IOP Publishing Journal of Neural Engineering
J. Neural Eng. 11 (2014) 056006 (21pp) doi:10.1088/1741-2560/11/5/056006

Useful properties of spinal circuits for learning and performing planar reaches

George A Tsianos, Jared Goodner and Gerald E Loeb

Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089, USA

E-mail: George.Tsianos@l-3com.com

Received 2 December 2013, revised 2 June 2014
Accepted for publication 5 June 2014
Published 1 August 2014

Abstract

Objective. We developed a detailed model of the spinal circuitry plus musculoskeletal system (SC + MS) for the primate arm and investigated its role in sensorimotor control, learning and storing of movement repertoires. *Approach.* Recently developed models of spinal circuit connectivity, neurons and muscle force/energetics were integrated and in some cases refined to construct the most comprehensive model of the SC + MS to date. The SC + MS's potential contributions to center-out reaching movement were assessed by employing an extremely simple model of the brain that issued only step commands. *Main results.* The SC + MS was able to generate physiological muscle dynamics underlying reaching across different directions, distances, speeds, and even in the midst of strong dynamic perturbations (i.e. viscous curl field). For each task, there were many different combinations of brain inputs that generated physiological performance. Natural patterns of recruitment and low metabolic cost emerged for about half of the learning trials when a purely kinematic cost function was used and for all of the trials when an estimate of metabolic energy consumption was added to the cost function. Solutions for different tasks could be interpolated to generate intermediate movement and the range over which interpolation was successful was consistent with experimental reports. *Significance.* This is the first demonstration that a realistic model of the SC + MS is capable of generating the required dynamics of center-out reaching. The interpolability observed is important for the feasibility of storing motor programs in memory rather than computing them from internal models of the musculoskeletal plant. Successful interpolation of command programs required them to have similar muscle recruitment patterns, which are thought by many to arise from hard-wired muscle synergies rather than learned as in our model system. These properties of the SC + MS along with its tendency to generate energetically efficient solutions might usefully be employed by motor cortex to generate voluntary behaviors such as reaching to targets.

Keywords: motor control, motor learning, reaching, spinal circutry, musculoskeletal system

1. Introduction

Even for seemingly simple tasks like locomotion, the underlying neural activity that controls the movement is quite complex and difficult to understand. Appropriate sequencing and modulation of muscle recruitment can be controlled entirely by the spinal cord through central pattern generators that make use of the abundant cutaneous and proprioceptive information that is continuously fed back to the spinal cord (McCrea and Rybak 2008, Stein 1978).

The spinal cord is clearly important for locomotor behavior, but its role in voluntary movements (such as made by the human arm) has long been debated. It has often been argued that most of the dynamics of muscle recruitment are preprogrammed and are specified explicitly in neural signals descending from various areas in the brain (Hallett *et al* 1975, Scheidt and Rymer 2000, Lillicrap and Scott 2013, Kargo and Nitz 2003). Such hypotheses are founded on past experiments showing that crisp sequencing of muscle activity occurred for ballistic movements in which there was insufficient time for feedback signals to influence the motor program via the brain (Hallett *et al* 1975). Scheidt *et al* (2000) had subjects perform two planar reaching movements to the same target, but in one case the shoulder

was constrained. Muscle activity was similar across situations, which led the authors to conclude that compensation of intersegmental dynamics was feedforward. Kargo and Nitz (2003) concluded that the brain controls the dynamics of groups of muscles or synergies directly based on correlational studies between brain and muscle activity. Such studies tend to overlook the local feedback circuits in the spinal cord, which are now known to be much more broadly connected than merely providing servocontrol loops for individual muscles (Pierrot-Deseilligny and Burke 2005). Efference copy and proprioceptive signals can influence even the fastest movements because they could be arising from postural muscles that are activated substantially prior to the onset of any overt motion (Aruin and Latash 1995, Lee 1980, Tyler and Karst 2004). Lillicrap and Scott (2013) incorporated state feedback into a neural network model of the motor cortex, which effectively assumes that such feedback is necessary but does not separate it into feedback occurring in cortical *versus* spinal loops, both of which certainly exist.

Any plausible theory of brain function in movement must account for the properties of the spinal circuitry plus musculoskeletal system (SC + MS). This study describes the most comprehensive model of the SC + MS for studying its contribution to sensorimotor control. Planar center-out reaching was studied because it is a popular paradigm with a large amount of validation data and because the musculature involved is kinetically redundant and its mechanics are influenced by intersegmental dynamics. Such complexities are present in MSs throughout the body and traditionally have been thought to pose major control challenges to the brain (Lackner and DiZio 1994). Using the model, the following questions were addressed to get insight into the mechanisms of controlling, learning, and storing movement repertoires:

(1) Can the spinal cord plus MS generate the physiological muscle dynamics of point-to-point movement? SC receives a rich set of proprioceptive feedback that it can flexibly combine through its highly interconnected network of neurons and descending input to generate highly modulated muscle activity. It is not clear, however, and it has not yet been demonstrated if such feedback and connectivity is sufficient to generate all of the dynamics of reaching movement that have been traditionally attributed mostly to brain activity. In order to determine the performance limits of SC + MS, brain commands were constrained to unmodulated step functions to force the SC + MS to generate all of the dynamics. Motor cortical activity associated with voluntary movements exhibits modulation before and during such movements, but it is generally not known to which spinal circuits (if any) such cortical neurons project or if their modulation represents outgoing commands or responses to sensory input.

(2) Can appropriate commands to the SC be determined easily through simple trial-and-error learning and appropriate cost functions? The redundancy, nonlinearity, and large number of control inputs of the SC + MS contribute to a large number of local minima of which some fraction corresponds to acceptable performance for a given task. This fraction depends on the collective properties of the SC + MS. A relatively large number of good-enough minima (compared to unsatisfactory local minima) would improve the probability and speed of learning good performance. The ease of finding commands that produce good-enough performance was assessed by employing a very simple learning algorithm. The feasibility and importance of incorporating energetics into the cost function were also assessed.

(3) Can commands to SC be interpolated to generate novel movement? In theory, any continuous system can be linearized over sufficiently small intervals and interpolation can be performed; however, it is not clear *a priori* how large this interval is for the SC + MS and if it is consistent with experimental observations. The extent of interpolability was investigated and implications for learning speed and memory capacity required for storing repertoires were assessed.

2. Methods

A simulation environment was created that incorporated detailed models of the SC + MS and an oversimplified model of the brain in order to force the SC + MS to generate all of the necessary dynamics. This model includes a refinement of the basic neuron presented in Raphael *et al* (2010) that incorporates scalability so that it is applicable to a larger range of MSs.

2.1. MS

The musculoskeletal model of the right arm has been described in detail in Tsianos *et al* (2011). Briefly, it consists of an elbow and shoulder joint that link the trunk (grounded), upper arm and lower arm. Each hinge joint is actuated by a pair of antagonist muscles that provide either flexion or extension torque in the horizontal plane. The model also has two biarticular muscles, one providing flexion and the other extension torques across both joints. Realistic muscle spindle models responding to muscle stretch and fusimotor control (Mileusnic *et al* 2006) and Golgi tendon organs responding to muscle tension (Mileusnic and Loeb 2009) provided continuous proprioceptive feedback to the model of SC.

2.2. Model of SC

Our model of the SC is called a spinal-like regulator (SLR) by analogy to multi-input–multi-output engineered systems that appear to have similar functionality (He *et al* 1991). The interneuronal pathways that comprise the model are detailed

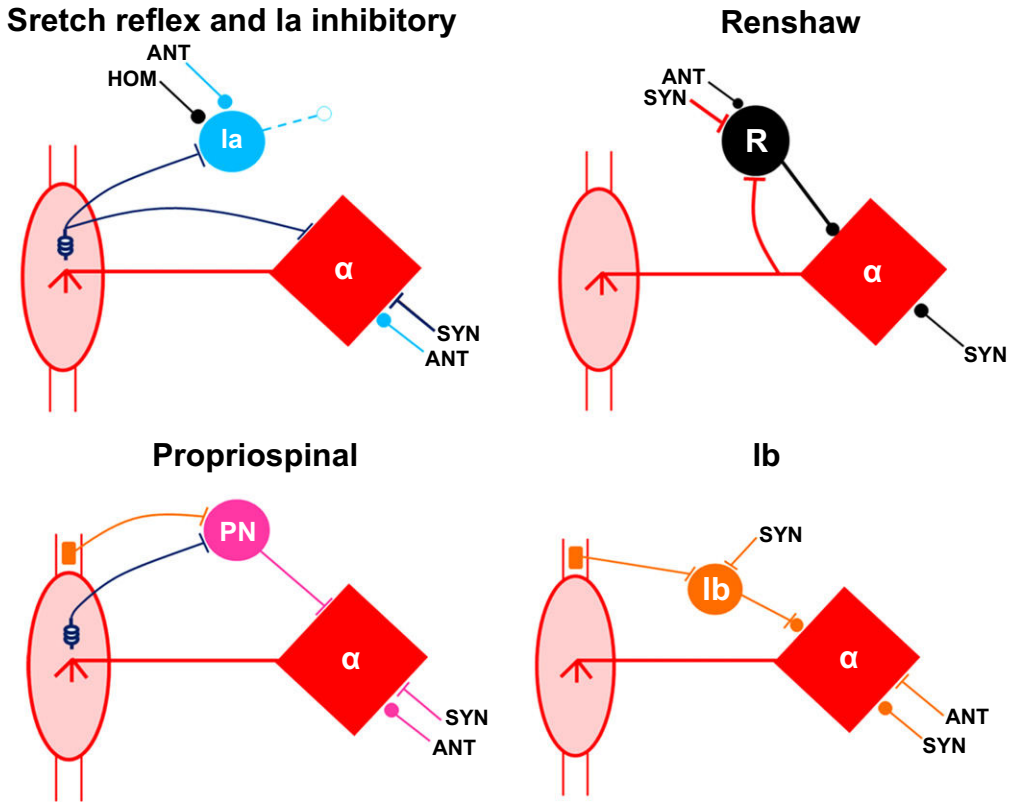

Figure 1. Connectivity model of spinal circuitry. Five classical interneuronal pathways that comprise the model from the perspective of a single muscle. Projections from neural elements associated with self (HOM) as well as synergist (SYN) and antagonist (ANT) muscles are shown.

in Raphael *et al* (2010), while their specific connectivity for the elbow–shoulder MS is described in Tsianos *et al* (2011). Briefly, the model of SC is based on a set of connectivity rules that have been derived from a combination of electrophysiological and anatomical studies (Pierrot-Deseilligny and Burke 2005). As shown in figure 1, the neural network of the spinal cord has been divided into five classical pathways as defined by their homonymous and heteronymous connectivity, which is different among synergists and antagonists. The monosynaptic Ia pathway excites alpha motoneurons directly while the rest of the modeled pathways influence motoneuron activity through the following interneuronal types: propriospinal, reciprocal Ia-inhibitory, Renshaw inhibitory and Ib-inhibitory. Synergist and antagonist projections are such that a given signal from a particular muscle has the same effect on synergist muscles and the opposite effect on antagonist muscles. The circuitry was identified typically between pairs of muscles that are functional synergists or antagonists and has been adapted for the more common case where a pair of muscles can have either relationship depending on the nature of the task. We call this relationship partial synergist, which is thought to be mediated by SC having both synergist and antagonist projections. The monoarticular muscles at each joint are modeled as strict antagonists while all other muscle pairings in the system are modeled as partial synergists. Even muscles that do not act across the same common joint need to be coordinated as partial synergists as a consequence of the intersegmental dynamics that they affect (Graham *et al* 2003, Zajac 1993). Descending commands (representing the lumped influence of many supraspinal neurons) project directly to the interneurons (but not the alpha motoneurons) and also modulate the gains of feedback pathways, corresponding to presynaptic modulation in the biological system.

