Routledge, 2012). doi:10.4324/9780203132746.ch1.
53. Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L. & Haith, A. M. Motor learning. *Comprehensive Physiology* **9**, 613–663 (2019).
54. Shadmehr, R. & Krakauer, J. W. A computational neuroanatomy for motor control. *Experimental Brain Research* **185**, 359–381 (2008).

55. Manley, H., Dayan, P. & Diedrichsen, J. When Money Is Not Enough: Awareness, Success, and Variability in Motor Learning. *PLoS ONE* **9**, e86580 (2014).
56. van Beers, R. J. Motor Learning Is Optimally Tuned to the Properties of Motor Noise. *Neuron* **63**, 406–417 (2009).
57. van Beers, R. J., Brenner, E. & Smeets, J. B. J. Random walk of motor planning in task-irrelevant dimensions. *Journal of Neurophysiology* **109**, 969–977 (2013).
58. Mussa-Ivaldi, F. A., Casadio, M., Danziger, Z. C., Mosier, K. M. & Scheidt, R. A. Sensory motor remapping of space in human-machine interfaces. *Progress in Brain Research* **191**, 45–64 (2011).
59. Mosier, K. M., Scheidt, R. A., Acosta, S. & Mussa-Ivaldi, F. A. Remapping Hand Movements in a Novel Geometrical Environment. *Journal of Neurophysiology* **94**, 4362–4372 (2005).
60. Nazarpour, K., Barnard, A. & Jackson, A. Flexible Cortical Control of Task-Specific Muscle Synergies. *Journal of Neuroscience* **32**, 12349–12360 (2012).
61. Nazarpour, K., Al-Timemy, A. H., Bugmann, G. & Jackson, A. A note on the probability distribution function of the surface electromyogram signal. *Brain Research Bulletin* **90**, 88–91 (2013).
62. Sanger, T. D. Bayesian Filtering of Myoelectric Signals. *Journal of Neurophysiology* **97**, 1839–1845 (2007).
63. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
64. Churchland, M. M. Neural Variability in Premotor Cortex Provides a Signature of Motor Preparation. *Journal of Neuroscience* **26**, 3697–3712 (2006).
65. Crouzier, M. *et al.* Do individual differences in the distribution of activation between synergist muscles reflect individual strategies? *Experimental Brain Research* **237**, 625–635 (2019).
66. Hug, F. Can muscle coordination be precisely studied by surface electromyography? *Journal of Electromyography and Kinesiology* **21**, 1–12 (2011).
67. Dyson, M., Barnes, J. & Nazarpour, K. Myoelectric control with abstract decoders. *Journal of Neural Engineering* **15**, (2018).
68. Berger, D. J., Gentner, R., Edmunds, T., Pai, D. K. & d'Avella, A. Differences in Adaptation Rates after Virtual Surgeries Provide Direct Evidence for Modularity. *Journal of Neuroscience* **33**, 12384–12394 (2013).
69. Radhakrishnan, S. M., Baker, S. N. & Jackson, A. Learning a Novel Myoelectric-Controlled Interface Task. *Journal of Neurophysiology* **100**, 2397–2408 (2008).
70. de Rugy, A., Loeb, G. E. & Carroll, T. J. Muscle Coordination Is Habitual Rather than Optimal. *Journal of Neuroscience* **32**, 7384–7391 (2012).
71. Vujaklija, I. *et al.* Online mapping of EMG signals into kinematics by autoencoding. *Journal of NeuroEngineering and Rehabilitation* **15**, 21 (2018).
72. Crevecoeur, F., Bollens, B., Detrembleur, C. & Lejeune, T. M. Towards a ‘gold-standard’ approach to address the presence of long-range auto-correlation in physiological time series. *Journal of Neuroscience Methods* **11** (2010).

73. van der Kooij, K., van Mastrigt, N. M., Crowe, E. M. & Smeets, J. B. J. Learning a reach trajectory based on binary reward feedback. *Scientific Reports* **11**, 2667 (2021).