Heterarchical Control in Sensorimotor Processing

Spencer R. Wilson

# 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 construct a high-dimensional electromyography recording setup as a platform on which a suite of motor control and learning experiments can be explored. With this novel setup, we hope to advance our understanding of trial-to-trial motor learning through the combination of experiment and theory.

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.[1](#ref-McNamee2019)

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[**Todorov2004?**](#ref-Todorov2004), reinforcement learning in continuous action space[2](#Xab6e78cce052d8d5fae8c0b3ccfc447833f4719), and detailed physiological studies[3](#X0f8dee2e4b88f61b0e2e514140c93bf67074d33). However, as the quote above suggests, a holistic understanding of the computations underlying the construction of skilled movement remains an exciting direction for research. Our aim is to progress 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.

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*. Soviet neuroscientist Nikolai Bernstein defined dexterity as the ability to “find a motor solution in any situation and in any condition.”[4](#ref-Bernstein1967) The crux of this definition is the flexibility of such solutions. This flexibility, or robustness[[1]](#footnote-20)[5](#ref-kitanoBiologicalRobustness2004), 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.

To explore dexterous movement, we will leverage recordings of muscles controlling the hand as a readout of flexible motor behavior. This is a step beyond recording hand kinematics as electromyography provides a physiological output of the nervous system. Surface electromyography recordings taken from the forearms controlling subjects’ dominant hands 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. Preliminary work in this direction is described in Section ??.

Using data from our experimental setup, we wish to understand both how the structure of muscle activation variability evolves during skill acquisition and how the motor system constructs skilled movement through the composition of component muscle activations. To begin, we review a sampling of current motor physiology research relevant to dexterous motor computations in Section [2](#sec:physiology). In Section ??, we cover our prototype hardware and experiments. With inspiration from physiology and our experiments, we hope to make progress in modeling sensorimotor control and learning in our experimental setup. We cover preliminary work in this direction in Section [4](#sec:theory), and discuss possible future directions in Section [5](#sec:next_steps).

# 2 Reverse Engineering the Movement Machine

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

— Bizzi & Ajemian, *2020*

## 2.1 Motor Units to Muscles

The quantum of motor output is the motor unit (MU), defined as a single motoneuron axon and the set of junctions the terminals of its axon branches form with one or more muscle fibers. The MU provides the motor system with spatial redundancy at the muscle level: multiple muscle fibers contract due to a single alpha motoneuron (AMN) spike in the spinal cord’s ventral horn, 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[6](#Xfa6370ebd0dadc4bc8befaa8bdaf259a580ed09). 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.

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[6](#Xfa6370ebd0dadc4bc8befaa8bdaf259a580ed09). 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[7](#ref-vanduinenConstraintsControlHuman2011). 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.

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 an incredible variety of movement patterns[8](#ref-yanUnexpectedComplexityEveryday2020),[9](#ref-Basmajian1963)[[2]](#footnote-22). The structure of the neuromuscular system that underlies this variety offers many clues as to the relevant computations required for dexterous movement. In Fig. [1](#fig:low_variance_PCs), Yan et al. show how even low-variance principle components of joint kinematics during object grasping and ASL signing display correlational structure and not merely noise. That is, the production of hand movement is highly task-specific, where individual tasks are linked to bespoke muscle activations patterns.

Figure 1: Taken from Yan et al. 2020. Plots show mean correlations between hand joint kinematic trajectories during grasp trials with the same (blue) and different (red) objects (a) and ASL signs (b) projected onto the same principle components. Correlations are averaged across 8 subjects. Within-object and within-sign correlations are systematically higher than their shuffled counterparts. Error bars denote SEM. This data supports the idea that low-variance components of kinematics data contain task-specific structure rather than merely reflecting noise. This is encouraging for our experiments, which hope to extend this idea into careful analyses of task specific features of EMG data across learning and in response to perturbations.

Figure 1: Taken from Yan et al. 2020. Plots show mean correlations between hand joint kinematic trajectories during grasp trials with the same (blue) and different (red) objects (a) and ASL signs (b) projected onto the same principle components. Correlations are averaged across 8 subjects. Within-object and within-sign correlations are systematically higher than their shuffled counterparts. Error bars denote SEM. This data supports the idea that low-variance components of kinematics data contain task-specific structure rather than merely reflecting noise. This is encouraging for our experiments, which hope to extend this idea into careful analyses of task specific features of EMG data across learning and in response to perturbations.