2.3. Model of the neuron

The motoneuronal output of SC that excites muscles in response to input signals from the brain and periphery depends not only on the connectivity of the circuitry but also on the biophysical properties of the constituent neurons. Neuronal axons influence the output spike train of their target neuron through a series of processes. Prior to reaching the synapse, activity within a given axon is modulated at the presynaptic terminal by spinal circuits that are themselves modulated by descending neural input (Rudomin and Schmidt 1999, Carroll *et al* 2005, Chen *et al* 2002). The resulting activity then causes current to flow into or out of the target cell through the process of synaptic transmission. The input current is accompanied by a change in membrane potential and can lead to output spiking activity depending on the properties of the cell (see figure 2 for a schematic illustrating these processes and their computational implementation). Our approach to modeling

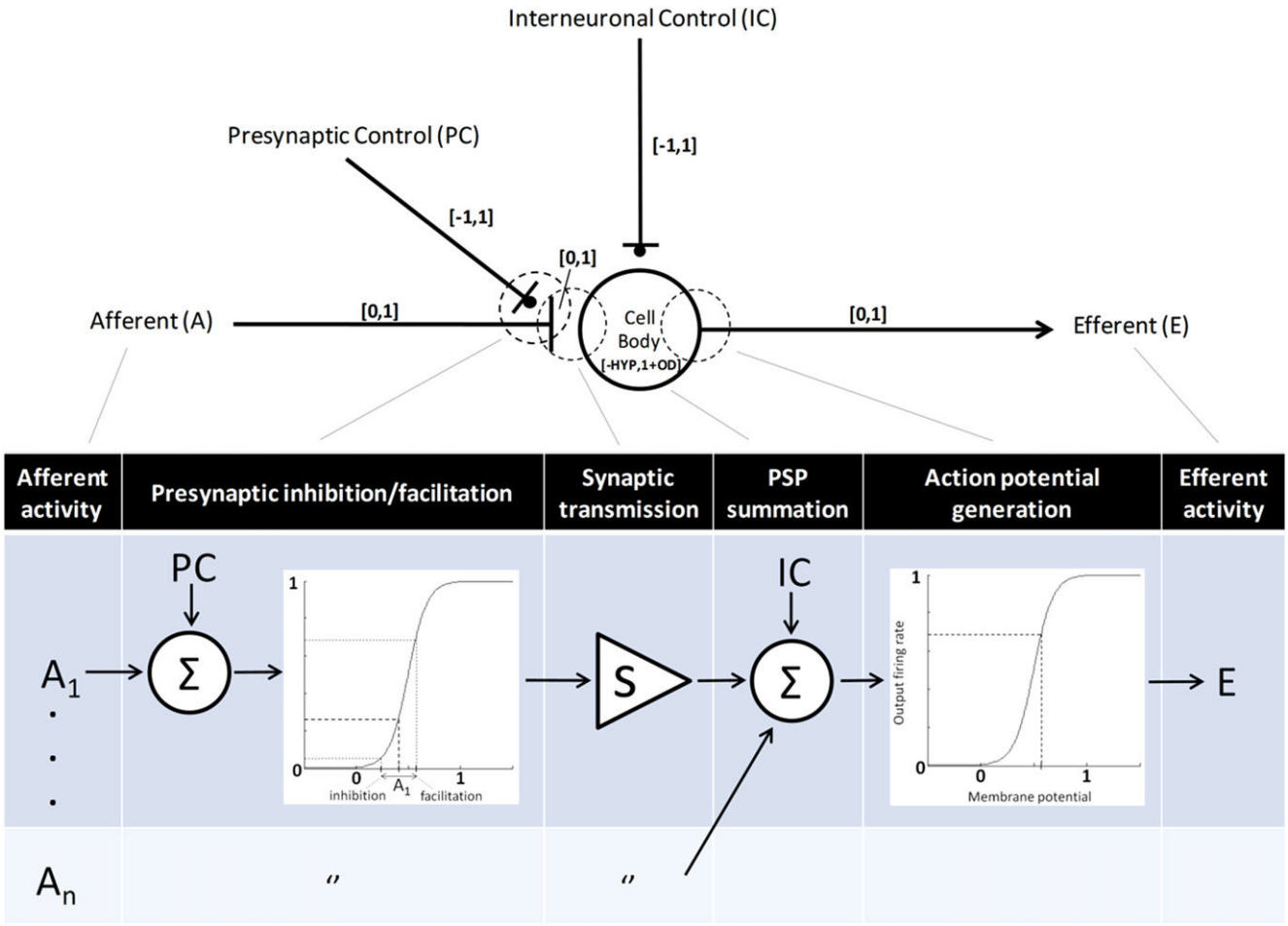

Figure 2. Model of the neuron. Schematic of the major physiological processes contributing to interneuronal output and corresponding block diagram that illustrates their computational implementation. Note that the gain 'S' in synaptic transmission is positive for excitatory input and negative for inhibitory input. This model also applies to motoneurons but with interneuronal control (IC) equal to zero. Refer to table 1 for a complete list of symbols and their definitions.

neurons aims to capture only the major computational properties while minimizing the number of arbitrary parameters and computational load.

2.3.1. Firing rate versus explicit spikes. In general, the higher the firing rate of a given neuronal input, the larger the effect on the firing rate of the target neuron's output. The higher the firing rate, the more neurotransmitter is released at the synapse over time, leading to more pulses of synaptic current entering the cell. If these currents are excitatory, for example, then sodium ions will flow inside the cell, depolarizing the membrane capacitance. Between synaptic events, the transmitter-gated ion channels close and active sodium–potassium pumps attempt to restore the resting membrane potential. At high enough rates of net excitatory synaptic input, the membrane potential will eventually exceed threshold, causing an action potential to occur. Following the action potential, the axonal membrane undergoes a refractory period where it is hyperpolarized to a large degree, making it unexcitable. The membrane potential gradually rebounds requiring progressively lower input currents to exceed its threshold and fire another action potential. Therefore, the larger the excitatory input current, the more depolarized the membrane potential, and the higher the firing rate of the neuron until it approaches the inverse of the refractory period, whereupon it saturates. This relatively simple transduction property is prevalent in neurons of the spinal cord, which generally lack dendritic spines or other histological features that have been associated with temporal sensitivity to local patterns of synaptic current such as in cortical neurons (Jadi *et al* 2012). Thus a simple mathematical relationship was used in this study that computes axonal output by summing the firing rates of all input axons rather than integrating post-synaptic potentials generated by individual spikes. It is worth noting that many classes of neurons in the brain exhibit subthreshold oscillations of membrane potential and are therefore more sensitive to intermediate frequencies of stimulation whose bandwidth can be very narrow. To predict the output of such neurons, it is important to account for the timing of individual spikes. Spike timing is also important when synaptic contacts are located close to each other. If a given synapse on a dendrite is active and its corresponding ion channels open, for example, subsequent

activation of an adjacent synapse would produce minimal changes to the neuron's membrane potential because the open ion channels nearby would create a shunt, thus, reducing the local input impedance (Segev *et al* 1990, Jadi *et al* 2012). This effect is not captured in our model because of the lack of data on the location of individual synapses in the spinal cord. There is also a lack of evidence for systematic spatial patterns of projections for a given population of neurons (McKeon *et al* 1984) suggesting that even if these spike-timing effects existed at the neuronal level, they would be averaged out by the population.

2.3.2. Presynaptic inhibition/facilitation. The spike train within the incoming axon will release different amounts of neurotransmitter in the synaptic cleft depending on the level of presynaptic inhibition/facilitation it experiences. This modulatory effect is mediated by neuronal (axoaxonal) projections to the presynaptic terminal whose electrical excursions affect the influx of Ca^{2+}, hence vesicle fusion and neurotransmitter release. These presynaptic gains are all assumed to be under cortical control, at least indirectly (Baudry *et al* 2010), and they constitute the majority of the inputs from the brain model (see section 2.4 'Model of the brain').

The maximal firing rate of a given axon is finite and depends on the duration of the action potential plus refractory period. For simplicity, maximal firing was set to 1 in the model. Presynaptic control (PC) input ranged from −1 to 1 to allow for either inhibitory or facilitating effects. The resulting signal represented the effective activation at the presynaptic terminal and lay within the range 0 to 1. This was accomplished by adding the afferent and PC signal and feeding their sum to a sigmoid function having a range of 0 to 1 (see figure 2).

2.3.3. Synaptic transmission. Neurotransmitter molecules from the presynaptic terminal diffuse across the synaptic cleft and bind to ligand-gated ion channels on the post-synaptic membrane. Ion channels open upon binding and allow the flow of specific ions along their electrochemical gradient, thereby generating synaptic current and altering the membrane potential. The extent to which the membrane potential changes depends on the input impedance of the cell, which further depends on the size of the cell and location of the synaptic contact. This simple, distributed model of cellular electrophysics seems to capture the most salient properties of those spinal neurons that have been studied (Segev *et al* 1990, Fleshman *et al* 1988), although it is clearly inadequate as a computational model of many neurons in the brain that have local circuit features such as dendritic spines (Jadi *et al* 2012, Polsky *et al* 2004).

2.3.4. Scaling for interneuron size. The six-muscle system described in this paper has almost twice as many synaptic inputs to each neuron as the four-muscle system studied by Raphael *et al* (2010). SLRs for other MSs will have yet other connectivities. In preliminary work, we discovered that the SLR models tended to become difficult to train or even unstable if the maximal post-synaptic effect that could be generated by a given synaptic input were too large or too small. This led us to introduce the following generalized scaling function, which is based on simple membrane biophysics.

The effect of an input current on the membrane potential of a given neuron is automatically scaled in the biological system according to the size of the cell, as first described for alpha motoneurons of the spinal cord (Henneman *et al* 1965). This is because each individual synapse acts as a constant current source while the complete set of synaptic inputs acts as a set of parallel load resistors and capacitors. From Ohm's law, the change in the membrane potential of the post-synaptic cell must be the product of synaptic current times the effective input impedance of the cell, which will scale approximately inversely with the total number of synapses. Similarly, the influence of each input on the model cell was scaled by a factor 's' (see figure 2) that depended on the total number of inputs to the cell. Rather than 's' having an arbitrary relation to the total number of inputs, it can be framed in terms of the maximal allowable hyperpolarization and overdrive of the neuron. These parameters are probably regulated in the nervous system through proper balance of excitatory/inhibitory inputs and input impedance to ensure that the membrane potential lies within its operating range and away from the saturation limits; otherwise, neurons with many inputs would tend to saturate easily.

In the model, a neuron receiving n_{exc} excitatory inputs and n_{inh} inhibitory inputs (each having a range of 0 to 1) that are not scaled (i.e. $s = 1$) could experience a change in membrane potential ranging from n_{inh} to n_{exc}. For this case, neurons with more inputs would tend to experience a larger change, thus being more likely to push the membrane potential beyond the limits that saturate the output firing rate. For a membrane potential operating range defined as 0 to 1 for simplicity, and maximal overdrive '*OD*' and hyperpolarization '*HYP*' limits (i.e. full range being –HYP to 1 + OD), it was necessary to scale the inputs by the following constants:

$$s_{inh} = -\frac{HYP}{n_{inh}};$$
$$s_{exc} = \frac{1 + OD}{n_{exc}}.$$

It was determined empirically that $HYP = OD = 2$ applied to all neurons in the model facilitated training of the model system and produced the most physiological results.

2.3.5. Action potential generation. Depending on the intrinsic dynamics of the cell and resistive paths between individual synapses and the cell body, incoming synaptic currents are integrated into a cell membrane potential which is then converted to a spike train depending on the threshold and refractory properties of the axon potential initiation site near the axon hillock.

Table 1. Symbols and definitions.

Symbols	Definitions
Output	Axonal output activity of a given cell (0–1)
A	Neural activity within afferent axons from proprioceptors, interneurons, or alpha motoneurons prior to presynaptic modulation (0–1)
PC	Descending drive that presynaptically upregulates (>0) or downregulates (<0) the nominal activity of a given axon (−1–1)
IC	Descending drive that controls interneuronal activity directly (−1–1)
exc_input	Magnitude of total axonal activity after presynaptic modulation that excites the target neuron (0–1)
inh_input	Magnitude of total axonal activity after presynaptic modulation that inhibits the target neuron (0–1)
s_{exc}	Scaling factor applied to exc_input to account for synaptic transmission (au)
s_{inh}	Scaling factor applied to inh_input to account for synaptic transmission (au)
n_{exc}	Total number of excitatory inputs to target neuron
n_{inh}	Total number of inhibitory inputs to target neuron
OD	Maximum amount a target neuron can be overdriven (relative to input that leads to output saturation)
HYP	Maximum amount a target neuron can be hyperpolarized (relative to input that leads to output saturation)
sig	Smooth saturation function that maps total input to a neuron into its axonal output and modulatory effects of presynaptic control onto incoming axonal activity

Each of the interneurons in the SLR represents a population of similarly connected but asynchronously active interneurons in the spinal cord. Neurons within a population having a variety of membrane potential thresholds and firing sensitivities collectively have a sigmoidal response (Eliasmith and Anderson 2003). As mentioned in 'Presynaptic inhibition/facilitation', the range of axonal firing rate was defined as 0 to 1 for simplicity. The sigmoid, therefore, accepted the combined effect of inputs on membrane potential (see 'Synaptic transmission') and produced a value from 0 to 1. The sigmoid function has the physiological attribute of nonlinearity in that a given change in input leads to different changes in output depending on the level of background activity, which defines the position of the system on the sigmoidal input–output function.

The output firing rate of a given neuron in the model was computed using the equations below (refer to table 1 for a complete list of symbols and their definitions). Besides some number of excitatory and inhibitory projections to interneurons that arise either from proprioceptive afferents, other interneurons and alpha motoneurons, there is also an input called interneuron control (IC) that represents the lumped influence of all descending signals on that interneuron. This value ranges between −1 and 1, as the lumped effect of descending signals can be either excitatory (the likely sign of direct corticospinal projections) or inhibitory (reflecting effects mediated by inhibitory interneurons under cortical control).

$$\text{output} = \text{sig}\left(s_{exc} * \sum_{i=1}^{n_{exc}} \text{exc\_input}_i + s_{inh} * \sum_{j=1}^{n_{inh}} \text{inh\_input}_j + \text{IC} \right),$$

where,

$$\begin{aligned} \text{sig}(x) &= \frac{1}{1 + e^{-11(x-0.5)}}, \\ \text{exc\_input}_i &= \text{sig}\left(A_i + \text{PC}_i\right), \\ \text{inh\_input}_j &= \text{sig}\left(A_j + \text{PC}_j\right). \end{aligned}$$

In the biological system, there is thermodynamic noise that affects proprioceptor transduction, synaptic transmission, as well as action potential generation and propagation. Muscle contraction is also subjected to noise. Noise was not accounted for in the model because it has not been well characterized and it was assumed that it would not alter the main results and implications of the work. Even if noise is substantial at the level of the single neuron, it is largely attenuated at the population level (as modeled here), because noise across individual neurons is poorly correlated. The effects of noise on the actual movement generated would be attenuated further by the inertia of the limb and damping properties of muscle. If significant noise was added to the model, it would bias the system towards solutions that distribute neural activity across more spinal circuits so that noise in any given neuron would have less of an effect on the overall performance. Spinal circuits would advantageously be activated in a way that increases cocontraction to stiffen the joints and therefore limit the effects of muscle force fluctuations on movement.

