A Computational Model of Cerebellar Learning for Limb Control *

Andrew H. Fagg[1] Nathan Sitkoff[1] Andrew G. Barto[1]
James C. Houk [2]

{fagg | sitkoff | barto}@cs.umass.edu
[1]Department of Computer Science
University of Massachusetts
Amherst, MA 01003-4610

houk@casbah.acns.nwu.edu
[2]Department of Physiology
Northwestern University School of Medicine
Chicago, IL 60611-3008

Abstract

We present a model of cerebellar learning for control of limb movements. The model learns to generate relatively smooth and direct arm movements, while not assuming (as some models do) a high-level input that specifies a pre-planned trajectory. The model controls a simulated two-joint arm, which is actuated by a set of six muscles. Control of the arm is shared by a learning cerebellar module and a hard-wired extra-cerebellar (**EC**) system. The cerebellar module consists of a group of independent submodules, referred to as *Adjustable Pattern Generators* (**APGs**). Each APG is made up of a set of Purkinje cells and a single cerebellar nuclear cell. Activation of an APG's nuclear cell drives a *muscle synergy* (a subset of the muscles). The EC system is responsible for producing low-quality corrective movements in situations where the cerebellar module is unable to bring the arm to the specified target. Via proprioceptive inputs, the modeled inferior olive assesses directional errors in cerebellar-generated movements by observing subsequent corrective movements produced by the EC system.

*This work is supported by the National Institutes of Health (grant #NIH 1-50 MH 48185-02)

1 Introduction

Many have argued for the cerebellum's role in the learning control of smooth, coordinated limb movements [6, 7, 10, 21]. However, there is less agreement as to the informational content of both the output of the cerebellum, and the training information received from the inferior olive.

Here, we present a model of cerebellar learning for limb control which represents one perspective on these questions. Specifically:

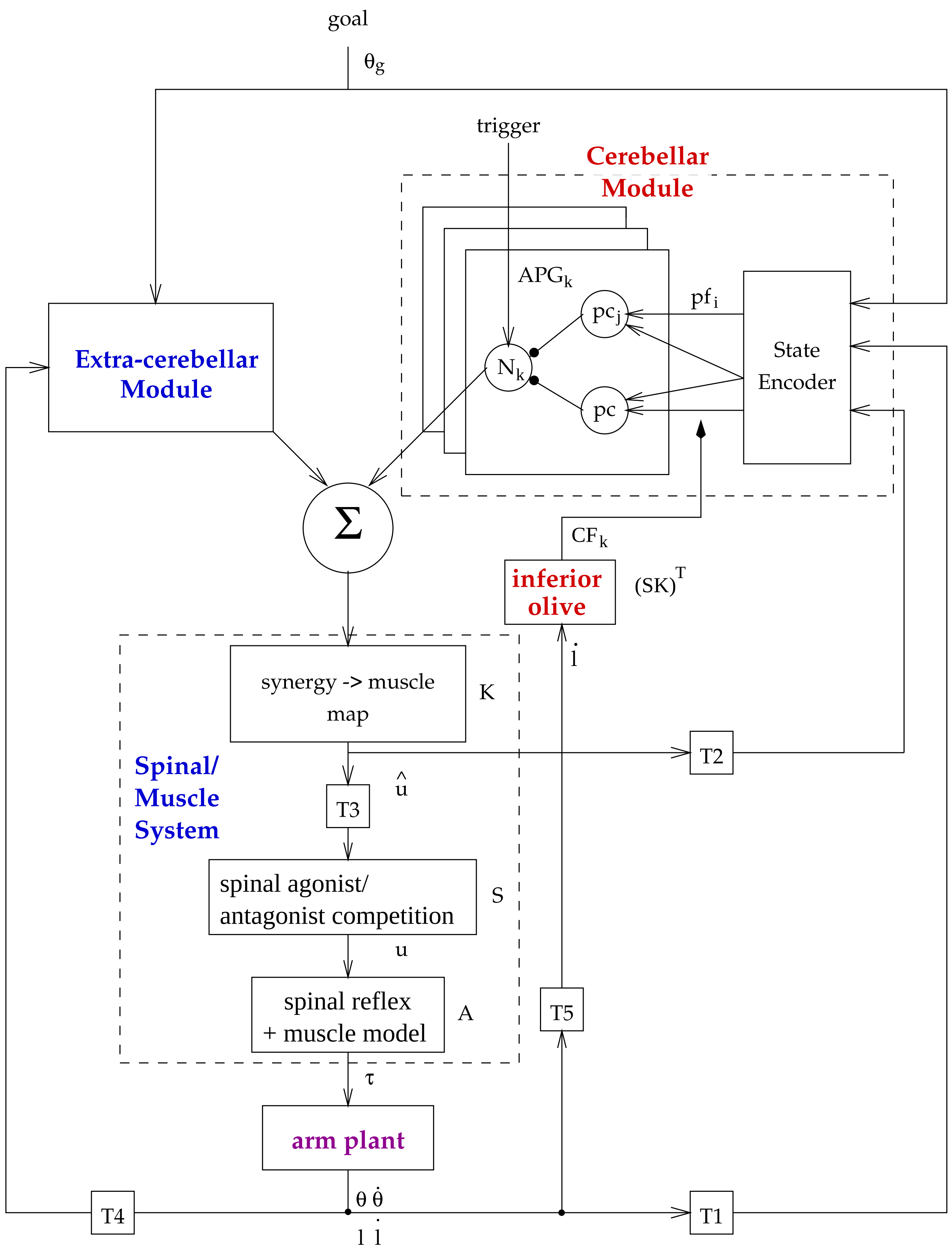
1. The cerebellum has access to the musculature via multiple routes, including the red nucleus and primary motor cortex [8], whose cells typically project to multiple muscles [4, 16]. Hence, it is important to think of the currency of cerebellar output in terms of a ***muscle synergy space***, rather than directly in terms of joint torques or the activation of individual muscles.