## 2.2 Coordinative Structures

Many studies have contributed to the concept of synergies as a hard-wired organizing feature of the motor system[10](#ref-DAvella2003). 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[11](#ref-giszterMotorPrimitivesNew2015),[12](#ref-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[13](#X004733ed781b28fb613c114f4c14a4def5a39c6). 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[14](#ref-Ingram2009),[15](#ref-TodorovDimensionality2005). 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. Note that the motor repertoire is hardly high-dimensional when compared to the dimensionality of the visual feature extraction system[8](#ref-yanUnexpectedComplexityEveryday2020). 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[8](#ref-yanUnexpectedComplexityEveryday2020). 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[16](#ref-bizziMotorPlanningExecution2020). 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[17](#X066222b83ea59dbe217ccd139f7bcfa4f55df9b). 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

Just as many muscle fibers may be innervated by a single AMN, up to thousands of neurons contact single AMNs through monosynaptic corticospinal, or corticomotoneuronal (CM), connections and other descending pathways through elaborate spinal circuitry. The hallmark of CM connections in particular is their influence over multiple muscle compartments as well as multiple muscles, though typically agonist or antagonist sets[18](#ref-cheneyFunctionalClassesPrimate1980). 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.[19](#ref-griffinMotorCortexUses2020) These findings confirm theories about the excitatory and inhibitory role of these connections dating back decades, and combine to suggest that variables encoded in cortical ensembles are more complex than kinematics or dynamics alone[18](#ref-cheneyFunctionalClassesPrimate1980).

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[20](#ref-Rathelot2009)–[22](#ref-Takei2017). 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[23](#ref-dumCorticospinalSystemStructural2011). 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”[21](#Xe4ce91f75c0ba1bd76301cfdd9381fcdf57f708). The hypothesis stemming from the previously described work is that CM connections override the “consolidated” patterns putatively generated via spinal interneuron circuitry. The setup devised in our work aims to measure fractionation by tracing motor unit correlations across learning. Whether fractionation in our experiments is due to the CM pathway can only be speculation, but our work may provide direction for future studies pairing intracortical recordings with careful electromyography.

## 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[3](#X0f8dee2e4b88f61b0e2e514140c93bf67074d33). Thinking of the structural architecture of M1 as an input-driven system with outputs along a spectrum of modularity from synergistic to fractionated, we can 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[24](#ref-graziano2006). 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 humunculus. Graziano refers to this as the cortical “action map”, that these stimulations tapped into the control mechanisms of the primate’s motor system[25](#X9356a0badb28708abdd9cfadb1fe3511ea7c662). These results has recently been confirmed by optogenetics work in marmosets and macaques.[26](#ref-ebina2019),[27](#ref-watanabeForelimbMovementsEvoked2020)

The motor map concept suggests interpreting activity in M1 as a field of feedback control microcircuits, integrating and transforming inputs, both internal and external, to sculpt ongoing movement[**wiltschkoMappingSubSecondStructure2015?**](#Xedd47360e20fcc75178835d98b09068eeeab2c3). 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. Fig. [2](#fig:strick_graziano) shows Graziano et al.’s stimulation results, what might be termed a functional view of the cortical motor system, next Strick er al.’s described above clarifying the structural view of modularity in this system.

Figure 2: 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 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 et al. 2009[20](#ref-Rathelot2009),[28](#ref-graziano2005).

Figure 2: 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 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 et al. 2009[20](#ref-Rathelot2009),[28](#ref-graziano2005).

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 this sense, a readout from a network transforming state-dependent inputs into movement goals. Rather than choosing muscle patterns in reconfigurable blocks, it creatively constructs and sculpts movement. The hierarchy of the motor system may not be rigidly organized around a particular set of variables. As shown in Fig. [3](#fig:motor_system), many loops exist connecting cortex with the spinal cord, the cerebellum, the basal ganglia, and the sensorimotor periphery. Each of these loops contributes information for the flexible activation of the relevant action maps. Put simply, prevailing evidence suggests that cerebellar loops provide predictive state information while basal gangliar loops provide state and/or action value information. Taken together, this work provides an image of the incredible complexity which generates dexterous movements of the hand. This is the foundation on which we can work to build experiments which elucidate the computations involved in the production of skilled movement. We aim to connect our results back to what is known about the system we are attempting to reverse-engineer in order to inspire future inquiries into the inner workings of the movement machine.