2.4. Model of the brain

The outputs of the brain model controlled fusimotor gain of muscle spindles, gain modulation of each presynaptic terminal and biasing activity of interneurons. The modeled outputs did not project directly to alpha motoneurons, which is consistent with the sparse cortical projections to the proximal limb musculature modeled in this study (Rathelot and Strick 2009). An oversimplified model of the brain was employed in order to force the SC + MS to generate all of the necessary dynamics. All commands to the SC were unmodulated step functions. A step input, initiated at the beginning of the simulation, was applied to each controllable element to set the background activity of the SLR prior to making a reach. The inputs projecting to the interneurons (interneuron control (IC)) were subjected to an additional step function whose onset was coincident with the simulated 'Go' cue of the movement. The amplitudes of all step functions (total of 438) were tuned by a coordinate descent optimization method.

2.4.1. Learning command programs. The amplitudes were initialized to random, previously learned, or interpolated values and tuned individually in a randomized sequence until

performance plateaued. Each amplitude was perturbed in the positive and negative direction by a predefined magnitude and was then set to the value that produced the best performance. Adjustment of a single amplitude marked one iteration and cycling through all amplitudes once corresponded to a learning cycle.

The model was trained to perform a variety of center-out reaches. The squared deviation from the desired state, integrated over the entire duration of the simulation, constituted the kinematic cost (in m^2 s units). The energetic cost (in Joules) was the sum of metabolic energy consumed by all muscles in the set over the entire simulation. The energy consumed by each muscle was computed using the validated model presented in Tsianos *et al* (2012). The nervous system presumably estimates energy consumption by combining information it receives from the contracting muscles such as chemoreceptor responses to metabolites with efference copy signals from their motoneurons. This information is likely integrated with additional sensory information related to heart rate, respiratory effort, body temperature, and perspiration. The acceptable limits of kinematic and energetic cost were determined empirically such that a solution within these limits tended to produce both a nearly straight reaching path that stabilized on the target and muscle activity that had a phasic burst pattern marked by low levels of cocontraction.

Initially, the model was trained to reach in the outward left direction (see figure 3). The endpoint had to follow a straight-line path to a target located 10 cm away in 0.5 s and maintained position for 1 s. Due to the complexity and nonlinearity of the system, initializations could not be determined *a priori*, other than setting the gains low initially to avoid instability. All amplitudes were initialized to random values within a range spanning −0.2 and 0.2. The perturbation magnitude applied to each amplitude was 0.2, 0.2, 0.1, and 0.1 for four cycles, respectively; performance typically plateaued by this point. The amplitudes were perturbed by large amounts in the beginning to avoid entrapment in regions of the solution space marked by poor local minima and were subsequently reduced to hone in on the local minimum. This is essentially an annealing function whose details have only modest effects on convergence and learning rates (Raphael *et al* 2010). Multiple initializations were used for each task. The system was first trained using a purely kinematic cost function and if converged performance was acceptable, training was continued with energetics added to the cost function. Kinematic cost and energetic cost use different units; it was determined through systematic analysis that scaling the energetics term by a factor of 2×10^{-4} caused them to have comparable effects on the learning process. Converged solutions for this particular reach were sometimes used as starting points for other reaches (see next section and section 3 'Results').

2.4.2. Interpolating command programs from existing solutions. In the biological system, command programs for useful tasks are stored for subsequent use, but the mechanism by which this occurs is poorly understood. The plausibility of

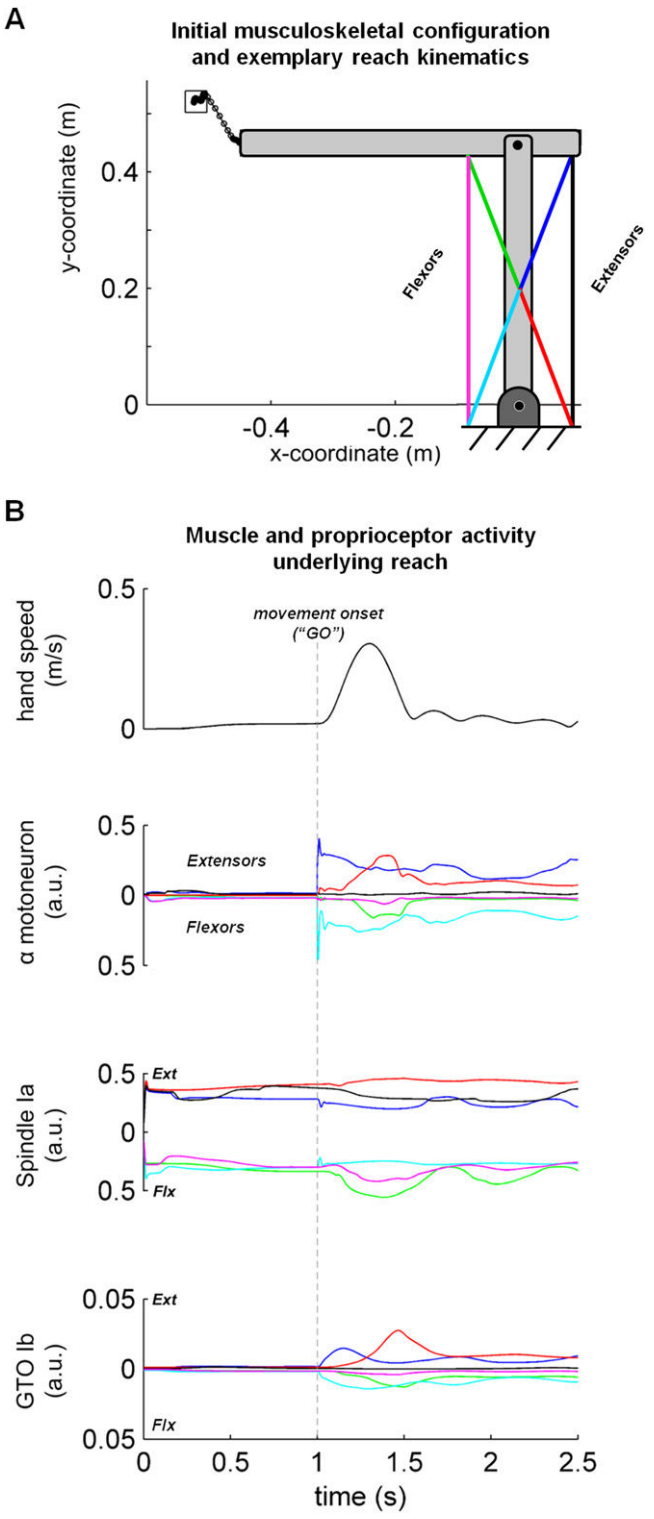

Figure 3. Exemplary reach to the outward-left target and underlying neural activity. (A) Reaching path in Cartesian space with circles spaced 50 ms apart in time. The overlaid diagram shows the initial posture of the arm. Colored lines represent muscle paths that correspond to the colored traces in the plots below. (B) Hand speed, alpha motoneuron, muscle spindle Ia, and Golgi tendon organ Ib activity. The vertical dashed line through these plots at 1 s marks the onset of the 'GO' signals that initiated the movement. Activity related to extensor muscles is plotted in the upward direction and downward for flexors.

a look-up table has been dismissed largely because the brain does not have enough storage capacity to memorize motor programs for all possible movements. As pointed out by Loeb (1983), however, instead of memorizing motor programs for all movements to a high degree of accuracy, it is more likely that the brain makes use of its finite memory capacity by allocating memory toward movements that are most important to the organism and are performed most frequently.

The storage capacity required for a lookup table would be reduced substantially if motor programs to previously learned tasks would generalize over a large range; or, equivalently, if motor programs for a relatively small set of movements could be combined in a simple manner to generate a larger set of motor programs. It is important to remember that the computational properties of the interneurons and the mechanical properties of the MS are all highly nonlinear. There is no *a priori* likelihood that any interpolation, much less a linear one, will be a useful strategy other than the general notion that any continuous system tends to be linear over sufficiently small parts of its operating space. The following method was used to explore the interpolability of command programs applied to the realistic model of the SC + MS. Interpolability was assessed for various task parameters including direction, distance, speed, and magnitude of a viscous curl field perturbation.

2.4.2.1. Assessing potential and limits of interpolability. Each solution plus corresponding cost can be fully described by a set of 439 scalar quantities (438 step function amplitudes plus 1 scalar cost), which can be represented as a point in 439 dimensional space. Together, all possible points form a hypersurface called the solution space. The likelihood of interpolating behavior that is intermediate to two tasks (T1 and T2) depends on the location of the solutions to these tasks (S1 and S2) in the solution space as well as the structure of the solution space in that local region. Because the SLR plus musculoskeletal plant is a highly parameterized and nonlinear system, the solution space is necessarily nonlinear and presumably has a large number of local minima. It was therefore hypothesized that pairs of solutions S1 and S2 that are located relatively close to each other in the hyperspace can be used to interpolate intermediate tasks with a higher likelihood of success. The first set of simulation experiments addressed the system's potential for interpolation by ensuring that solutions to T1 and T2 (S1 and S2) were as close as possible. The second set examined the limits of interpolation, which occurred when solutions lay relatively far from each other in the solution space.

To improve the likelihood that solutions would interpolate, we trained the system to learn pairs of tasks (T1 and T2) in a specific sequence. T1 was learned from a randomly initialized set of inputs while T2 was learned with the solution of T1 as a starting point. To further ensure interpolability, the learning algorithm's search space (i.e. perturbation size used when adjusting parameters) was as small as possible while still enabling acceptable performance for T2. In preliminary research, the appropriate set of perturbation magnitudes applied to the amplitudes during learning was determined to be 0.1, 0.1, 0.05, 0.05, respectively for each cycle. These values represent an annealing curve whose magnitude is half that used when learning from randomly initialized SLR gains.

Table 2. Interpolation test summary. List of reach parameters that were tested for interpolability as well as their specific values that were learned (T1 and T2) and interpolated (Ti).

	Learned task 1 (T1)	Learned task 2 (T2)	Interpolated task (Ti)
Reach direction (deg)	135	90	112.5
Reach distance (cm)	10	15	12.5
Flight duration (ms)	500	1000	750
Curl-field mag (N s m^{-1})	0	6.5	3.25

Task T1 was always the same while T2 varied depending on the reaching parameter that was tested for interpolability. T1 was a reach in the outward-left direction (135° with respect to the horizontal axis; see figure 1) over a 10 cm distance, with a flight duration of 500 ms. Reach direction, distance, duration and curl-field magnitude parameters of T1 were varied to generate T2. An intermediate task, Ti, was defined according to parameters that were halfway (linearly) between T1 and T2. Interpolability was assessed by variously weighting all step function amplitudes associated with T1 and T2 and comparing the resulting cost with respect to the mean parameters of Ti. See table 2 for a summary of the reaching parameters tested for interpolation and the learned tasks whose solutions were used to generate them.

The viscous curl-field consists of perturbing force applied to the endpoint whose magnitude is proportional to hand velocity and direction is orthogonal to the direction of motion (Mattar and Ostry 2007). The curl-field direction used in this study was chosen arbitrarily to be in the clockwise direction, i.e. it was 90° with respect to hand motion in the clockwise direction. Reaching in the midst of a curl-field is a popular experimental paradigm for studying motor learning. It is widely hypothesized that the brain builds an internal model of the field and computes the dynamic compensatory commands that are necessary to counteract its effects (see for example Conditt *et al* 1997).

To test the limits of the system's interpolability, interpolation was attempted using learned solutions located far from each other in the solution space (or more precisely, were near different local minima). One way to achieve this condition was to interpolate solutions from the same training sequence whose directions are far apart (e.g. 90° versus 45°). Another way was to interpolate two solutions for the same reach that were learned from different random initializations.

2.4.2.2. Interpolation method. In theory, there are many ways that the solutions of the two learned tasks can be combined to obtain the solution to an intermediate task. We

used a very simple method in which all amplitudes were swept simultaneously within the range of their respective learned values and the optimal setting was determined.

The interpolation method was implemented as follows:

(1) Solutions (S1 and S2) were obtained for a pair of learned tasks (T1 and T2) that differed only in terms of one reach parameter (see previous section).
(2) For each corresponding amplitude in S1 and S2, i.e. S1 ($amplitude_k$) and S2($amplitude_k$), a set of n interpolated values that evenly spanned the range of S1($amplitude_k$) to S2($amplitude_k$) were computed. This generated a set of *n* interpolated solutions that were all candidates for the new task (Ti) to be performed.
(3) Model performance was computed for all *n* solutions and the solution producing the best performance (i.e. lowest cost) was determined.
(4) If the performance of the best solution was at least good-enough (see section 2.4.1 'Learning command programs'), then the new task Ti was deemed interpolable.

The number of interpolated solutions to be tested, *n*, was chosen to be sufficiently high to eliminate the possibility that an interpolated solution that led to good-enough performance was not tested. It was determined empirically that a value of 30 was appropriate. Note that even if an interpolation function that is appropriate consistently across all situations is not found, the fact that at most only 30 trials would be necessary to achieve acceptable performance makes this an extremely efficient learning process compared to starting from scratch. If such a space were nonlinear but monotonic, it could be explored by successively halving the interpolation factor *n*; a successful solution would necessarily be found in five or fewer trials. If learning is rapid enough, unused movements might be 'forgotten' and then regenerated upon demand by interpolation and relearning faster than can be detected by psychophysical experiments, which typically require averaging of many trials.