2. Although human reaching movements demonstrate very regular behavior (e.g. bell-shaped velocity profiles), we believe it to be unreasonable that a detailed trajectory (consisting of position, velocity, and acceleration of the joints) is planned prior to the execution of the movement. Rather, we believe that such a trajectory naturally emerges from a combination of a gross plan of action, a learned motor program, and the dynamics of the arm.

3. In learning simple reaching movements, a high-quality training signal (from the inferior olive) may not be necessary. Instead, our model indirectly derives cerebellar training information by observing crude corrective movements that follow cerebellar-generated movements.

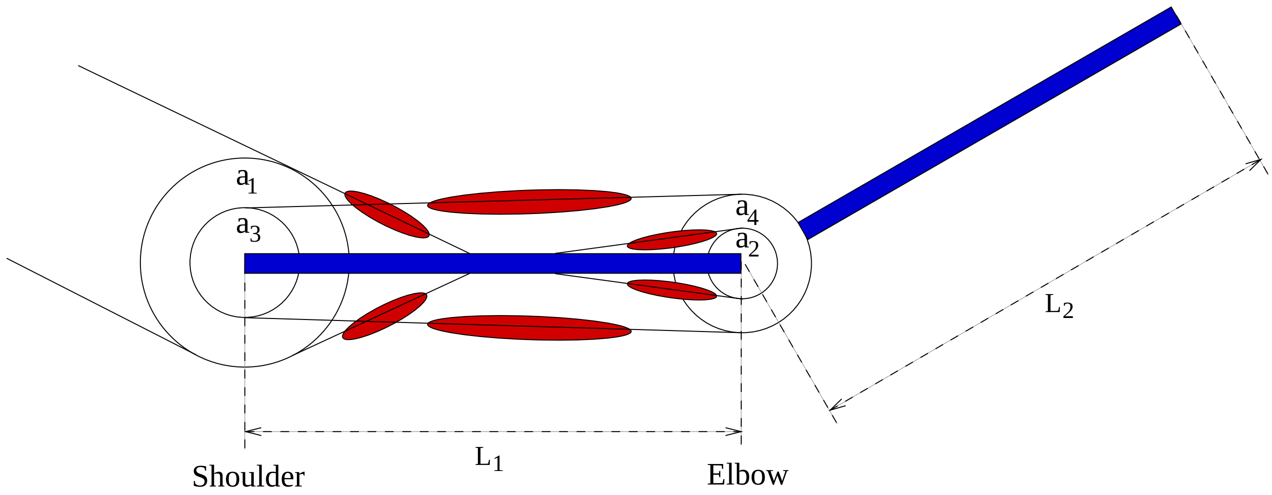
The model presented below learns to perform point-to-point reaching movements in a trial-and-error fashion. The modeled cerebellum acquires an implicit inverse model of the limb, which combines delayed sensory inputs, a motor efference copy, and contextual information in order to compute the current motor output.