Figure 3: Overview sketch of the motor system depicting the the redundancy of the system both hierarchically (multiple muscle fibers are innervated by the same motor neuron, many motor neurons innervate the same muscle) as well as heterarchically (parallel spinal, corticomotoneuronal, cerebellum, basal gangliar feedback loops). Parallel reflex responses can be classified as long latency (approximately 60-150ms) and short latency (approximately 60ms). We hope to consider the parallelism and redundancy of the motor system to inspire our data analyses and models of motor computation.

Figure 3: Overview sketch of the motor system depicting the the redundancy of the system both hierarchically (multiple muscle fibers are innervated by the same motor neuron, many motor neurons innervate the same muscle) as well as heterarchically (parallel spinal, corticomotoneuronal, cerebellum, basal gangliar feedback loops). Parallel reflex responses can be classified as long latency (approximately 60-150ms) and short latency (approximately 60ms). We hope to consider the parallelism and redundancy of the motor system to inspire our data analyses and models of motor computation.

## 2.5 John Rothwell and Jens Bo Nielson: Voluntary Control

In the vast majority of cases, cortical inputs fi rst contact interneurones which then relay the commands to motoneurones. Since the same interneurones also receive con -tinuous input from sensory receptors (and hence might be thought to participate in spinal refl exes) as well as from interneurones from other parts of the spinal cord, this means that by the time cortical input reaches motoneurones it has been fi ltered by multiple lower level systems. In higher primates and in man, cortical input can access some motoneurones via a special direct pathway (the corticomotoneuronal pathway), which is often supposed to play a critical role in volitional movement. However, even if this input is strong (and there is little comparative evidence on this) the excitability of motoneurones will have been biased by the multiple other inputs that each one receives. Thus, even this connection does not guarantee the brain a straightforward control of muscle.

We argue that distributed cortical projections allow for flexibility of connections between muscle representations, and therefore are critical to the flexibility of movements unrestricted by postural demands. Physiologically, the cortex is the main gateway for visual inputs to enter and infl uence motor control. This is particularly relevant during reaching with the arm and during the swing phase of gait for the leg. In both cases, the limbs are relatively free from feedback control from gravitational and contact force sensors, and can therefore be driven to a large extent by visual inputs.

Posture and contact dictates much of corticospinal function, and we would expect that these demands influence the architecture of the underlying circuitry. “Conscious” control is likely simply the availability of visual and propriospinal sensation in the course of the movement.

The anatomy and physiology of the [corticospinal] connections mean that if a volitional command were formulated in some hypothetical centre, it would be extremely diffi cult to predict the consequences with any certainty unless the state of every interposed connection were known in advance.

Nielson argues that hierarchy doesn’t function straightforwardly, that all corticospinal loops can be seen as a collection of distributed, overlapping modulation of motor neuron activity.

About 40 percent of the corticospinal fi bers come from the primary motor cortex, whereas the cingulate and supplementary areas supply only about 20 percent each and the premotor areas some 10 percent (Lemon, 2008). All of these areas of cortex also project to brainstem areas that give rise to reticulospinal tracts, giving them an indirect, cortico- reticulo-spinal route to the spinal cord as well as the direct corticospinal route. The primary motor cortex is thought to have fewer of these indirect connections than other motor areas, suggesting that its output represents the most ‘favored’ cortical access to spinal cord circuits.

In primates most of the terminations of the corticospinal tract are on interneurones in the spinal grey matter with a smaller number of direct monosynaptic inputs to motoneurones, particularly those inner-vating the distal muscles of the extremities. These connections represent the only way the cortex can interact directly with the motor apparatus (corticomotoneuronal connections).

after pyramidotomy […] movements remain slower and fatigued more rapidly.

This may be connected to synchronization of the CST? synchronous, driving input for rapid reactions.