3. Results

3.1. Exemplary solution to a reaching task

The SLR model was trained successfully to reproduce the dynamics of a rapid point-to-point reach (figure 3). This requires a tri-phasic burst pattern of muscles in order to generate agonist torques to accelerate the limb toward the target, antagonist torques to decelerate it and corrective torques to stabilize it on the target. Note that temporally modulated muscle activity must be produced by the SLR in response to unmodulated step functions applied only to the interneurons; thus, precise timing of the excitatory bursts is generated entirely by distributed proprioceptive feedback and its nonlinear integration by spinal circuits. Figure 3 shows the dynamical behavior of all feedback to the SC model, namely muscle spindle Ia, GTO Ib and alpha motoneuron activity, which inhibits itself via the Renshaw pathway.

3.2. Learning curves

The coordinate descent algorithm tended to converge on acceptable kinematic performance after three to four cycles (i.e. 1300–1700 iterations; typical learning curve in figure 4, left). In this particular case, but not always (see 'Variability of performance'), the energetic cost of the movement was also within acceptable limits even though energetics were not included in the cost function.

A more compact way to visualize the learning curve is introduced in figure 4 (right) that represents performance as a point in 2D space, showing how kinematic and energetic behavior coevolved with time. Small open circles in the figure are spaced one learning cycle (438 iterations) apart, that is, the trajectory bounded by a given pair of small circles corresponds to the evolution of performance over one cycle of the algorithm. From this perspective, the goal is to guide the trajectory rapidly into a good-enough region in which kinematic and energetic cost are both acceptable.

3.3. Variability of performance

The system converged to good-enough performance from many different Monte Carlo initializations. Figure 5(A) shows learning curves for eight representative initializations in which only kinematic cost was considered. All of these learning trials converged eventually to good-enough performance; however, the energetic quality of the solutions was not consistent. About half of the converged solutions had a reasonable energetic cost while the rest were energetically wasteful, exhibiting high levels of cocontraction. A similar result was seen for 2DOF wrist movements generated by an SLR (Raphael *et al* 2010). The system's tendency to converge to energetically efficient solutions despite the use of a purely kinematic cost function suggests that the structure of the SC may favor these types of solutions, probably as a consequence of the large number of reciprocal inhibitory circuits.

3.4. Effects of energetics on learning

To investigate whether the SLR can be trained to generate acceptable kinematic behavior but with physiological patterns of muscle excitation more consistently, the converged solutions were tuned after incorporating an energetics term into the cost function. This reduced the energetic cost of the solutions substantially while maintaining acceptable kinematic performance (figure 5(B)). The evolution of muscle recruitment strategies is shown on the sides of figure 5(B) for two exemplary situations: one in which the initial energetic cost is marginally acceptable (right) and another in which it is clearly excessive (left). The resulting muscle excitation patterns are marked by less cocontraction and exhibit a phasic burst pattern more clearly.

Training the system with a purely kinematic cost initially and subsequently incorporating energetics ultimately led to acceptable kinematic performance with physiological patterns of muscle excitation. Training with a combined cost function from a random initialization generally led to worse kinematic performance (results not shown), which suggests that varying

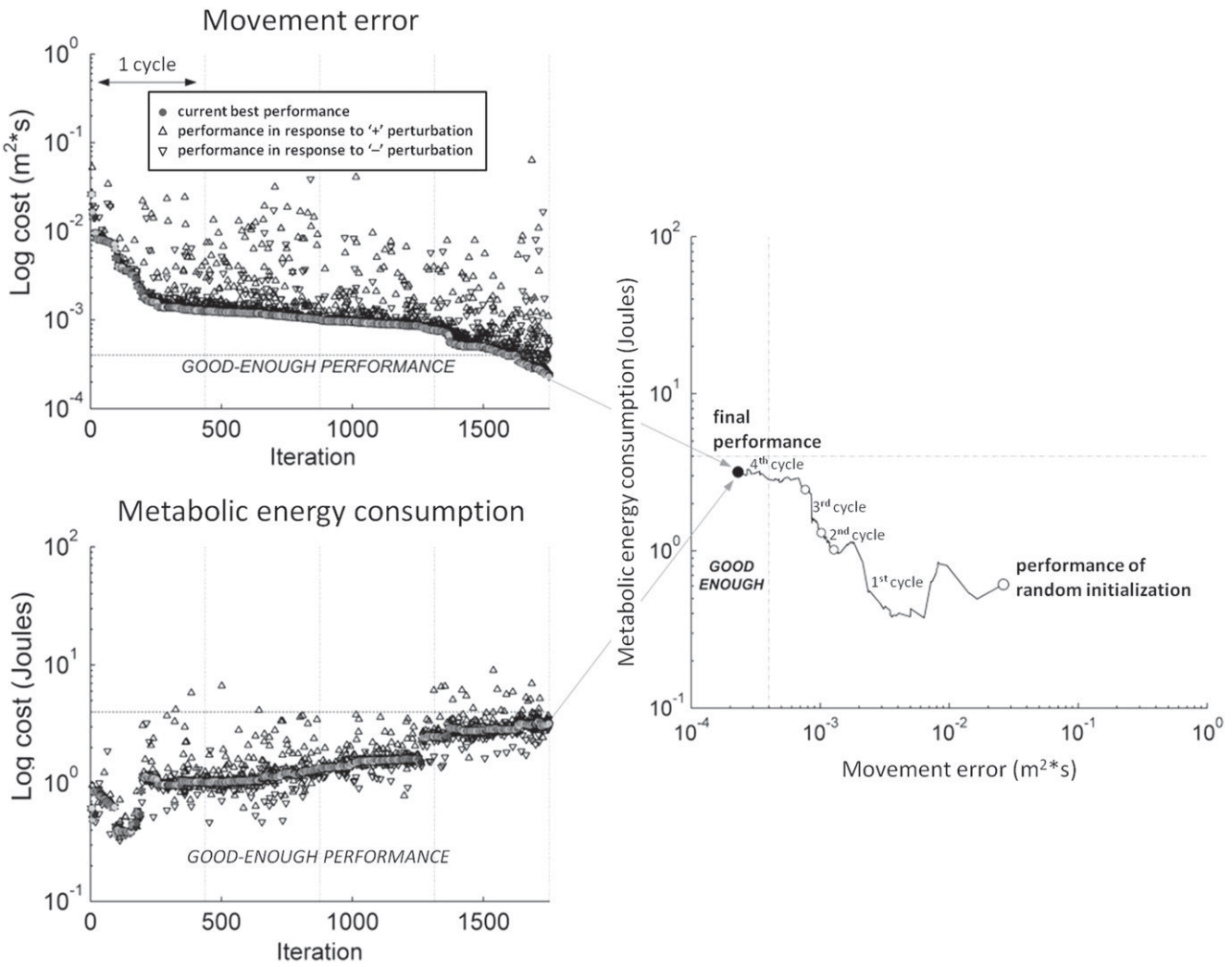

Figure 4. Exemplary learning curve. Evolution of kinematic and energetic performance over four cycles of the optimization algorithm (438 iterations per cycle). The plots on the left show the evolution of movement error and metabolic energy consumption separately. Filled circles represent the best performance achieved until that point in the learning trial while open triangles represent performance at that state when a randomly chosen input was perturbed in the positive direction (upward triangle) or in the negative direction (downward triangle). Dotted horizontal lines correspond to acceptable performance. The plot on the right shows energetic versus kinematic cost, with open circles being one cycle apart. Good-enough performance is marked by the region where both energetic and kinematic cost are acceptable as denoted by the vertical and horizontal dotted lines, respectively.

performance criteria throughout learning has important effects on converged performance.

3.5. Learning to reach in other directions

The inertial properties of a linkage having nonconcentric joints that link segments of nonzero mass are nonisotropic. For example, the effective inertia when reaching in a direction that requires a large excursion at the shoulder, as opposed to the elbow, is relatively large because rotating the shoulder requires moving a larger inertial load (mass of upper arm plus forearm, whose center is located far away from the axis of rotation). Furthermore, at every point in the movement, motion about each joint is subject to interaction torques, which depend on arm posture, joint velocities and accelerations (Hollerbach and Flash 1982). The demands placed on torques and underlying muscle activity, therefore, vary substantially with direction (Graham *et al* 2003). To test if the model can cope with these varying demands, it was trained to perform center-out movements to eight evenly spaced directions in space. One of the converged solutions shown previously (figure 5(B); trial 1) was used as the starting point for each of the other directions, which seems more physiological than starting from a random initialization.

For every direction, the simple optimization algorithm successfully tuned the cortical inputs to achieve acceptable performance (figure 6). The muscle excitation patterns that emerged also appear physiological despite using a purely kinematic cost function. Furthermore, it took a smaller number of iterations to converge (avg = 418 iterations <1 cycle, stdv = 251) than the case in which each reach direction was trained from random initializations (avg = 939, stdv = 584).

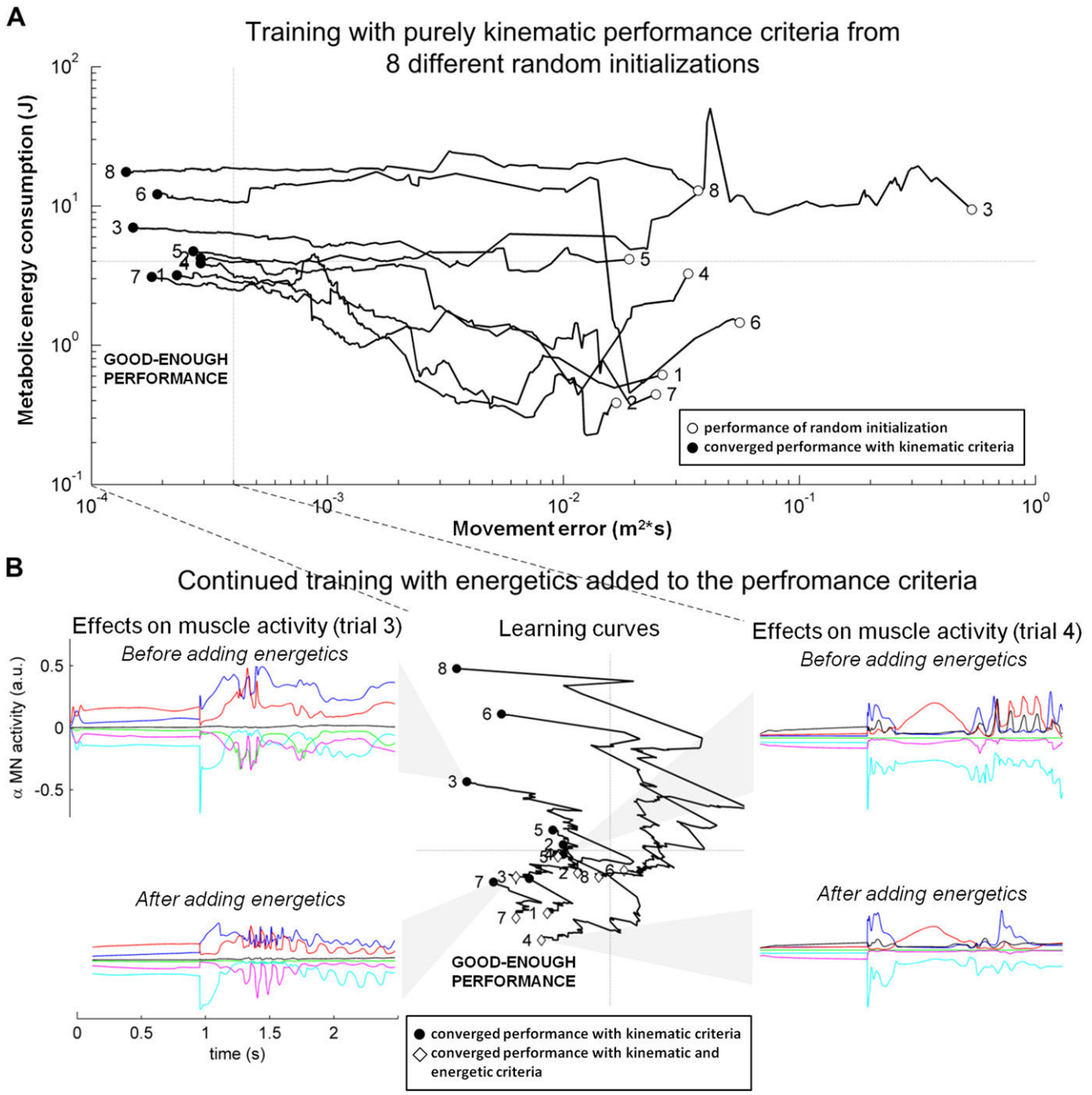

Figure 5. Multiple learning curves. (A) The evolution of learned performance is shown using a purely kinematic cost function from eight different random initializations. Good-enough performance corresponds to the region in performance space where both energetics and kinematics are acceptable as denoted by the dotted lines. (B) The evolution of learned performance is shown after adding an estimate of metabolic energy consumption to the cost function. Alpha motoneuronal traces corresponding to before and after training are shown for two exemplary cases.

3.6. Solution space analysis

Many different sets of inputs to the SLR produced acceptable performance. Random initialization of the model resulted in converged solutions that were located in substantially different parts of the solution space, as evidenced by the different converged values (figure 7(A)) and the highly variable energetic cost of the movements (figure 5(A)), reflecting variable control strategies. In fact, figure 7(B) illustrates that the various solutions for the same reach direction are actually further apart than the solutions for different reach directions that were derived from a common seed.