2 Model Architecture

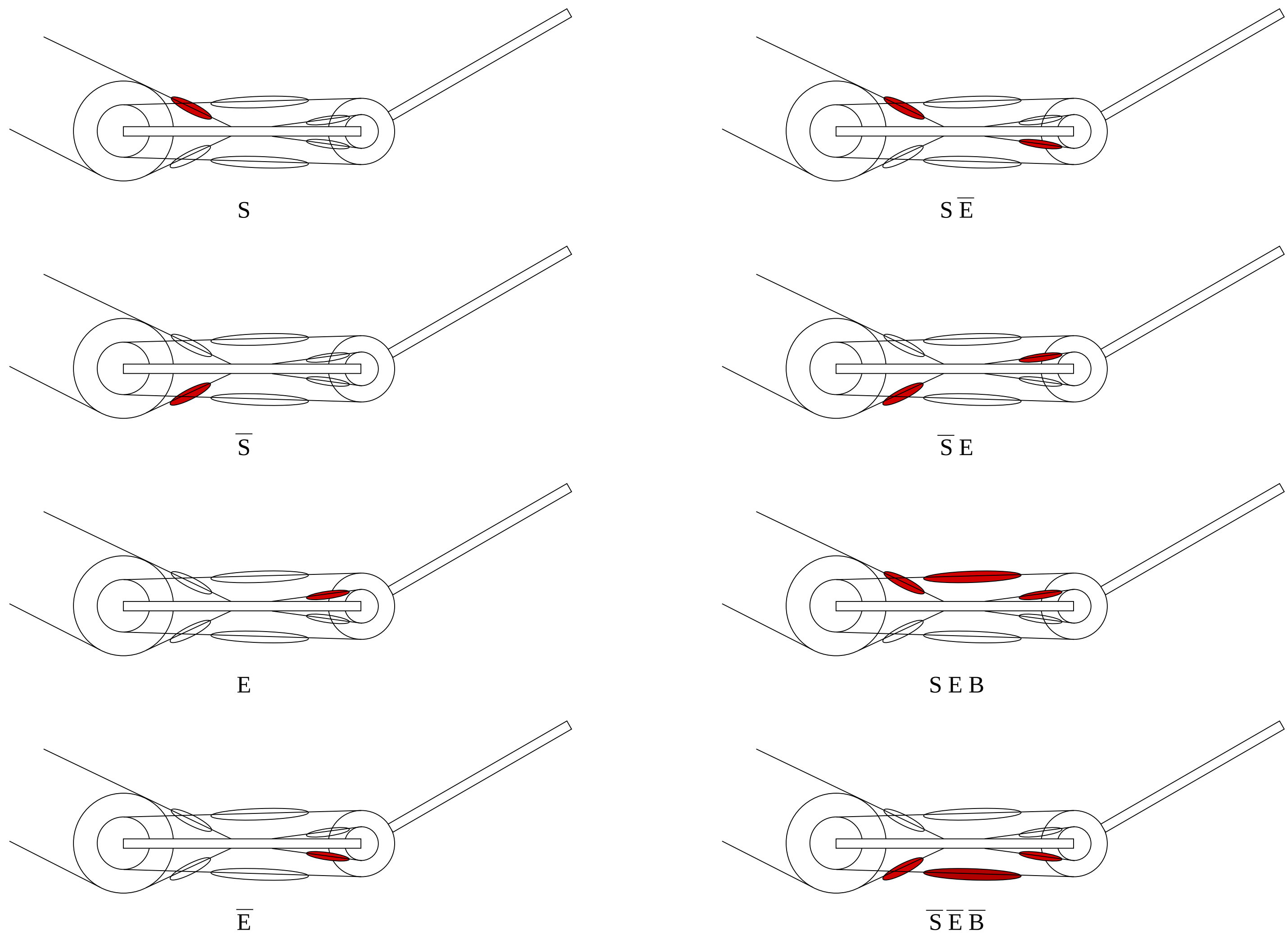


An overview of the system architecture is shown in the above figure. Two control modules (*cerebellar* and *extra-cerebellar* (EC) systems) produce motor signals in a muscle synergy space. These signals are transformed by the *spinal/muscle* system into joint torque commands that drive the limb. Errors in movement made by the cerebellar module are detected by the *inferior olive* (IO). This is accomplished indirectly by observing the proprioceptive signals that result from subsequent corrective movements that are generated by the the EC controller.