We do not know the rules that specify spinal organization in any detail. However, one striking observation is that most of the connections between sensory input and motoneurone output are indirect, going via interneurones rather than direct sensory- motor pathways ( Jankowska, 2001; 2008). An obvious exception to this is the monosynaptic connection between primary muscle spindle afferents and their homony-mous motoneurones. However, this seems very much to be a special case rather than the rule.

One advantage of having interposed interneurones is that they are an effective way of allowing the spinal circuitry to switch between different states. For example, in the two funda-mental states of stance and gait, connectivity during posture should be arranged in order to resist perturbations of the body whereas during gait postural control must be released and movement allowed. Going from posture to movement means turning off the connections that assure stability and turning on those that allow movement. […] A second advantage of interneurones is that they can specialize in producing different patterns or rhythms of activity. This could be a special property of individual neurones or a property of an interconnected network of neurones, such as envisaged for the locomotor pattern generator.

Drawing a picture where spinal circuits are autonomous, but modulated, by cortical input.

t is a general fi nding that every single interneu-rone receives input not only from the sensory modality which is the basis for its classifi cation (e.g. as a ‘Ia inhibitory interneurone’, ‘Ib inhibitory interneurone’, ‘gr. II interneurone’ or ‘fl exor refl ex afferent interneurone’), but also from a number of other sensory afferent modal-ities, other interneurones and a number of descending pathways (e.g. corticospinal, vestibu-lospinal, reticulospinal).

It is not an unrealistic possibility that spinal interneurones with a slight turn of events could have been classifi ed based on their supraspinal input as taking part in different voluntary movements rather than the current clas-sifi cation based on afferent input as taking part in different refl ex actions. This was realized already by Sherrington (1906) more than 100 years ago when he wrote: “A simple refl ex is probably a purely abstract conception, because all parts of the nervous system are connected together and no part of it is probably ever capable of reaction without affecting and being affected by various other parts, and it is a system certainly never absolutely at rest. But the simple refl ex is a convenient, if not a probable, fi ction.”

**The discharge of every single motoneurone and thus the activation of every single muscle fi ber is determined by the integrated depolarization from on average 10,000 synaptic inputs arising from a number of different sensory modalities, spinal interneu-rones and supraspinal control centres.**

Cortical input to the spinal cord should be viewed as using or modulating the output of the spinal circuitry itself. There is no separation between ‘spinal refl exes’ and ‘cortical voluntary movement’. Instead, it is important to focus on how the neuronal machinery in the spinal cord may provide an extremely fl exible tool for the execution of voluntary movements.

We hypothesize that there are at least two advantages of cortical control. The fi rst is adaptability which emerges as a consequence of the anatomy of the cortical motor representation. The second is integration of visual input which is particularly important in shaping the hand to manipulate objects. Individuation and precision are secondary consequences of this organization.

Begs the question of a no-visual experiment?

EMG recordings show that very short synchronous bursts of activity are characteristic of many fractionated fi nger movements, such as writing and tool use […] Interposing interneuronal synapses in these connections would tend to remove synchrony and smooth out the command. This is indeed what is seen following corti-cospinal lesion (Farmer et al., 1993). The CM system may thus also be at the heart of human evolution in view of the evolutionary advantage of being able to throw something at an animal in order to kill and eat it.

One- third of the cortex, particularly in the parietal and premotor areas, is devoted to visual processing. A considerable part of this is used for shaping/orienting our hand in preparation for grasping and manipulating objects

the motor cortex may differ from the spinal cord in degree of fl exibility and a larger possibility of integrating visual input, but otherwise there are no differences between the motor cortex and the spinal cord circuitries (after all CM cells project to motoneurones and receive sensory input much like any good old- fashioned interneurone) that could warrant a signifi cantly different role in our conscious experience of control of the movements that we perform. What we are proposing is that it is not the degree of perceived volition which determines to what extent the motor cortex is involved in a given task, but rather the need for fl exible visual control.

We hope that we have made it clear that there is little to support this distinction between automatic and voluntary tasks. We need to consider the integration between supraspinal and spinal control centers for any specifi c task in order to understand how that task is controlled by the nervous system, and try to avoid putting into it volition and voluntary which in any case are terms that belong in philosophy or in specifi c inquiries aimed at unravelling the mechanisms of our cognitive abilities and our conscious experiences.