Although all of these solutions produce similar performance, they may differ in terms of their utility for learning new tasks. Using a solution having a relatively low energetic cost, for example, was a good initialization for learning to reach in all other directions (see 'Learning to reach in other directions'). The location of solutions in the solution space also affects the ability to combine them to generate solutions

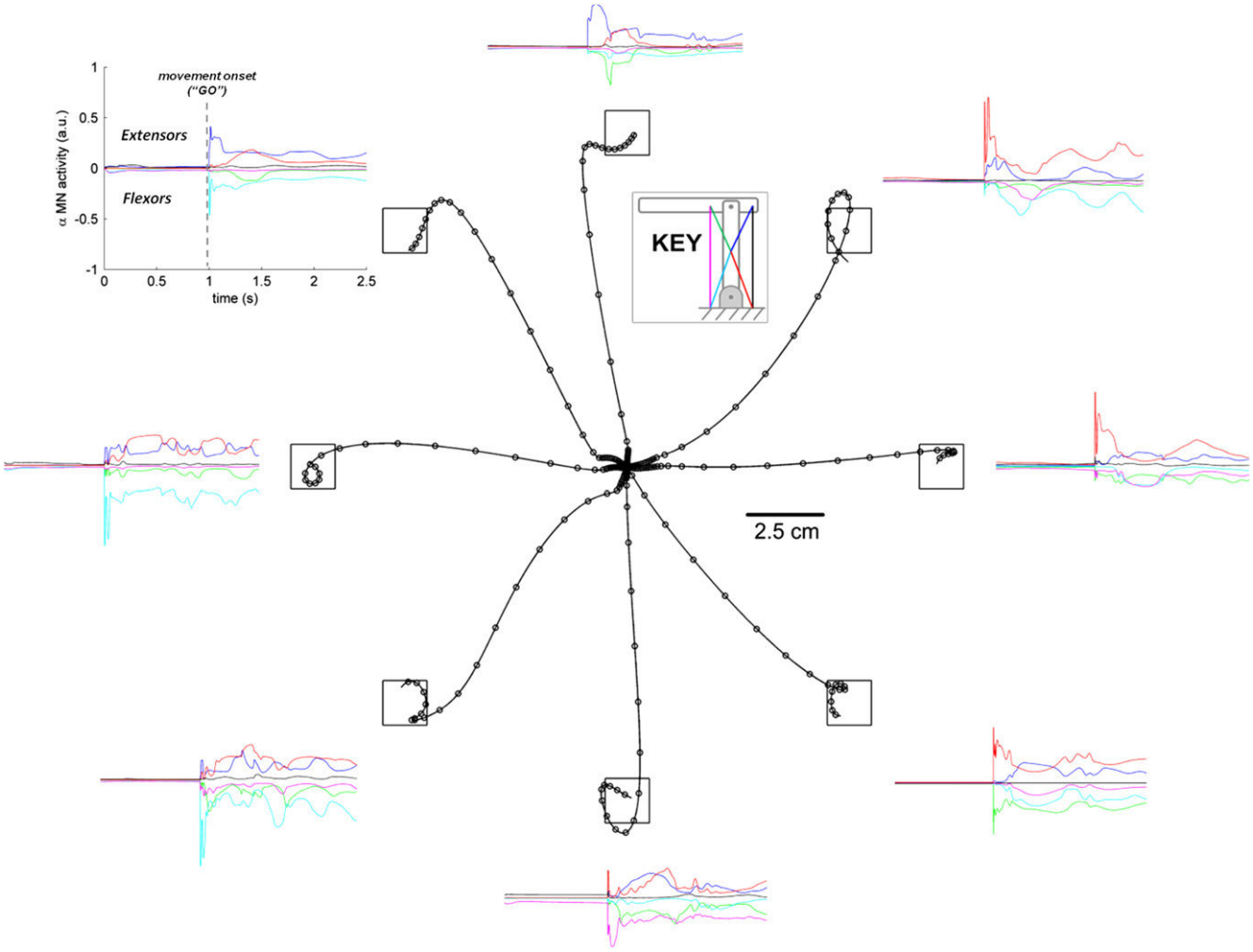

Figure 6. Exemplary kinematics and alpha motoneuron activity for eight reaching targets. Targets were spaced evenly along the perimeter of a 10 cm radius circle centered on the initial position of the hand. Overlaid circles on reach paths are spaced 50 ms apart to illustrate the timing characteristics of the movement. Alpha motoneuronal traces of all six muscles are shown beside each target. Extensor activity is plotted in the upward direction and flexor activity is plotted downward. The key for the alpha motoneuronal traces is shown in figure 3.

to new tasks (see section 3.8 'Interpolability of learned command programs').

3.7. Validation of model performance

Neural control of movement is typically assessed at the muscle level via EMG. Because an electromyogram of a given muscle reflects the aggregate activity of its alpha motoneurons, in principle, it is comparable to the signal of the alpha motoneuron population model.

In order to compare muscle activity emerging in the model with that reported in the experimental literature, it was important to verify that the joint kinematics of these movements were at least qualitatively similar to those observed experimentally. As mentioned in the previous section, kinematics can have large effects on the torques necessary to produce a movement due to their contribution to interaction torques. They also modulate muscle force substantially, which would require proper compensation at the neural level to generate the joint torques required for a given task.

Modeling results were compared against a thorough kinematics and kinetics analysis of center-out reaching that was performed on monkeys (Graham *et al* 2003). Some deviation from the modeling results was expected simply due to the different inertial properties and dimensions of the limbs. Nevertheless, the topology of the monkey arm is similar to our idealized human model and key features of reaching kinematics agree closely with human behavior, so musculoskeletal kinetics should also be similar. Kinematics and kinetics from the model were in qualitative agreement with observations presented in Graham *et al* (2003). A detailed comparison and analysis is presented in Tsianos (2012) (chapter 9; figures 7–8).

Muscle recruitment for a given task can vary substantially across as well as within subjects, which precludes validating the fine details of EMG. Some features, however, such as initial timing and magnitude of EMG bursts are fairly consistent for center-out reaching. Wadman *et al* 1980 reported timing and magnitude of individual muscles that depended systematically on reach direction and were

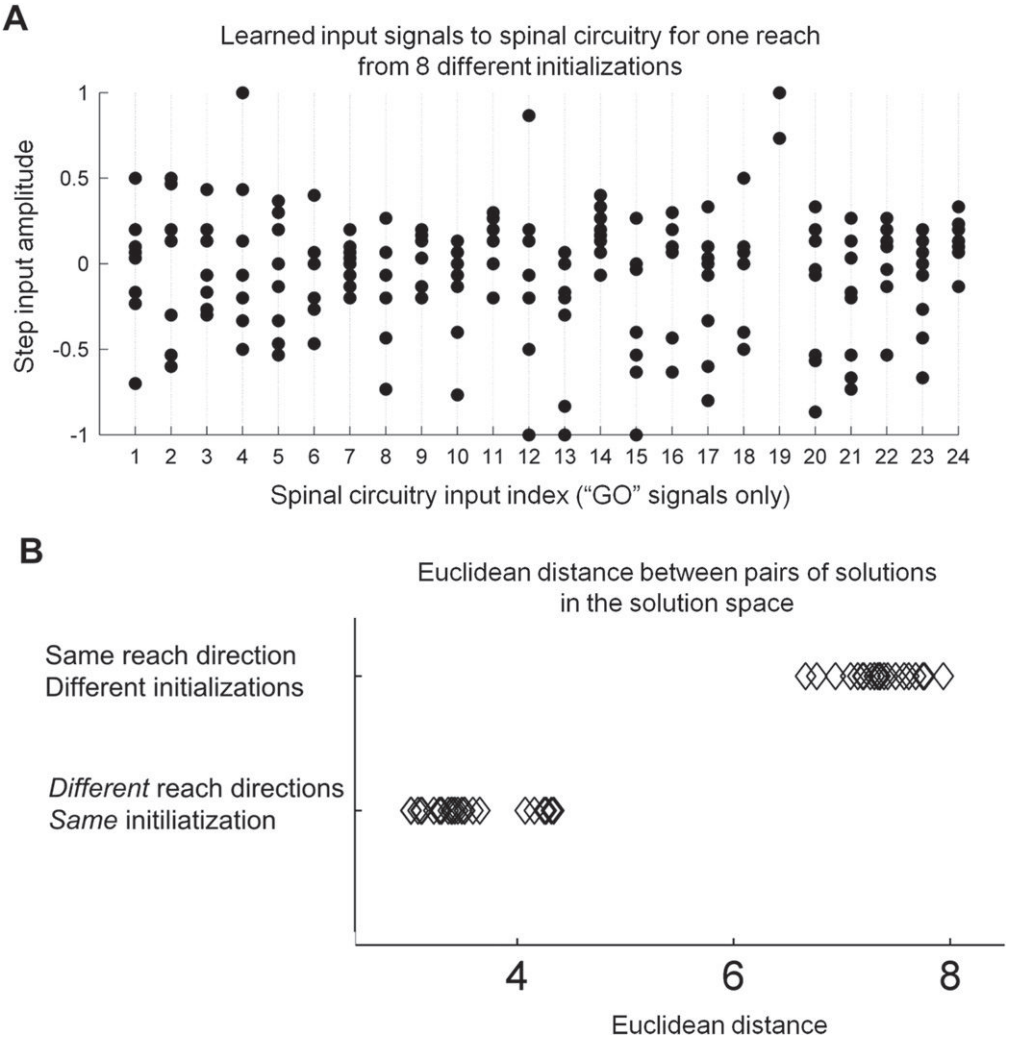

Figure 7. Solution space analysis. (A) Distribution of 'GO' signal amplitudes for one reaching target (outward-left) learned from eight different initializations. Note: there were 24 'GO' signals applied to the model of spinal circuitry, one for each interneuron. (B) Euclidean distance between all possible pairs of solutions for the same reach using different initializations and between all possible pairs of solutions to reaches in different directions using the same initialization.

qualitatively similar to results from the model (Tsianos 2012; chapter 9; figures 10–11). Karst and Hasan (1991) observed systematic coordination of initial elbow and shoulder agonist bursts. In particular, the initial direction of torque (flexor or extensor) produced at each joint correlated with the angle between the long axis of the forearm and the straight path to the target. The relative timing and magnitude between agonist bursts at the shoulder and elbow also correlated with this angle. Such correlations emerged in the model as well and are presented in detail in Tsianos (2012) (chapter 9; figure 9).

3.8. Interpolability of learned command programs

Solutions having different direction, distance, speed, or curl-field magnitude could be used to generate intermediate behaviors via the simple interpolation method but only if the learned solutions were part of the same training sequence and, presumably, located close together in the hyperspace.

3.8.1. Interpolation of direction. Learned solutions for reaches that were 45° apart could be used to interpolate the solution for reaching toward an intermediate direction. Figure 8(A) shows the performance of the best interpolated solution for one particular training sequence (see 'Assessing potential and limits of interpolation'). The performance of this solution was well within the acceptable limit and the appropriate excitation patterns to the muscles emerged from the interpolation. The solution to the intermediate reach was also interpolable for two other training sequences tested (Tsianos 2012; chapter 10; figures 2–3), suggesting that this result is robust. As illustrated by the alpha motoneuron patterns (figure 8(A)) and the pattern of inputs (figure 8(C)) to the model of SC, both the learned solutions and interpolated solutions differed substantially across training sequences. Every interpolated solution that was generated using the technique described in 'Interpolation method' exhibited similar speed and distance and the direction was bounded by the two learned directions (figure 8(B)).