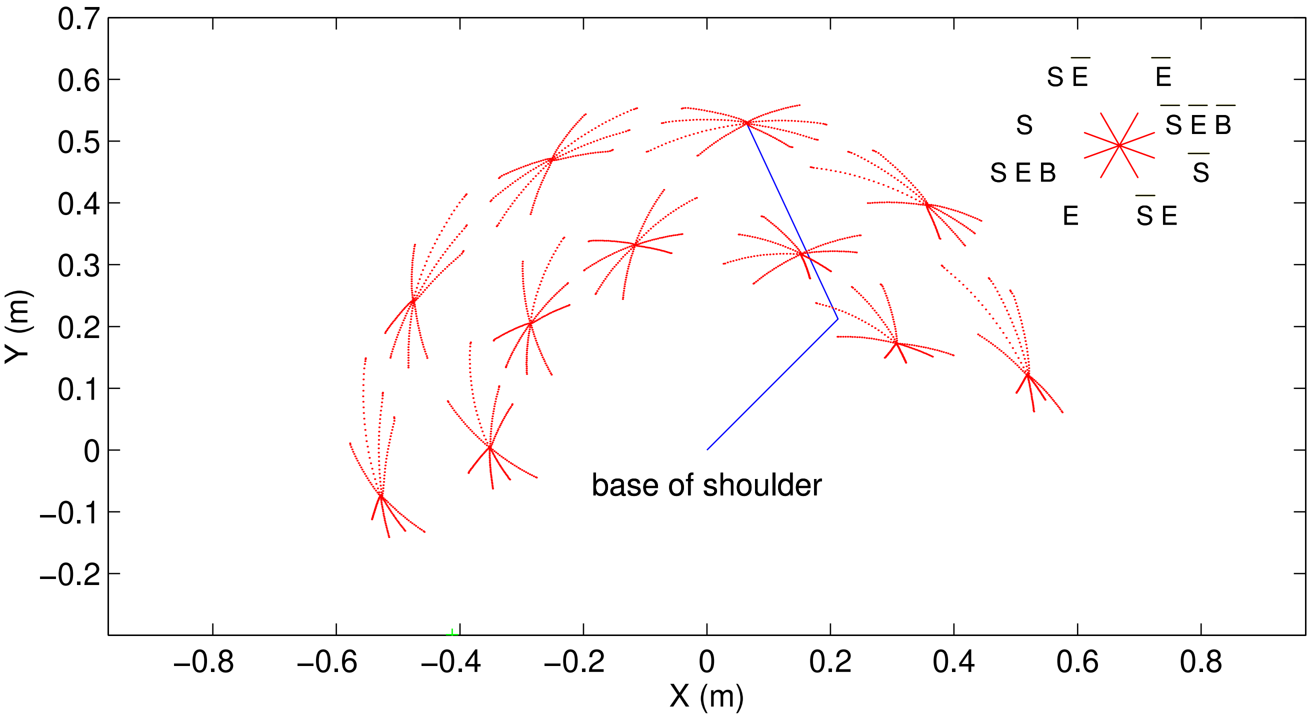
2.1 Arm Model and Muscle Synergies



The plant used in our experiments is a simulated, two degree-of-freedom, planar arm (adapted from [12]). Actuation of the arm is performed by three pairs of opposing muscles. Two pairs of muscles individually actuate the shoulder and elbow joints, while the third pair actuates both joints simultaneously. Each pair of muscles pull against one-another. Activation of the *agonist* muscle results in a flexion of the joint(s); activation of the *antagonist* results in joint extension. A single muscle behaves as a non-linear visco-elastic (spring) element. A particular activation level of a muscle translates linearly into a setting of the muscle's rest length.



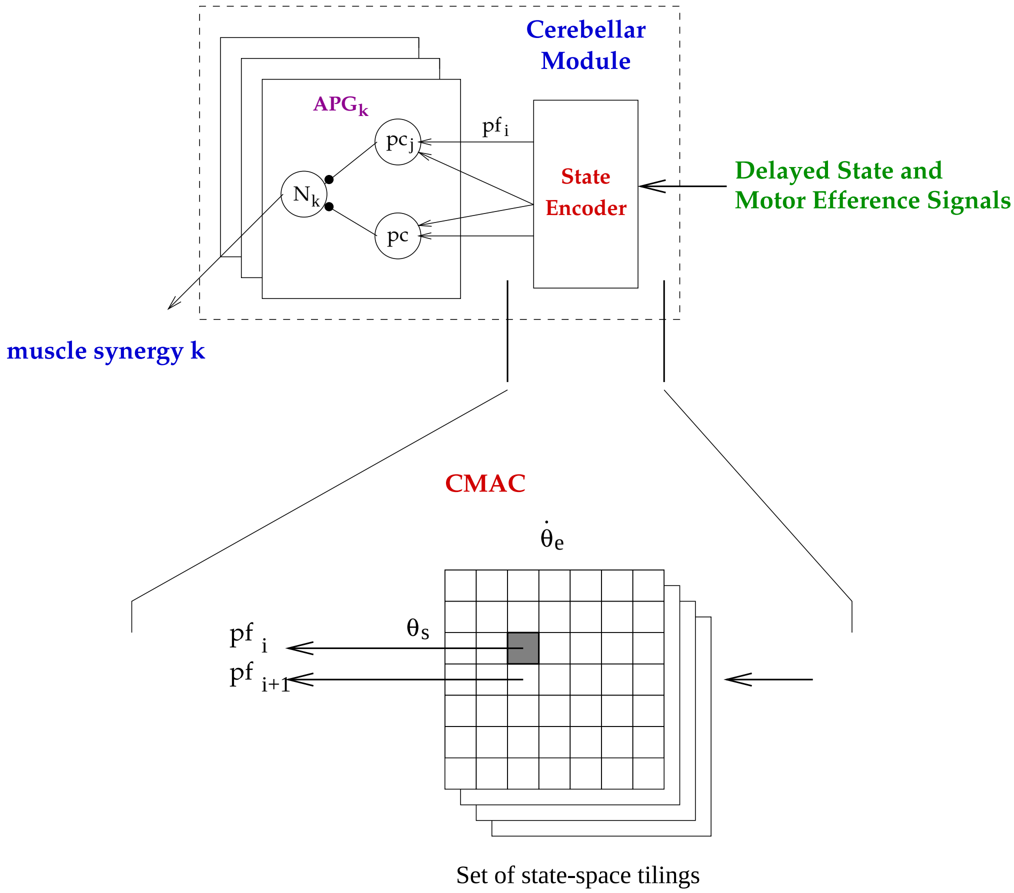
Control of the arm is achieved through the recruitment of mixtures of the eight muscle synergies (above). Muscle synergy S recruits only the shoulder agonist; $\bar{S}$ recruits the antagonist; and $S\bar{E}$ recruits a combination of the shoulder agonist and the elbow antagonist.



The behavior of each muscle synergy as a function of position in peripersonal space is demonstrated in the above figure. Each "spider" corresponds to a single starting point. Individual legs of a spider are the result of tracking the position of the wrist in response to activation of one muscle synergy for 50ms. The muscle synergy labels (upper right of the figure) correspond spatially to the upper, center spider.

Note that movement in arbitrary directions can be obtained by appropriately mixing several muscle synergies. In addition, as the shoulder rotates through its range, the spiders also rotate in a corresponding fashion.

2.2 The Cerebellar Module



The learning cerebellar control module consists of an array of *Adjustable Pattern Generators* (APGs), and a *state encoder*. Each APG is constructed from a group of Purkinje Cells (PCs) and a cerebellar nuclear cell. An APG is responsible for learning the appropriate state-dependent activation level of the muscle synergy that it drives.