Compelling case to look at the entirety of the system, focusing on the contributions that cause motor neurons to fire.

# 3 Data Analysis

[CAPTION [CITATION].](images/data_analysis2023/behavior.png)

# 4 Preliminary Theory

*An interesting open question is how to relate trial-to-trial dynamics of learning to asymptotic predictions regarding optimal adaptation.*

— Todorov, *2007*

What are “normative models”?

Normative models suggest that the nervous system optimally adapts when faced with an error. To determine this optimal adaptation, the normative model must specify two key features of the world. First, how different factors, such as tools or levels of fatigue, influence the motor system — the so-called generative model. Second, how these factors are likely to vary over both space and time — that is the prior distribution. The structure of the generative model and the prior distribution together determine how the motor system should attribute an error to the underlying causes and, therefore, how it should adapt.[**wolpertPrinciplesSensorimotorLearning2011?**](#X766c804d862d147f44dc688b8c3ceb4178b2abd)

## 4.1 Optimal Feedback Control

The optimal feedback control framework remains the strongest normative model of human movement control. The first step of our theoretical work is to build from the simplest optimal feedback control models, working towards constructing our own variants of such models in order to capture aspects of our experimental findings. The first model we will investigate is the fully-observable, discrete-time, infinite-horizon linear quadratic regulator (LQR) problem with additive Gaussian noise. Given a state and an control input , the state evolves in discrete time according linear dynamics

where The LQR problem is to find a sequence of controls which minimize a cost

according to some chosen state and control cost parameters and . The optimal cost-to-go or state value function for the problem is

The optimal control law which solves this problem is linear function in the state

This control law is found through iterating a Riccati equation involving , , , and to find the optimal which is a quadratic function in state. There is a functional relationship between and that can be derived algebraically. Fig. [4](#fig:control_field) shows simulated trajectories of the infinite-horizon problem from random initial positions. The vector field depicts the two-dimensional control vector. In this example, the model describes a second-order point mass dynamical system where the control input acts as a force on the particle. The state vector contains two dimensions each of position, velocity, and force of the particle, each updated via Euler integration. Fig. [5](#fig:cost_field) shows the same simulated trajectories of the point mass atop the quadratic value function. The goal state in these simulations is (0.5,0.5). Note that there are many free dynamics parameters in such simulations within and which drastically alter the resulting simulations. The values here were chosen to match the motor control literature.

Figure 4: Simulation of trajectories from uniform random initial positions for the simplest LQR controller. The diffusions are controlled such that their inputs are proportional to the positional error. The plain LQR controller is invariant (up to a translation) to the goal state, as explained in the text. Here the goal state is (0.5,0.5) denoted by a white star. red circles denote the initial position of the trajectory and green circles denote the endpoint after 200 increments. Arrows show the state-dependent control signal (force) vector .

Figure 4: Simulation of trajectories from uniform random initial positions for the simplest LQR controller. The diffusions are controlled such that their inputs are proportional to the positional error. The plain LQR controller is invariant (up to a translation) to the goal state, as explained in the text. Here the goal state is (0.5,0.5) denoted by a white star. red circles denote the initial position of the trajectory and green circles denote the endpoint after 200 increments. Arrows show the state-dependent control signal (force) vector .

Figure 5: The same trajectory simulations as in Fig. [4](#fig:control_field) atop the quadratic cost field.

Figure 5: The same trajectory simulations as in Fig. [4](#fig:control_field) atop the quadratic cost field.

Assuming that the original system is controllable and stabilizable, the linear optimal control problem is a matter of trading off eigenvalues of the closed loop system which drive the steady-state error of the system to 0 under the control cost. This system can be made to drive to a given target by feeding back evolving the dynamics in terms of target error or by augmenting the state vector with terms which compute this error implicitly. These yield identical solutions for the control law, the latter typically being more convenient. The implication of this is that only one control law is required for the problem with a point target in the state space. If the problem were to model human movement, it might be said that we only require the internalization of a single control law up to a translation in the target’s desired state. Only the error is required in this simple case. As discussed in Section [5](#sec:next_steps), variants of this basic model are more interesting in that they do not display this target invariance.