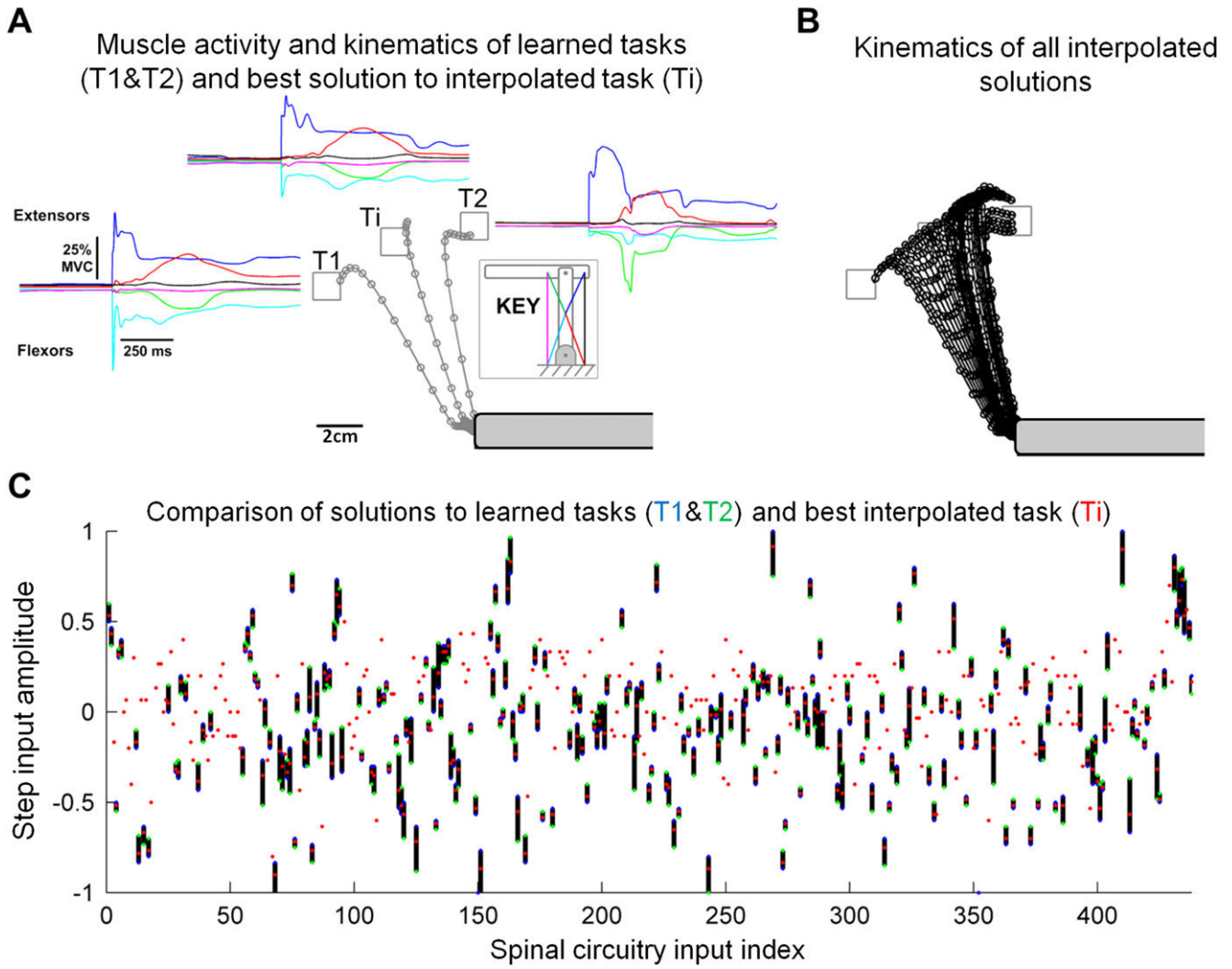

Figure 8. Interpolation results for reach direction. (A) Endpoint kinematics and alpha motoneuron patterns of the solutions to learned reaches (T1 and T2) spaced 45° apart as well as the best interpolated solution to a direction half way in between (Ti). Overlaid circles on endpoint paths are spaced 50 ms apart in time. (B) Reaching trajectories of all 30 interpolated solutions. (C) Values of all input parameters to the SLR for solution to T1 (S1—blue), solution to T2 (S2—green) and solution to Ti (Si—red). Black vertical lines connect the learned values for each input to highlight the set of input values that are changed to accomplish the three tasks; isolated red dots correspond to inputs that were identical for all three solutions. Note that the location of each red dot with respect to green and blue dots is the same for every input and reflects the interpolated solution that produced the best results.

3.8.2. Interpolation of distance. Solutions to reaching movements with a distance that was between two previously learned distances were also interpolable (figure 9). The direction of all reaching paths of the interpolated solutions was roughly the same and their reaching distances were bounded by the distances of the two learned solutions (figure 9(B)). Note that for the longer reaching distance of T2, the tri-phasic burst pattern of muscle activity is necessarily more pronounced with a shorter duration between agonist peaks in order to traverse a longer distance over the same period of time (figure 9(A), top left). By combining S1 and S2, the interpolated solution to reach Ti exhibits a tri-phasic burst pattern with intermediate magnitude and duration.

3.8.3. Interpolation of speed. Solutions for slow and fast reaches could be used to interpolate solutions for intermediate flight durations. All interpolated solutions have the same direction and distance (not shown)and are bounded by the flight durations of movements T1 and T2 (figure 10(B)). The solution to T2 is marked by lower muscle excitation and phasic muscle activity that is more spread out in time. Mixing the right proportions of S1 and S2 led to appropriate muscle excitations for Ti that were essentially intermediate in terms of magnitude and timing.

3.8.4. Interpolation of curl fields. Using a solution to an unperturbed reach (T1) as a starting point, the model was trained successfully to perform reaches in the midst of a strong curl-field applied to the endpoint. Interestingly, the model was able to compensate for the strong and highly

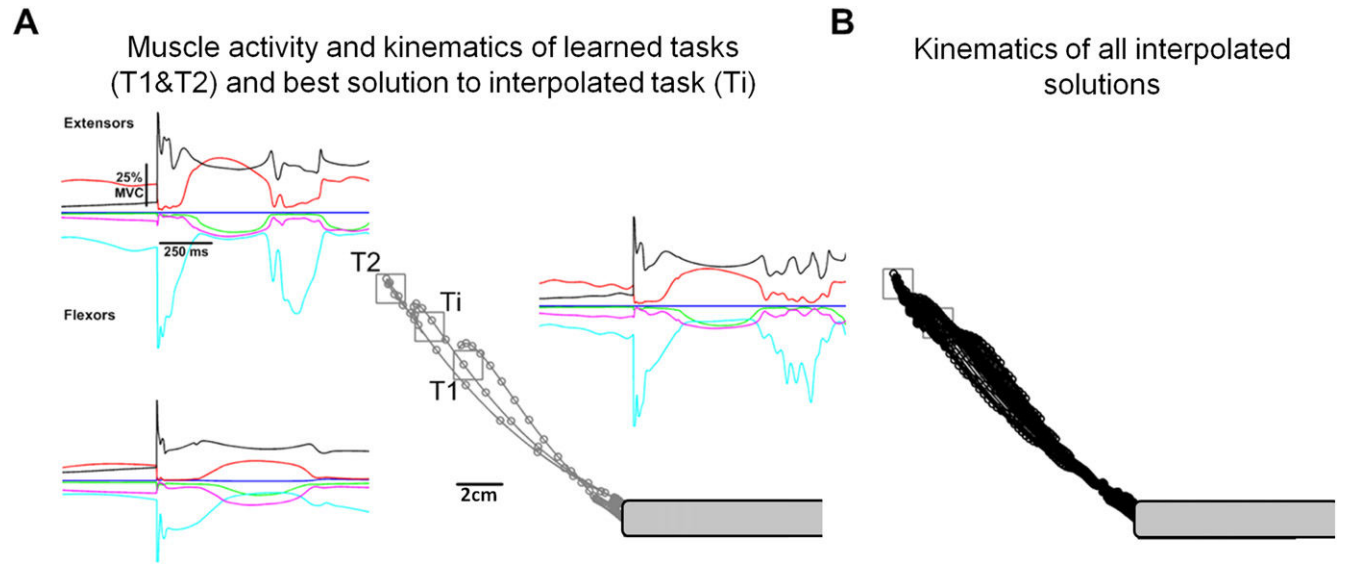

Figure 9. Interpolation results for reaching distance. (A) Endpoint kinematics and alpha motoneuron patterns of the solutions to learned tasks (T1 and T2) from an exemplary training sequence as well as the best interpolated solution to the new task (Ti). Overlaid circles on endpoint paths are spaced 50 ms apart in time. (B) Reaching trajectories of all 30 interpolated solutions.

dynamic perturbation applied to the hand by simply changing the amplitudes of the step functions applied to the SLR; it was not necessary to modulate the descending signals during the movement to achieve the dynamic compensation. Surprisingly, a solution for a free reach (zero curl) and a solution to the same target but with a strong curl-field could be used to interpolate a solution for a curl-field having half the magnitude (figure 11). As expected, when the learned solution S1 drove the model and the clockwise curl-field was turned 'on,' the trajectory deviated in a clockwise direction with respect to the ideal trajectory. The deviation increased with the magnitude of the curl field. When the learned solution to the reach in the presence of a strong field (S2) was tested in a weaker field, the trajectory deviated in the opposite (counterclockwise) direction, and when the field was turned off completely, the deviation increased even more. This result is consistent with the 'after-effects' phenomena observed experimentally, which indicates that the compensation to the curl-field was a result of changes in the dynamics of the muscle activity instead of a simple stiffening of the joints. If the model was simply stiffening the arm by cocontracting the muscles to a high degree, then removing the curl-field would have less of an effect on performance. Interpolating a reach in the midst of a curl-field having a modest magnitude resulted in acceptable performance that also exhibited a sophisticated and metabolically efficient muscle activation strategy that relied more on neurally mediated 'reflexes' rather than cocontraction and 'preflexes' (Brown and Loeb 2000; figure 11(B)).

Note that unlike the after-effects observed experimentally, the endpoint did not eventually stabilize on the target within the allowed duration of the task. This discrepancy probably arose because the model's inputs were constrained to be unmodulated step functions and therefore could not issue corrective action as experimental subjects appear to do early in the movement (Shadmehr and Mussa-Ivaldi 1994).

3.8.5. Limits of interpolation. When interpolating reaches using learned solutions to reaches spaced 90° apart, the resulting performance was not acceptable. As demonstrated by the exemplary result shown in figure 12, the performance of even the best interpolated solution was worse than the acceptable limit, marked by reaching trajectories having inappropriate timing and excursion. Nevertheless, the performance of the interpolated solution to the new task Ti was closer to the target than either of the original solutions. This level of generalization would be useful if the interpolated reach enables a substantial savings as a starting point for subsequent trial-and-error learning. Indeed, training the model using Si as a starting point was substantially faster at reaching good-enough performance (avg = 67.75 iterations, std = 76.65) than initializing the model to S1 (avg = 226, std = 82.4).

To ensure successful interpolation it was necessary to use learned solutions that were obtained from the same training sequence. When learned solutions for reaches spaced 45° apart were obtained from different training sequences, the resulting endpoint kinematics of the interpolated solutions were not bounded by the kinematics of the learned reaches as reported above. They varied wildly both in terms of direction and reaching distance (figure 13(A)). In fact, similar observations were made even when solutions to the same task, from different training sequences, were interpolated (figure 13(B)). The results are consistent with the highly nonlinear nature of the whole plant (SLR + MS). They support the hypothesis that solutions to the same task that were trained from different random initializations are likely to be located near local minima that are distinct and therefore, using them

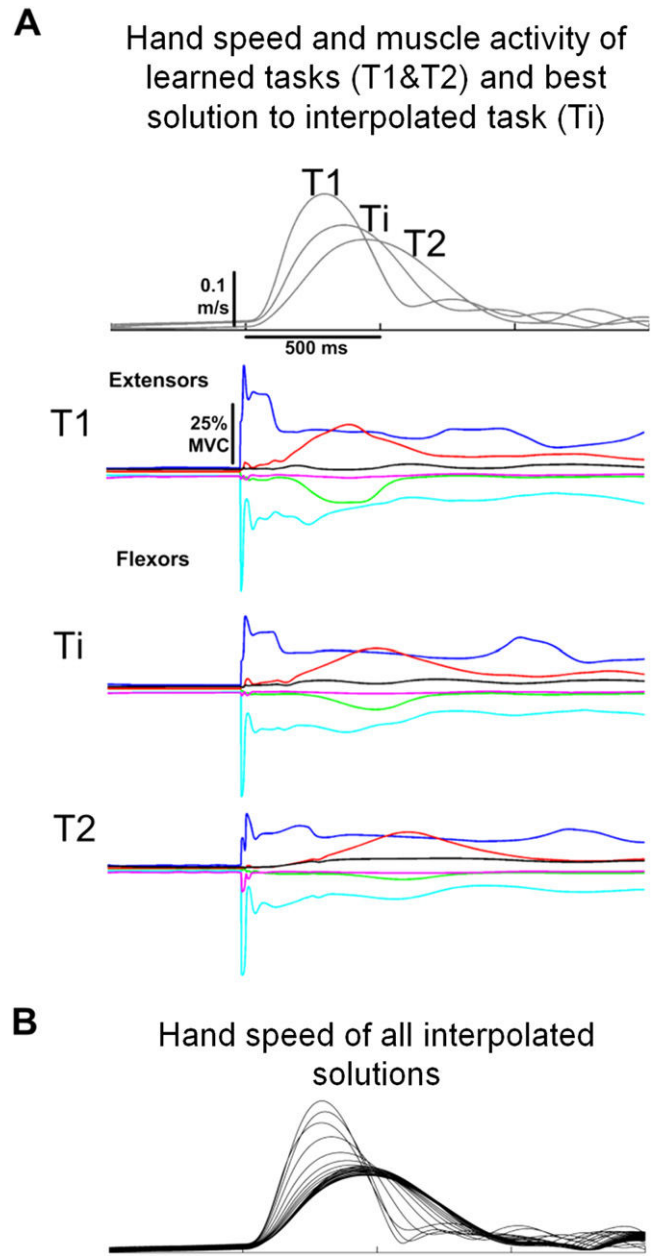

Figure 10. Interpolation results for reaching duration. (A) Tangential speed profiles and alpha motoneuron patterns of the solutions to learned tasks (T1 and T2) from an exemplary training sequence as well as the best interpolated solution to the new task (Ti). (B) Tangential speed profiles of all 30 interpolated solutions.

for interpolation is likely to generate interpolated solutions that lie far from either favorable region in the solution space.

4. Discussion

The models of the musculoskeletal plant, the SLR and the trial-and-error learning employed in this study are all idealized simplifications designed to reveal emergent properties that could influence the interpretation of experimental data from biological systems. Observations that appear to be counter-intuitive tend to provoke complex explanations and conjectures about neural constraints and behavioral strategies, but intuition can be misleading in a system that is highly complex and nonlinear. It is important to consider both the potential and the limits of idealized but realistic models to account for such observations.

4.1. Temporal structure of descending commands

Because the descending commands were modeled as unmodulated step functions, the dynamics of muscle activity were driven by ongoing proprioceptive feedback throughout the movement. This suggests that the brain does not need to specify these dynamics in its commands; instead, it could recruit the appropriate spinal circuits to do this by exploiting the naturally occurring feedback from the periphery. These circuits would have the added potential advantage of automatically generating relatively short-latency responses to perturbations that might arise during the execution of such movements. When complex modulation of cortical activity is observed during movements, it may reflect a monitoring rather than a control function, particularly if those observations come from cortical neurons that are not identified as having corticospinal projections.

