Sensorimotor Learning in Virtual Environments

Spencer R. Wilson

February 4, 2021

1 Introduction & Aims

> *Movement is nothing but the quality of our being.*
>
> — Sunryu Suzuki, *Zen Mind, Beginner's Mind*

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. For example, Moravec's Paradox encourages us to not look past the stunningly complex computations generated by the human motor apparatus. Following Moravec's suggestion, this work focuses on the human motor system, the most advanced control apparatus in the known universe.

A recent review corroborates this perspective and provides a clear call to action:

> 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.[1]

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[2], reinforcement learning in continuous action space[3], and detailed physiological studies[4]. 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."[5] The crux of this definition is the flexibility of such solutions. This flexibility, or robustness[16], 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.

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.

---

[1]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.

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[2]. 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 it's 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 these contractions during skill acquisition and subsequent goal-oriented muscle activations. 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 3.

how are value computations connected to action and policy selection how are feedback controllers adapted to motor errors, new environments, how are they learned as well as combined?

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 it's relation to dexterity will provide a scaffold on which we can hang our experimental and theoretical investigations detailed in Section 3 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 neuro-muscular 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 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.

---

[2]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?

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[7].

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[7,8].

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 metaphalangeal 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[7]. Adduction and Abduction of the fingers is produced by the 19 instrinsic muscles of the hand, each of which has their origin and insertion points within the hand itself[9]. The instrinisc 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[10]. 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[11,123]. The structure of the neuromuscular system that underlies this variety offers many clues as to the relevant computations required for dexterous movement.

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.[5] 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 is evolutionarily advantageous 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". 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

---

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

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[13]. 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[14].[4] 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[15,Gao2017?]. 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.[16] 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[17,18]. 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.^[It has been argued that the motor repertoire is hardly high-dimensional when compared to the dimensionality of the visual feature extraction system[11]. 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[11]. 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[19]. 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. 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 computation and structure 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[20–22]. The key insight of the work is that while "the organization of the spinal cord is based on relatively rigid muscular modes, a mechanism 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[23]. 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[24,25].

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

---

[4]I should really have more studies here, or a really nice review.

CM connections is their influence over multiple muscle compartments as well as multiple muscles, though typically agonist or antagonist sets[26]. This may seem counter-intuitive as a means to produce individuated movement, but experimental evidence in primates has show 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.[27] 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.[26].

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[28–30]. 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[31]. 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".[29]

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[32]. These results are 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[7].

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.

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.[4] This finding is relevant for our purposed 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 reliabaly produced stereotyped movements in primates, shown in Fig. 1[33]. 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 homonculus. Graziano refers to this as the cortical

"action map", that these stimulations tapped into the control mechanisms of the primate's motor system[34].
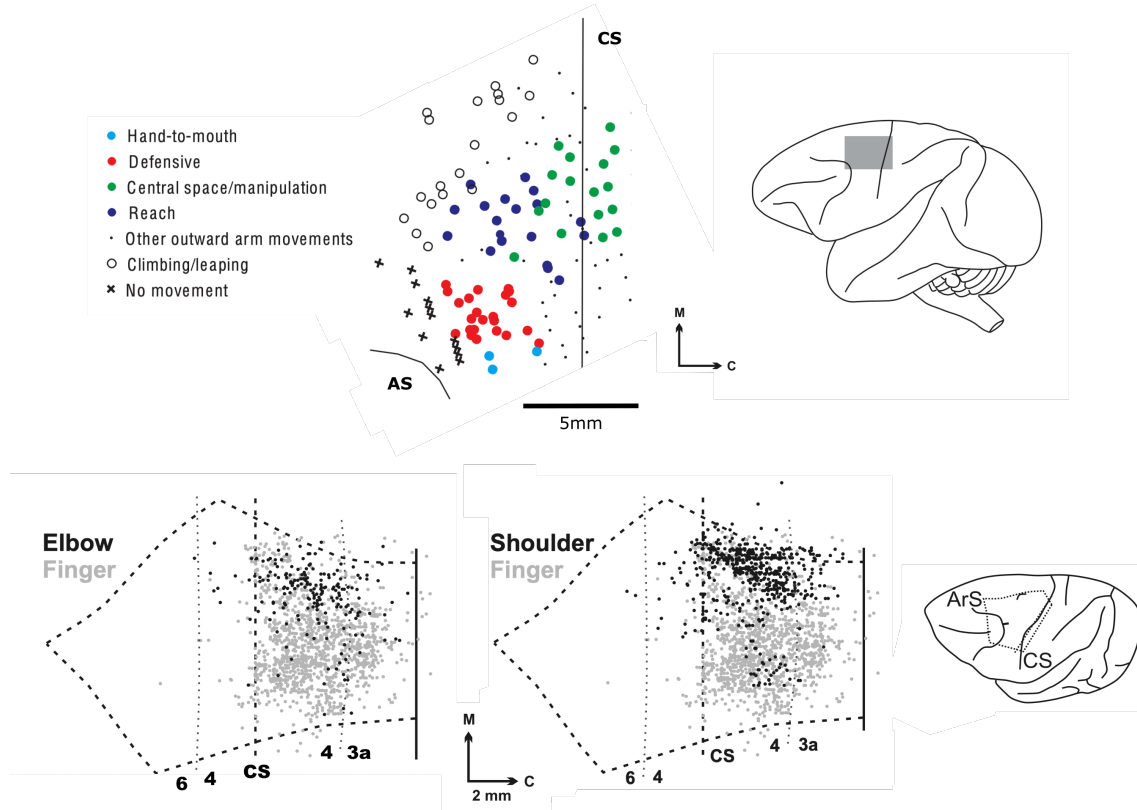

Figure 1: Similarities between electrical stimulation on behavorial timescales and rabies tracing identification of CM cells. CM cells are largely confined to the caudal half of M1, while this region tends to evokes complex manipulatory movements when electrically stimulated. (Top Left) Corticomotoneuronal (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*

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.[35] 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 wields 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 flexible activate the relevant action maps. Prevailing theory suggests that cerebellar loops provide predictive state information while basal gangliar loops provide state and/or action value information.