## 4.2 Internal Model Adaptation for Linear Quadratic Control

In experiments within the setup described in Section ??, subjects are faced with a novel muscle-to-environment mapping that they must ostensibly learn in order to achieve their goals. Here I investigate the effects of approximating dynamics models within the LQR framework. This short experiment is a first step in modeling how subjects may use endpoint error in each trial to update or adjust their internal approximations of the environment’s dynamics.

Our state space is denoted and our control space where . Each trial, we move from state to x(N) in timesteps. Each trial, we have a goal state and a resulting endpoint error . We follow the same LQR setup as defined in the previous section. We can write the controlled, closed-loop system dynamics for the final time step :

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

This error may be due to several sources. Our internal dynamics model might have error relative to the true dynamic . Our control gain may be optimal relative to our internal model but not with respect to the true dynamic . Finally, we might have an approximate model and a suboptimal control gain . 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 is optimal for our approximate internal model (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 on trial :

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

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

Fig. [6](#fig:gradient_descent) shows the LQR simulations across gradient descent updates to the 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 dimensions of the state vector. Unfortunately, this optimization alters the dynamics incompatibly. The routine is also very fragile to parameter changes. This experiment highlights the difference in loss landscapes between the optimal control problem and the gradient descent simulated here. There are many directions for this work to proceed as discussed in Section [5](#sec:next_steps).

Figure 6: Iterations of gradient descent on the matrix of an infinite-horizon LQR where the original A is corrupted with Gaussian noise. Each dotted line is a sampled trajectory using a recomputed control gain with an updated matrix. Red circles denote the initial state, the star denotes the goal state, and the colored circles denote the endpoints of each trajectory sampled at each iteration. Note that the initial solution diffuses directly towards the target, and the gradient updates for the dynamics model alter this trajectory in a nontrivial way. As discussed in the main text, the gradient descent is optimizing for a different cost than the controller optimization, and thus this divergence might be expected.

Figure 6: Iterations of gradient descent on the matrix of an infinite-horizon LQR where the original A is corrupted with Gaussian noise. Each dotted line is a sampled trajectory using a recomputed control gain with an updated matrix. Red circles denote the initial state, the star denotes the goal state, and the colored circles denote the endpoints of each trajectory sampled at each iteration. Note that the initial solution diffuses directly towards the target, and the gradient updates for the dynamics model alter this trajectory in a nontrivial way. As discussed in the main text, the gradient descent is optimizing for a different cost than the controller optimization, and thus this divergence might be expected.

Learning LQR controllers

Q-learning for LQR  
policy gradient for LQR  
what is an LQR-SR? what does this help us do?

where does LQR break? - goal shift (is this true?) - task shift (different goal? this isn’t true) - goal uncertainty (this can’t be modeled…) - LQR variants break more easily?

where does KL-LQR / control break? - one policy at a time…? re-optimize your single policy per task? (task could have multiple goals) - could have multiple possible (terminal) goal states - not continuous – code this up and understand it in continuous would be a good result - selection is done beforehand… can this shift online…? - allows you to construct more interesting policies… - can we connect this to KL between passive and dynamic? change this to planned and replanned?

is there a multiplicative LQR composition? - sergey levine multiplicative paper?

# 5 Next Steps

*When it comes to the problem of skilled movement, the algorithm is simply not known.*

— Wolpert & Ghahramani, *2000*

## 5.1 EMG Hardware

Our preliminary data confirms the working principle of the setup and highlights the next steps for producing quality datasets. This is in accordance with the literature, where more advanced use of EMG is emerging as an important tool in understanding the complexities of motor computation[29](#ref-Hug2011). Our next steps are to build a new version of the EMG hardware doubling the number of channels to 64 electrodes placed across the forearm and to provide better hand constraints to ensure completely isometric contractions. The next hardware version will also include investments in shielding to provide proper noise mitigation. Most EMG in the literature is smoothed and trial-averaged due to noise, but we are confident that our records can be analyzed at the level of single trials, much like recent developments in neural data analyses.[30](#X77ca8e76944af68cc3f74dd2f0c86402f7e5cc2),[31](#Xfa55a75565b601f4c3f89bca569707e371d2725)