The step functions used to excite interneurons in the model may be inadequate for more complex movements that require abrupt changes in muscle recruitment such as reaching out to an object and then interacting with it. Indeed, Yakovenko *et al* (2011) showed that the behavior of cats reaching out to a lever with their forelimb, manipulating it, and then returning to initial position can be broken down into sequences of simple movements that are controlled by sequential changes in both the muscle recruitment and motor cortical activity. Thus, such movements in our model would necessarily require multiphasic commands to the interneurons, but nevertheless less modulated than the muscle excitation patterns required to achieve them.

The relative roles of reflexes mediated by spinal cord and by transcortical loops remains contentious. The shortest latency responses must arise from the monosynaptic projection of muscle spindles directly onto motoneurons but such reflexes appear to be relatively weak during similar reaching behaviors (Kurtzer *et al* 2008). The oligosynaptic pathways through the spinal interneurons are more numerous and could contribute to the larger reflexes observed at longer latencies, but these latencies also permit cortical contributions based strictly on the conduction times to and from motor cortex. Furthermore, there are many other circuits in the brainstem and cerebellar nuclei that could be programmed by motor cortex to contribute to the observed motor output under both perturbed and unperturbed conditions. Some of the functionality ascribed to our SLR could be performed or augmented by similarly programmable gains in a heteronymous set of proprioceptive feedback circuits in those other subcortical structures. That is particularly likely for coordination over longer distances than the projections of typical spinal interneurons; such coordination is required to deal with trunk-limb and interlimb mechanics. The connectivity of such circuits is well-known only for the spinal cord, but the existence

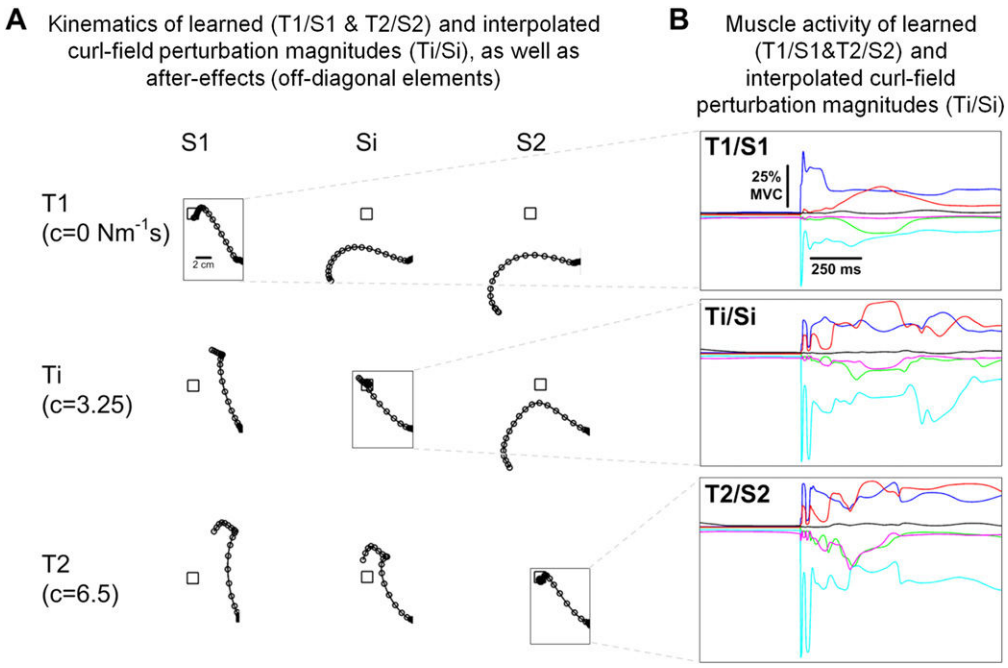

Figure 11. Interpolation results for curl-field magnitude. (A) Endpoint kinematics of the solutions to a learned reach in no curl-field (T1/S1) and strong curl-field (T2/S2) as well as the best interpolated solution for a moderate curl-field (Ti/Si) are shown along the diagonal. Off-diagonal elements show the performance of a particular solution on an untrained magnitude of the curl-field. Ti/S1 and T2/S1 subplots show the system's response to two different curl-field magnitudes prior to adaptation. Ti/S2 and T1/S2 show after-effects upon halving and completely removing the learned strong curl-field, respectively. T1/Si shows after-effects upon removing the learned moderate curl-field and T2/Si shows the response upon doubling the curl-field. (B) Alpha motoneuron patterns for the learned reach in no curl-field (T1/S1) and strong curl-field (T2/S2) as well as the best interpolated solution for reaching in a moderate curl-field (Ti/Si).

of any such circuits anywhere in the neuraxis has similar implications for the control problem that must be solved by the motor cortex.

4.2. Solution space afforded by the SC + MS

The genetically specified and phylogenetically conserved circuitry of the spinal cord appears to be crafting a solution space with a high density of good-enough local minima (see 'Solution space analysis' and Raphael *et al* 2010). In fact, this property holds even when relatively strict performance criteria are imposed such as minimizing both kinematic and energetic costs (figure 5(b)). Such high density of good solutions improves the likelihood and speed that they will be found by any learning algorithm.

Another useful property of the solution space afforded by the SC + MS concerns a natural propensity toward energy-efficient motor strategies even when energy consumption is not included in the cost-function or is added after prolonged training without it. Previous studies of the emergent behavior of a similar model SC + MS for the wrist demonstrated a propensity for strategies that produced relatively little cocontraction (Raphael *et al*, 2010). Initial consideration of kinematic criteria followed by later emphasis on energetics is consistent with experimental observations showing that subjects tend to adopt high levels of cocontraction during the early phases of learning a new task and gradually reduce it once acceptable kinematic performance is attained (see Thoroughman and Shadmehr 1999, for example). A similar effect of ordered changes in the cost function was observed when a similar SLR model to that employed here was trained to resist perturbations (Tsianos *et al* 2011).

4.3. Do similar patterns of muscle recruitment across tasks arise from muscle synergies?

Successfully interpolated command programs in the model exhibited muscle excitations that had distinct similarities (figures 8–11). This could be interpreted as reflecting hard constraints placed on muscle excitations, a concept known as synergies (see Tresch and Jarc 2009 for review). In the model at least, it is simply a result of incremental adjustment of an existing motor repertoire to improve generalization of learning (see final section).

If synergies did underlie center-out reaching, they would have to reproduce the diverse patterns of muscle activity that give rise to various movements. Muscle activity varies substantially across different reaching directions in a plane (Wadman *et al* 1980) and it has been shown that a relatively modest number of muscle synergies can account for most of this activity (d'Avella *et al* 2006). Unloaded reaching in a plane, however, represents only a portion of the very large range of typical reaching movements. Different accuracy and stiffness requirements, for example, result in substantially different muscle recruitment strategies (Osu and Gomi 1999, Cheng and Loeb 2008, Gribble *et al* 2003). Even when a group of synergies can be combined to closely reconstruct activity of a larger number of muscles, the resulting task

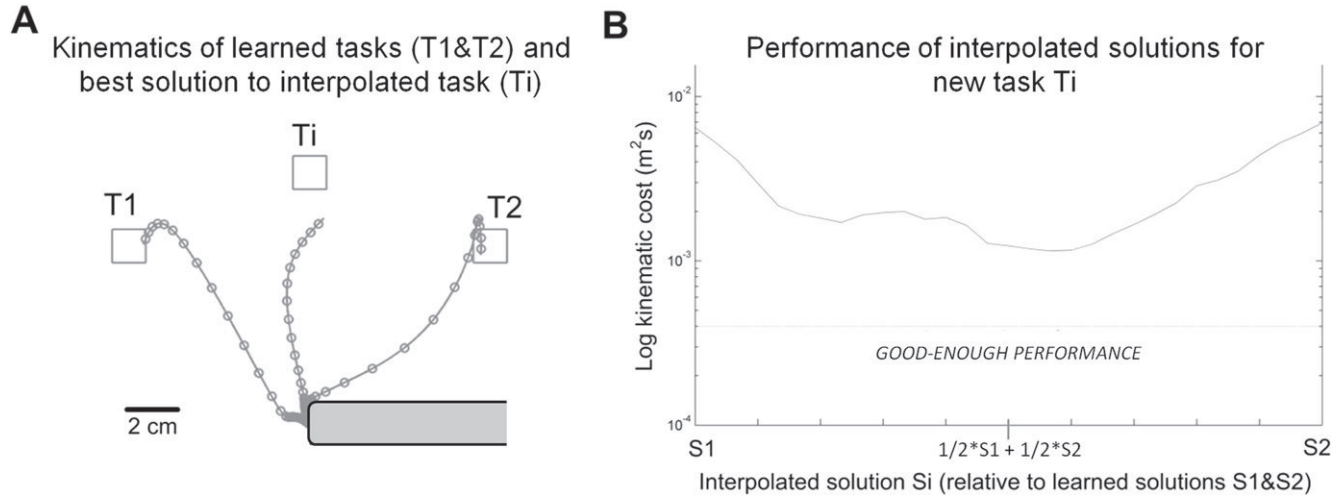

Figure 12. Interpolation results using reaches spaced 90° apart. (A) Endpoint kinematics of the solutions to learned reach directions (T1 and T2) from an exemplary training sequence as well as the best interpolated solution to a direction half way in between (Ti). Overlaid circles on endpoint paths are spaced 50 ms apart in time. (B) Performance of interpolated solutions as a function of different weighted averages of learned solutions S1 and S2.

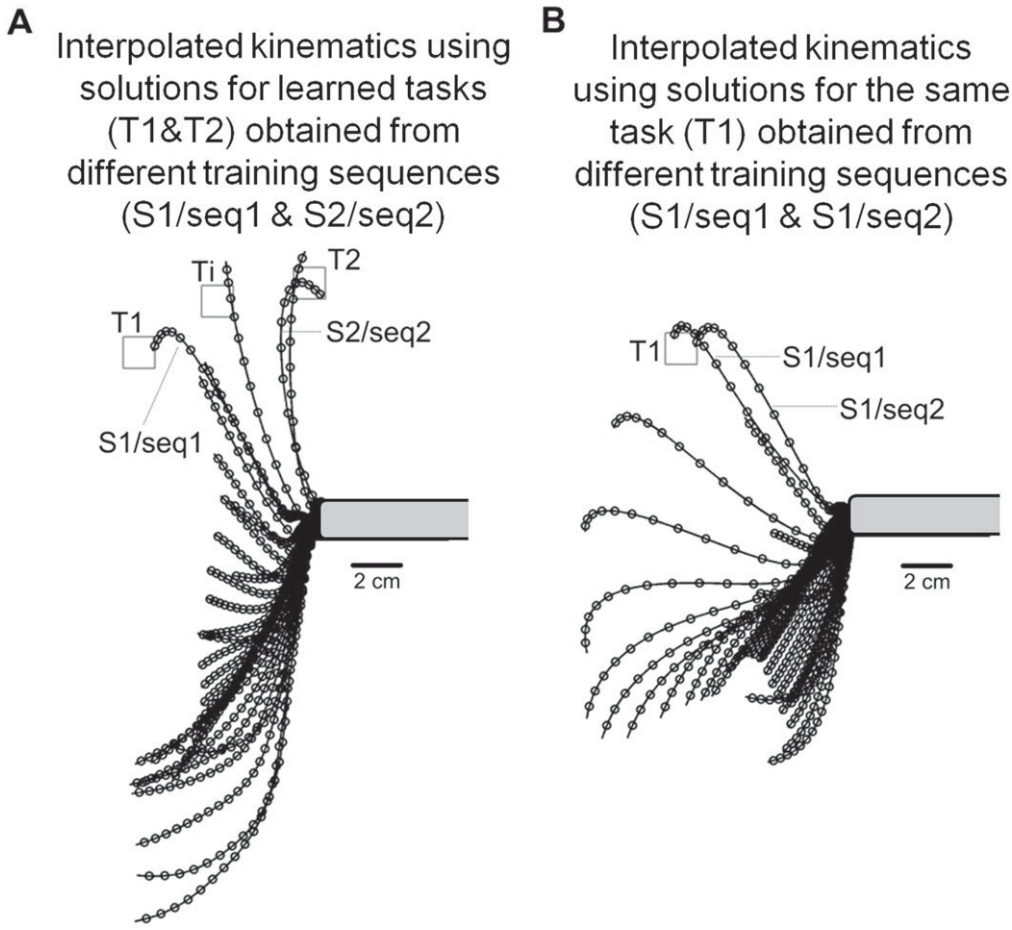

Figure 13. Effects of location of learned solutions in the solution space on performance of interpolated solutions. (A) Endpoint kinematics of the solutions to learned tasks (T1 and T2) as well as all interpolated solutions to the new task (Ti). The solution to T1 was obtained from training sequence 1 and is denoted in the figure as S1/seq1 while the solution to T2 was obtained from training sequence 2 and is denoted as S2/seq2. Overlaid circles on endpoint paths are spaced 50 ms apart in time. (B) Endpoint kinematics of the solutions to a learned task (T1) obtained from training sequence 1 and 2 (denoted as S1/seq1 and S1/seq2, respectively in the figure). The remaining trajectories were obtained by mixing different percentages of the two solutions as described in Interpolation method.

performance may deviate substantially from actual performance (de Rugy *et al* 2013). Bernstein (1967) originally introduced motor synergies to solve the problem of motor redundancy, in which there are more muscles than necessary to control the degrees of freedom of the skeleton. This problem only arises if motor programs must be computed by inverse dynamics, as discussed next. Learning them incrementally, as done here, avoids the problem of redundancy while giving rise to families of solutions that might appear to support the hypothetical solution to the nonexistent problem.