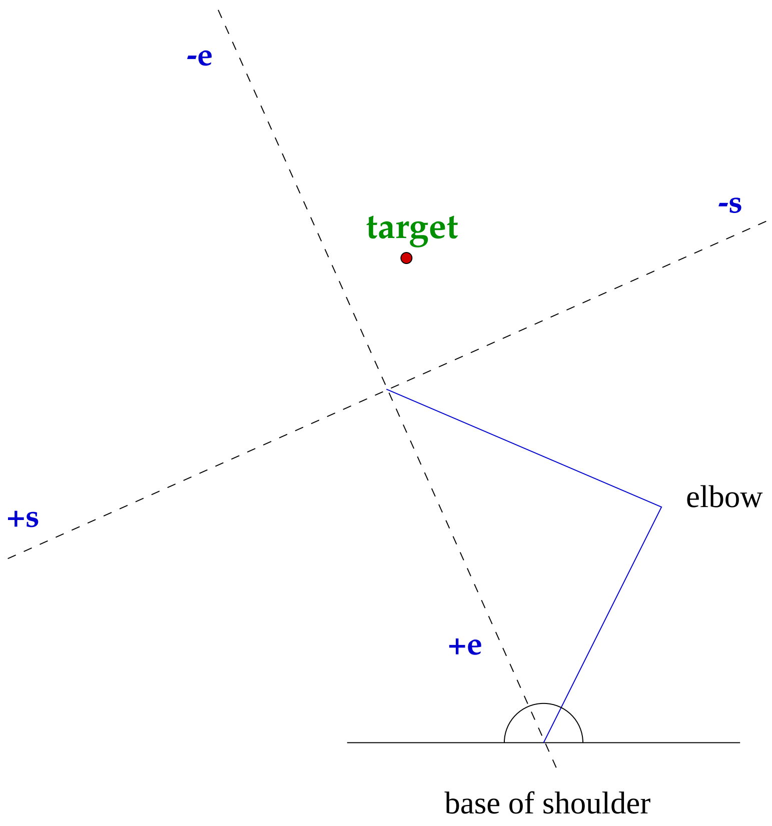
At an abstract level, the state encoder models the transformation that is performed by granule cells and associated inhibitory interneurons with information delivered by the mossy fibers. This transformation is thought to provide (via the *parallel fibers* (PFs)) a sparse, expansive encoding of the incoming state information [1]. In the model, this information consists of delayed plant state, recent motor commands, and other contextual signals. The sparse, expansive code is computed using a *CMAC* (Cerebellar Model Arithmetic Computer) [1, 17].

2.3 The Extra-Cerebellar Error Corrector

For a given movement, the cerebellar control module may generate a pattern of muscle activity that does not bring the arm to rest at the target position. In these cases, the role of the hard-wired extra-cerebellar (EC) module is to produce one or more corrective movements. However, it is not necessary, that a perfect correction be produced. Rather, the minimal constraints are that:

1. The corrective movement makes non-trivial progress toward the goal.
2. The movement generally does not overshoot the goal.

In the model, corrective movements are generated by short, constant bursts of muscle activity. The muscle burst pattern is determined as a function of the position of the wrist relative to the target. Errors along the Y-axis $(+e/-e)$ translate directly into movements of the elbow, whereas errors along the X-axis yield movements of the shoulder. This crude algorithm does not necessarily move the wrist directly toward the goal, but does sufficiently satisfy both the constraints stated above.



The EC module functionality may be implemented by a variety of brain regions. A very crude hard-wired mechanism, which serves to bootstrap the infant system may involve pathways such as the tectal-spinal tract. As the system matures, the basal ganglia may be involved in the triggering of corrective movements that are stored in the motor cortices and/or the cerebellum itself [9].

3 Learning Algorithm

When the cerebellum does not bring the arm to the target (and the extra-cerebellar module generates a corrective movement) the learning algorithm updates the PF-to-PC synapses in order to improve the cerebellum's movement generation performance. The *credit assignment problem* consists of two components [6]. The *structural* credit assignment problem is that of determining *which* synapses should be updated. The *temporal* credit assignment problem is that of properly attributing error signals received at one instant to synaptic events that have taken place in the recent past.