## 5.2 EMG Analyses

### 5.2.1 Preprocessing

Preprocessing of the EMG signal per-channel is another key area for improvement. EMG signal is the convolution of motor unit action potentials terminating near the electrode sites. Ideally we could filter each channel of raw EMG to infer the signal envelope. There is precedent in the literature for Bayesian filtering of EMG signal using a Laplacian distribution. Sanger used a Laplacian distribution to filter a one-dimensional EMG signal[32](#Xeac09f1e71d13d9ad6729947f274e364e95bf68). In accordance with this choice, Nazarpour found that as more motor units are recruited, the EMG distribution shifts from super-gaussian to gaussian following the central limit theorem[33](#Xae2caa71bd19493f31b49296b7a380b5051bb76). That work suggests methods for estimating high-order statistics for better filtering at low contraction levels relevant to our experiments. Such methods may prove to be more rigorous for inferring motor unit activations from our raw signal.

### 5.2.2 Calibration

With a filtered raw signal per channel, our goal is to devise mapping from EMG space to task space which are biophysically achievable by our subjects but require a degree of learning over many trials. In our preliminary task, we hardcoded mappings. Our next step would be to design a calibration task which asks subjects to actively explore the biophysical limits of EMG space that can be captured by our electrodes. This is akin to extracting features of spontaneous activity and passive viewing used in cortical BMI experiments to generate “intuitive” mappings.[34](#ref-sadtlerNeuralConstraintsLearning2014),[**Clancy2014?**](#ref-Clancy2014)

With high-dimensional EMG, we would ideally devise a principled method of extracting modes from raw EMG that accurately reflect modes of neural drive, demixing neural modes across channels. Here we used PCA as a first step. One starting point would be to align with the cortical BMI literature and use factor analysis and Kalman filtering[34](#ref-sadtlerNeuralConstraintsLearning2014). The issue, however, is how to design a task which evokes the available modes of possible EMG activity before using dimensionality reduction to generate learnable yet non-intuitive mappings.

## 5.3 Task Design and Data Collection

With a working calibration task, our next goal is to use this calibration data to generate mappings to track learning. Our center-hold, reach-out task combined these two steps into one task for validation of the setup using hardcoded mappings. We can maintain the center-hold, reach-out style of the task but with data-driven mappings, as well as construct novel tasks to test predictions of our models. Scaling this task up to multiple subjects across days will provide a dataset with which we can test hypotheses concerning the evolution of complexity and correlations across learning. This will also give us an opportunity to study inter-subject variability, for which there is precedent in the literature which suggests individual strategies[35](#X04cd25f3dc92c6f30b9469b98499a09359290a9).

## 5.4 Modeling Control

One relevant aspect of the basic optimal feedback control model is that the optimal controller that arises from specifying a quadratic state and control cost is invariant to the target state. In spite of this, we can use the aforementioned task to test predictions of the basic LQR model with respect to state and control noise and imperfect dynamics.

We expect to validate the basic optimal control models for our setup as we’ve designed the learning environment specifically for the EMG signal provided by the subject acts as the input to chosen virtual dynamics, which can be chosen in accordance with our model. We can then test perturbations in our task with respect to noise, goal, and dynamics and compare subjects’ responses to our models.

The question then becomes: when might subjects need to internalize a new control policy? When might they need to internalize multiple control policies? We hope to work towards answers of these questions alongside models of compositional control. Such control models could deal with, for example, target uncertainty as well as multiple competing targets[36](#ref-gallivanActionPlanCooptimization2015),[37](#Xa5065fcfd70d3e6e7b53602e39795554440bd4e).

## 5.5 Modeling Learning

Ultimately, our goal is to adapt optimal control models which begin as coarse approximations and are updated both within and across trials. Adaptation typically refers to online alterations to control policies while learning might refer to across-trial policy changes. Our theoretical aim is to devise models of learning and movement construction which extend the optimal feedback control framework through additions of composition and error-based updates.