4.4. Are motor programs generated using an internal model of the plant?

If the nervous system had an internal inverse model of the plant, then it would be able to compute the optimal motor program for a given task online. Such a model would have to account for the properties of the SC + MS plus the convergence patterns of all of the brain's outputs onto the SLR, which seems implausible. Furthermore, experimental subjects often adopt motor strategies that are far from optimal (de Rugy *et al* 2012) even when they have been exposed previously to the optimal strategy (Ganesh *et al* 2010, Latash *et al* 1998). If the brain were capable of learning an internal model of the plant, then forming an internal model of a velocity dependent force field applied to the hand should be easy given its relative simplicity. Because an internal model by definition captures the intrinsic properties of the plant, its validity should generalize across all possible behaviors of the plant. It has been shown consistently, however, that learning to resist a novel force field for a particular reach does not generalize well across reaches having different distances (Mattar and Ostry 2010) and directions (Gandolfo *et al* 1996, Mattar and Ostry 2007). Substantial generalization was only observed across different postures and only for a specific type of force field that produced the same velocity-dependent torque perturbations in both postures (Shadmehr and Mussa-Ivaldi 1994). This observation has led to the popular hypothesis that the sensorimotor system develops an internal model of the force field as a mapping between joint velocities and torque perturbations, i.e. in intrinsic coordinates. Better generalization is expected, however, simply because torque perturbations across postures are similar; thus a given change in muscle recruitment learned in one posture is more likely to be effective in a different posture (see Thoroughman and Shadmehr 1999, Shadmehr and Moussavi 2000).

It has been reported that adaptation to a force field in the workspace for one movement generalizes across other movements, like training on point-to-point reaches and then moving in a circular path (Conditt *et al* 1997). The circular motion, however, could have been a product of the learned discrete reaches stitched together in time. Goodbody and Wolpert (1998) reported that adaptation to a curl-field for one reaching movement generalizes across untrained movements having either shorter duration or amplitude, but Mattar and Ostry (2010) claimed that this was an experimental artifact.

4.5. Interpolability constrains motor learning

The ranges of task parameters over which interpolability was observed, while not unlimited, were substantial and surprisingly consistent with experimental results. This is a necessary condition for it to be feasible to use a learned and stored repertoire of programs to generate the virtually unlimited set of motor behaviors of which humans are capable.

Local generalization has been observed experimentally in subjects that were trained to counteract a force field in one portion of the workspace and subsequently tested for movements that differed in direction (Gandolfo *et al* 1996, Mattar and Ostry 2007). In fact, Mattar and Ostry (2007) showed that generalization was very weak for reach directions that differed by more than 45° from a single trained direction and was nearly nonexistent when the two reaches differed by 90°. When the test reach was between a pair of trained directions that were 90° apart, generalization was good but not complete, which is consistent with our modeling results.

If the interpolability of solutions is indeed a property that is exploited by the brain to speed up learning and save storage capacity, then this would reinforce the exploration of motor programs that were similar to those stored from previous experience even if other strategies might satisfy the current task better (de Rugy *et al* 2012, Ganesh *et al* 2010, Debicki and Gribble 2005). This sort of learning strategy has interesting implications for the tendency to develop and persist in motor habits even when they interfere with optimizing performance for the immediate criteria of the task. It appears that such incremental learning of motor repertoires is useful for fulfilling a larger goal of the nervous system that is more than simply completing the task at hand in an optimal way, but considers also the rate at which new tasks are learned and the memory capacity that they will require (Loeb 2012).

Acknowledgements

We thank N Li for assistance with implementing the model in software. This work was funded by the DARPA REPAIR program.

References

Aruin A S and Latash M L 1995 Directional specificity of postural muscles in feed-forward postural reactions during fast voluntary arm movements *Exp. Brain Res.* **103** 323–32

Baudry S, Maerz A H and Enoka R M 2010 Presynaptic modulation of Ia afferents in young and old adults when performing force and position control *J. Neurophysiol.* **103** 623–31

Bernstein N A 1967 *Human Motor Actions: Bernstein Reassessed* Translation ed H T A Whiting (Amsterdam: Elsevier)

Brown I E and Loeb G E 2000 *Neuro-Control of Posture and Movement* ed J Winters and P Crago (New York: Springer) pp 148–63

Carroll T J, Baldwin E R L and Collins D F 2005 Task dependent gain regulation of spinal circuits projecting to the human flexor carpi radialis *Exp. Brain Res.* **161** 299–306

Chen X Y, Carp J S, Chen L and Wolpaw J R 2002 Corticospinal tract transection prevents operantly conditioned H-reflex increase in rats *Exp. Brain Res.* **144** 88–94
Cheng E J and Loeb G E 2008 On the use of musculoskeletal models to interpret motor control strategies from performance data *J. Neural Eng.* **5** 232–53
Conditt M A, Gandolfo F and Mussa-Ivaldi F 1997 The motor system does not learn the dynamics of the arm by rote memorization of past experience *J. Neurophysiol.* **78** 554–60
d'Avella A, Portone A, Fernandez L and Lacquaniti F 2006 Control of fast-reaching movements by muscle synergy combinations *J. Neurosci.* **26** 7791–810
de Rugy A, Loeb G and Carroll T 2013 Are muscle synergies useful for neural control? *Front. Comput. Neurosci.* **7** 19
de Rugy A, Loeb G E and Carroll T J 2012 Muscle coordination is habitual rather than optimal *J. Neurosci.* **32** 7384–91
Debicki D B and Gribble P L 2005 Persistence of inter-joint coupling during single-joint elbow flexions after shoulder fixation *Exp. Brain Res.* **163** 252–7
Eliasmith C and Anderson A H 2003 *Neural Engineering: Computation, Representation, And Dynamics in Neurobiological Systems* (Cambridge, MA: MIT)
Fleshman J W, Segev I and Burke R 1988 Electronic architecture of type-identified motoneurons in the cat spinal cord *J. Neuropathol. Exp. Neurol.* **60** 60–85
Gandolfo F, Mussa-Ivaldi F A and Bizzi E 1996 Motor learning by field approximation *Proc. Natl. Acad. Sci. USA* **93** 3843–6
Ganesh G, Haruno M, Kawato M and Burdet E 2010 Motor memory and local minimization of error and effort, not global optimization, determine motor behavior *J. Neurophysiol.* **104** 382–90
Goodbody S J and Wolpert D M 1998 Temporal and amplitude generalization in motor learning *J. Neurophysiol.* **79** 1825–38
Graham K M, Moore K D, Cabel D W, Gribble P L, Cisek P and Scott S H 2003 Kinematics and kinetics of multijoint reaching in nonhuman primates *J. Neurophysiol.* **89** 2667–77
Gribble P L, Mullin L I, Cothros N and Mattar A 2003 Role of cocontraction in arm movement accuracy *J. Neurophysiol.* **89** 2396–405
Hallett M, Shahani B T and Young R R 1975 EMG analysis of stereotyped voluntary movements in man *J. Neurol. Neurosurg. Psychiatry* **38** 1154–62
He J, Levine W S and Loeb G E 1991 Feedback gains for correcting small perturbations to standing posture *IEEE Trans. Autom. Control* **36** 322–32
Henneman E, Somjen G and Carpenter D O 1965 Functional significance of cell size in spinal motoneurons *J. Neurophysiol.* **28** 560–80
Hollerbach J M and Flash T 1982 Dynamic interactions between limb segments during planar arm movement *Biol. Cybern.* **44** 67–77
Jadi M, Polsky A, Schiller J and Mel B W 2012 Location-dependent effects of inhibition on local spiking in pyramidal neuron dendrites *PLoS Comput. Biol.* **8** e1002550
Kargo W J and Nitz D A 2003 Early skill learning is expressed through selection and tuning of cortically represented muscle synergies *J. Neurosci.* **23** 11255–69
Karst G M and Hasan Z 1991 Timing and magnitude of electromyographic activity for two-joint arm movements in different directions *J. Neurophysiol.* **66** 1594–604
Kurtzer I L, Pruszynski J A and Scott S H 2008 Long-latency reflexes of the human arm reflect an internal model of limb dynamics *Curr. Biol.* **18** 449–53
Lackner J R and DiZio P 1994 Rapid adaptation to coriolis foce perturbations of arm trajectory *J. Neurophysiol.* **72** 299–313
Latash M L, Gelfand I M, Li Z M and Zatsiorsky V M 1998 Changes in the force-sharing pattern induced by modifications of visual feedback during force production by a set of fingers *Exp. Brain Res.* **123** 255–62
Lee W A 1980 Anticipatory control of postural and task muscles during rapid arm flexion *J. Mot. Behav.* **12** 185–96
Lillicrap T P and Scott S H 2013 Preference distributions of primary motor cortex neurons reflect control solutions optimized for limb biomechanics *Neuron* **77** 168–79
Loeb G E 1983 Finding common ground between robotics and physiology *Trends Neurosci.* **6** 203–4
Loeb G E 2012 Optimal isn't good enough *Biol. Cybern.* **106** 757–65
Mattar A A G and Ostry D J 2007 Modifiability of generalization in dynamics learning *J. Neurophysiol.* **98** 3321–9
Mattar A A G and Ostry D J 2010 Generalization of dynamics learning across changes in movement amplitude *J. Neurophysiol.* **104** 426–38
McCrea D A and Rybak I A 2008 Organization of mammalian locomotor rhythm and pattern generation *Brain Res. Rev.* **57** 134–46
McKeon B, Gandevia S and Burke D 1984 Absence of somatotopic projection of muscle afferents onto motoneurons of same muscle *J. Neurophysiol.* **51** 185–94
Mileusnic M P, Brown I E, Lan N and Loeb G E 2006 Mathematical models of proprioceptors. I. control and transduction in the muscle spindle *J. Neurophysiol.* **96** 1772–88
Mileusnic M P and Loeb G E 2009 Force estimation from ensembles of Golgi tendon organs *J. Neural Eng.* **6** 036001
Osu R and Gomi H 1999 Multijoint muscle regulation mechanisms examined by measured human arm stiffness and EMG signals *J. Neurophysiol.* **81** 1458–68
Pierrot-Deseilligny E and Burke D C 2005 *The Circuitry of the Human Spinal Cord: Its Role in Motor Control and Movement Disorders* (Cambridge: Cambridge University Press)
Polsky A, Mel B W and Schiller J 2004 Computational subunits in thin dendrites of pyramidal cells *Nat. Neurosci.* **7** 621–7
Raphael G, Tsianos G A and Loeb G E 2010 Spinal-like regulator facilitates control of atwo-degree-of-freedom wrist *J. Neurosci.* **30** 9431–44
Rathelot J A and Strick P L 2009 Subdivisions of primary motor cortex based on cortico-motoneuronal cells *PNAS* **106** 918–23
Rudomin P and Schmidt R F 1999 Presynaptic inhibition in the vertebrate spinal cord revisited *Exp. Brain Res.* **129** 1–37
Scheidt R A, Reinkensmeyer D J, Conditt M A, Rymer W Z and Mussa-Ivaldi F A 2000 Persistence of motor adaptation during constrained, multi-joint, arm movements *J. Neurophysiol.* **84** 853–62
Scheidt R A and Rymer W Z 2000 Control strategies for the transition from multijoint to single-joint arm movements studied using a simple mechanical constraint *J. Neurophysiol.* **83** 1–12
Segev I, Fleshman J W and Burke R E 1990 Computer simulation of group Ia EPSPs using morphologically realistic methods of cat x-motoneurons *J. Neurophysiol.* **64** 648–60
Shadmehr R and Moussavi M K 2000 Spatial generalization from learning dynamics of reaching movements *J. Neurosci.* **20** 7807–15
Shadmehr R and Mussa-Ivaldi F A 1994 Adaptive representation of dynamics during learning of a motor task *J. Neurosci.* **14** 3208–24
Stein P S G 1978 Motor systems, with specific reference to the control of locomotion *Annu. Rev. Neurosci.* **1** 61–81
Thoroughman K A and Shadmehr R 1999 Electromyographic correlates of learning an internal model of reaching movements *J. Neurosci.* **19** 8573–88
Tresch M C and Jarc A 2009 The case for and against muscle synergies *Curr. Opin. Neurobiol.* **19** 601–7
Tsianos G A 2012 Investigating the role of muscle physiology and spinal circuitry in sensorimotor control *PhD Dissertation* University of Southern California, Los Angeles, USA

Tsianos G A, Raphael G and Loeb G E 2011 Modeling the potentiality of spinal-like circuitry for stabilization of a planar arm system *Prog. Brain Res.* **194** 203–13

Tsianos G A, Rustin C and Loeb G E 2012 Mammalian muscle model for predicting force and energetics during physiological behaviors *IEEE Trans. Neural. Syst. Rehabil. Eng.* **20** 117–33

Tyler A E and Karst G M 2004 Timing of muscle activity during reaching while standing: systematic changes with target distance *Gait Posture* **20** 126–33

Wadman W J, Denier van der Gon J J and Derksen R J A 1980 Muscle activation patterns for fast goal-directed arm movements *J. Hum. Mov. Stud.* **6** 19–37

Yakovenko S, Krouchev N and Drew T 2011 Sequential activation of motor cortical neurons contributes to intralimb coordination during reaching in the cat by modulating muscle synergies *J. Neurophysiol.* **105** 388–409

Zajac F E 1993 Muscle coordination of movement: a perspective *J. Biomech.* **26** 109–24