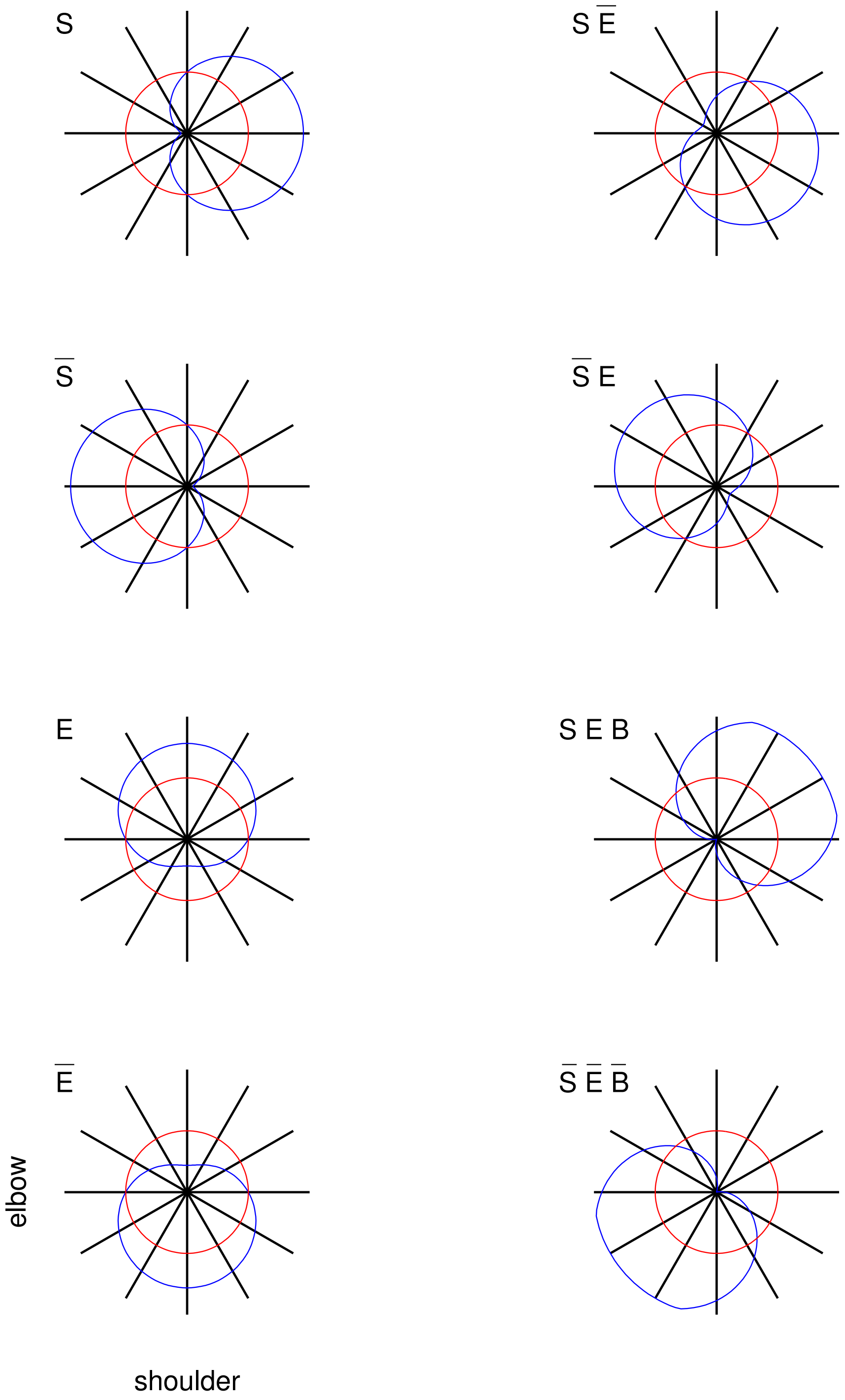
3.1 Climbing Fibers Carry Error-Corrective Information

The inferior olive (IO) is known to be the source of CF inputs into the cerebellum, which are thought to carry some form of error information related to the current movement [10]. Studies in cat show that when a limb is passively moved by an external force, CFs fire in response to the proprioceptive inputs from the the limb [5]. In addition, the CFs are differentially responsive to the direction of limb movement. During active movements (such as stepping), some CF responses are inhibited, except at the very end of movement, where certain perturbations are likely to occur (e.g. due to misanticipation of the height of the step).

In the model, CFs deliver error information to the cerebellum's APGs (there is exactly one CF for each APG). The CF behavior draws in two key ways from the behavior observed in cat:

1. From the perspective of the cerebellar-generated movement, the corrective movement made by the EC module is seen essentially as passive (or unexpected) [2]. Thus, CF signals only deviate from their background firing rate when corrective movements are initiated.
2. The increase (or decrease) in CF firing probability is a function of the rate of stretch of the muscles. Each CF is maximally responsive to stretches in the muscles that are activated by the corresponding APG. In other words, the corrective movement generated by the EC system is spatially coded in the array of CFs.

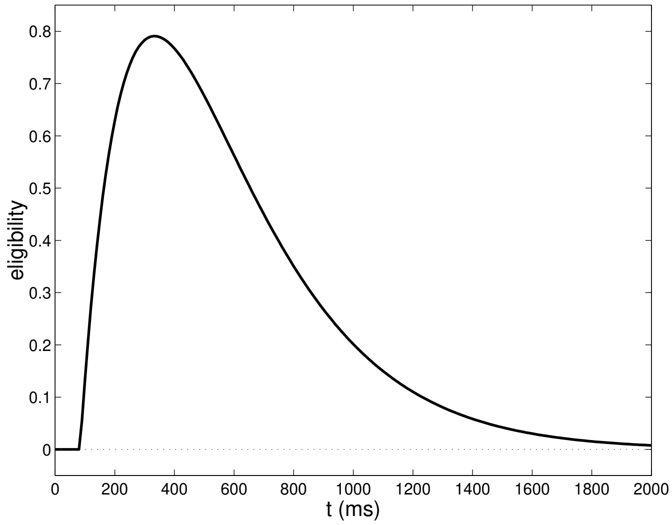
The figure to the right demonstrates the modeled CF responses to movements in joint space. Each sub-plot corresponds to the CF of the indicated APG. For each possible direction in joint space (all combinations of shoulder and elbow movements), the distance from the origin to the blue curve represents the probability of a CF spike (PR_k). Purely horizontal vectors correspond to shoulder movements, while vertical vectors correspond to movements of the elbow. The red circles indicate the base-line firing probability.