Stemming from our work using simple gradient descent to update internal dynamics models, we would like to gain a better understanding of the loss landscape. It may be possible to compute the optimum analytically and to corrupt the dynamics matrix in a more principled way. We will also explore the action of the resulting gradient, and compute second-order derivatives, and compute derivatives with respect to the control law as a comparison. These results can then be compared with results from the reaching adaptation literature. This work can be guided by analyzing our empirical data to understand what aspects of our trajectories in EMG and task space are changing over trial.

## 5.6 Open Questions

* What is the best calibration task to find the boundaries of the available EMG space?
* What are the defining features of learning in our EMG-based task?
* Are our experiments well-modeled by the optimal control framework?
* How do subjects efficiently use error information from each trial and feedback from each time step to update their forward model and control policy/policies? How do subjects balance policy updates with model updates?
* How does a subject sample the state space to efficiently learn? Do subjects sample optimally?
* How can you empirically disentangle system identification (model estimation) and policy learning? If subjects are suboptimal, is it due to model mismatch or a suboptimal policy?

## 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. Kober, J., Bagnell, J. A. & Peters, J. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research* **32**, 1238–1274 (2013).

3. Sauerbrei, B. A. *et al.* Cortical pattern generation during dexterous movement is input-driven. *Nature* (2019) doi:[10.1038/s41586-019-1869-9](https://doi.org/10.1038/s41586-019-1869-9).

4. Bernstein, N. *The coordination and regulation of movements*. (Pergamon, 1967).

5. Kitano, H. Biological robustness. *Nature Reviews Genetics* **5**, 826–837 (2004).

6. 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).

7. 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).

8. 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).

9. Basmajian, J. V. Control and Training of Individual Motor Units. *Science* **141**, 440–441 (1963).

10. 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).

11. Giszter, S. F. Motor primitivesnew data and future questions. *Current Opinion in Neurobiology* **33**, 156–165 (2015).

12. Gao, P. *et al.* *A theory of multineuronal dimensionality, dynamics and measurement*. (2017) doi:[10.1101/214262](https://doi.org/10.1101/214262).

13. 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).

14. Ingram, J. N. & Wolpert, D. M. The statistics of natural hand movements. *Brain* **188**, 223–236 (2009).

15. 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).

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

17. Bruton, M. & O’Dwyer, N. Synergies in coordination: A comprehensive overview of neural, computational, and behavioral approaches. *Journal of Neurophysiology* **120**, 2761–2774 (2018).

18. 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).

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

20. 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).

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

22. 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).

23. 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](https://doi.org/10.1002/cphy.cp120106).

24. Graziano, M. The Organization of Behaviorial Repertoire in Motor Cortex. *Annual Review of Neuroscience* **29**, 105–134 (2006).

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

26. 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).

27. Watanabe, H. *et al.* Forelimb movements evoked by optogenetic stimulation of the macaque motor cortex. *Nature Communications* **11**, 3253 (2020).

28. Graziano, M. S. A., Aflalo, T. N. S. & Cooke, D. F. Arm Movements Evoked by Electrical Stimulation in the Motor Cortex of Monkeys. *Journal of Neurophysiology* **94**, 4209–4223 (2005).

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

30. Churchland, M. M. Neural Variability in Premotor Cortex Provides a Signature of Motor Preparation. *Journal of Neuroscience* **26**, 3697–3712 (2006).

31. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).

32. Sanger, T. D. Bayesian Filtering of Myoelectric Signals. *Journal of Neurophysiology* **97**, 1839–1845 (2007).

33. 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).

34. Sadtler, P. T. *et al.* Neural constraints on learning. *Nature* **512**, 423–426 (2014).

35. 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).

36. Gallivan, J. P., Barton, K. S., Chapman, C. S., Wolpert, D. M. & Randall Flanagan, J. Action plan co-optimization reveals the parallel encoding of competing reach movements. *Nature Communications* **6**, 7428 (2015).

37. Gallivan, J. P., Logan, L., Wolpert, D. M. & Flanagan, J. R. Parallel specification of competing sensorimotor control policies for alternative action options. *Nature Neuroscience* **19**, 320–326 (2016).

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. [↑](#footnote-ref-20)
2. In a classic study, Basmajian and colleagues showed that it is possible to activate single motor units in the thumb abductor. [↑](#footnote-ref-22)