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[36,37]. This has some precedence in the literature and will be discussed further in **??**. This section has attempted to illustrate the complexity 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.

3   Experimental Methods

3.1   Prior Art

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.

D. J. Berger, R. Gentner, T. Edmunds, D. K. Pai, and A. d'Avella. Differences in Adap- tation Rates after Virtual Surgeries Provide Direct Evidence for Modularity. Journal of Neuroscience, 33(30):12384–12394, 2013. Zachary Danziger, Alon Fishbach, and Ferdinando A Mussa-ivaldi. Learning Algorithms for Human – Machine Interfaces. 56(5):1502–1511, 2009. Kristine M. Mosier, Robert A. Scheidt, Santiago Acosta, and Ferdinando A. Mussa-Ivaldi. Remapping Hand Movements in a Novel Geometrical Environment. Journal of Neurophys- iology, 94(6):4362–4372, December 2005. K. Nazarpour, A. Barnard, and A. Jackson. Flexible Cortical Control of Task-Specific Muscle Synergies. Journal of Neuroscience, 32(36):12349–12360, 2012. Saritha M. Radhakrishnan, Stuart N. Baker, and Andrew Jackson. Learning a Novel Myoelectric-Controlled Interface Task. Journal of Neurophysiology, 100(4):2397–2408, Oc- tober 2008.

3.1.1   Arbitrary Visuomotor Mappings

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

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?

3.2   High-Dimensional EMG Feedback

Electromyography is the detection of changes in chemical potential using electrodes. In my setup, we use a total 64 monopolar surface electrodes and monopolar needle electrodes to record chemical potentials from muscles in the forearm and hand.

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.

Francois Hug. Can muscle coordination be precisely studied by surface electromyography? Journal of Electromyography and Kinesiology, 21(1):1–12, February 2011.

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 thee 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.

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 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?

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

We then would like to present fixed mappings between hand output (either through direct muscle activity or through a controller such as a force pad).

3.3 Unsupervised Feature Extraction

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. Expore 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 addition to) physiological structure. This needs to be done incredibly carefully to escape criticism of hard-wired synergy enthusiasts.
   - See *de Rugy 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:

$$X = W \cdot H$$
$$W^{-1} \cdot X = H$$
$$W^T \cdot X = H$$

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. :

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

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.

3.4   Task Formalization

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:

$$y_t = Hx_t + v_t \text{ (LQG)}$$
$$y_t = Mx_t + v_t. \text{ (experiment)}$$

The state dynamics in the task are of the form:

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

where D is a diagonal decay matrix of with terms $\mathrm{e}^{-\lambda}$ 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:

$$y_t = MDx_{t-1} + Mu_{t-1} + Mw_{t-1} + v_t$$
$$= A'x_{t-1} + B'u_{t-1} + Mw_{t-1} + v_t.$$

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 MRM^T. This may make things difficult.

3.4.0.1 Questions

- 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?

3.4.1 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 $i \in \{0 \ldots N\}$
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".

3.4.1.1 Questions

- 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?

4 Data Analysis

5 Theory

6 Next Steps

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. Bizzi, E. & Ajemian, R. From motor planning to execution: A sensorimotor loop perspective. *Journal of Neurophysiology* **124**, 1815–1823 (2020).

20. Lemon, R. N. Cortical control of the primate hand. *Experimental Physiology* **78**, 263–301 (1993).

21. Lemon, R. N. Mechanisms of cortical control of hand function. *Neuroscientist* **3**, 389–398 (1997).

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

23. Lemon, R. & Kraskov, A. Starting and stopping movement by the primate brain. *Brain and Neuroscience Advances* **3**, 239821281983714 (2019).

24. Kawasawa, Y. I. *et al.* Control of species-dependent cortico-motoneuronal connections underlying manual dexterity. *Science* **357**, 400–404 (2017).

25. Murabe, N. *et al.* Higher primate-like direct corticomotoneuronal connections are transiently formed in a juvenile subprimate mammal. *Scientific Reports* **8**, 1–10 (2018).

26. 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).

27. Griffin, D. M. & Strick, P. L. The motor cortex uses active suppression to sculpt movement. *Science Advances* **6**, eabb8395 (2020).

28. 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).

29. Griffin, D. M., Hoffman, D. S. & Strick, P. L. Corticomotoneuronal cells are "functionally tuned". *Science* **350**, 667–670 (2015).

30. 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).

31. 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.

32. Furuya, S. & Altenmüller, E. Flexibility of movement organization in piano performance. *Frontiers in Human Neuroscience* **7**, (2013).

33. Graziano, M. THE ORGANIZATION OF BEHAVIORAL REPERTOIRE IN MOTOR CORTEX. *Annual Review of Neuroscience* **29**, 105–134 (2006).

34. Graziano, M. S. A. *The intelligent movement machine: An ethological perspective on the primate motor system.* (Oxford University Press, 2009).

35. Wiltschko, A. B. *et al.* Mapping Sub-Second Structure in Mouse Behavior. *Neuron* **88**, 1121–1135 (2015).

36. Bahl, S., Mukadam, M., Gupta, A. & Pathak, D. Neural Dynamic Policiesfor End-to-End Sensorimotor Learning. 5 (2020).

37. 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).