For the APG that drives the "S" muscle synergy (upper-left panel), the CF is most likely to fire in response to flexion of the shoulder joint (to the right of the origin), and is suppressed below baseline when an extension of the shoulder is sensed. Furthermore, the CF remains at baseline for pure movement of the elbow.

3.2 Temporal Credit Assignment and Eligibility Traces

The temporal credit assignment problem results, in part, from inherent delays in the sensorimotor system. This problem is addressed through the use of an *eligibility trace* on each synapse [15, 20]. This trace acts as a local memory of recent synaptic activity. The memory marks the synapse as being *eligible* for potential update in the event that later training inputs arrive from the IO.



In our model, we make use of a second-order eligibility trace [14, 19], whose time-course is shown above.

3.3 Changes to Synaptic Strengths

Updates to the PF-to-PC synaptic strength require the coincidence of eligibility and corresponding CF input:

$$\Delta w_{ij} = -\alpha \; g(\bar{e}_{ij}) \; (CF_k - CF_{base}),$$

where CF_k is the state of the climbing fiber that projects to APG_k (which contains PC_j), α is the learning rate, $g()$ is a non-linear thresholding function, and CF_{base} is the background firing rate of the climbing fiber. The binary state of CF_k is randomly selected at each timestep as a function of its probability of firing (PR_k). This form of weight update is similar to that used by [13, 19].

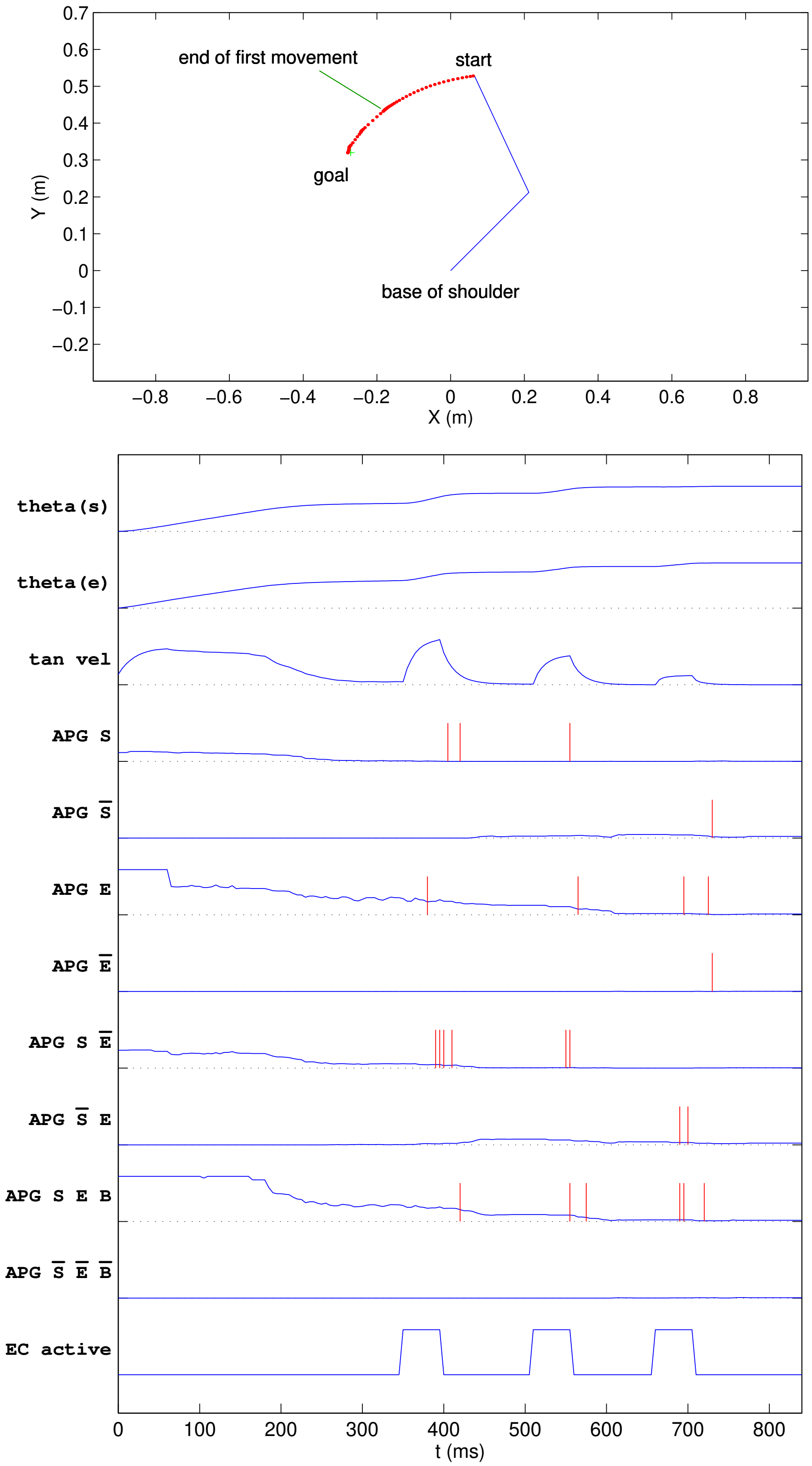
In cases where the corrective movement is aligned with APG_k, PR_k will be high. Thus, on average, the eligible synapses from the PFs to the PCs within APG_k will be decremented via LTD [1, 3, 11]. The next trial that the same movement is executed, the PCs of APG_k will be less active, resulting in a higher level of activity for APG_k. On the other hand, corrective movements in exactly the opposite direction will induce a small amount of LTP [18]. In essence, the trajectory that results will be a blend of the original and corrective movements. Over the course of repeated trials, the cerebellar module will ultimately be able to bring the arm to the target, and corrective movements will no longer be necessary.

4 Model Behavior

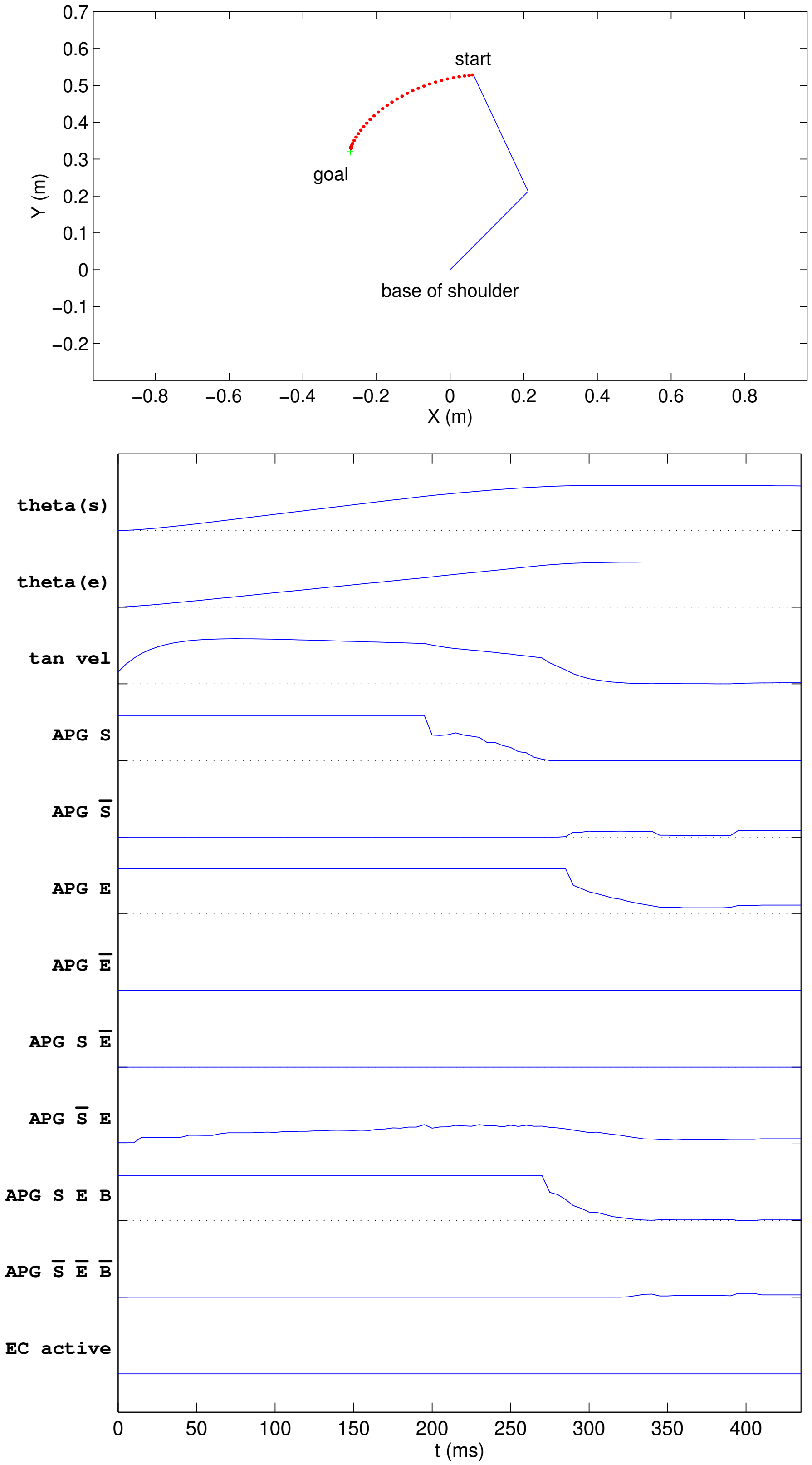
At the beginning of a trial, the arm is placed in an initial starting configuration (*start*), and the target location is presented to the controller (*goal*). The movement is then triggered, and proceeds under cerebellar control until the module turns off all descending motor commands, and the arm comes to rest. If the arm has stopped moving at a position that is not near the target, the EC system is triggered, which initiates a corrective arm movement. In this example, the same start and goal locations are chosen repeatedly.

On the sixth trial of learning, the movement produced by the cerebellar module has a duration of 300 ms, and brings the arm most of the distance to the target. The primary contributors to this movement are the APGs that activate muscle synergies E and SEB.

The three subsequent movements are produced by the EC module. The inferior olive responds to these movements by producing multiple CF firing events for the APGs corresponding to E, S, $S\bar{E}$, and SEB.

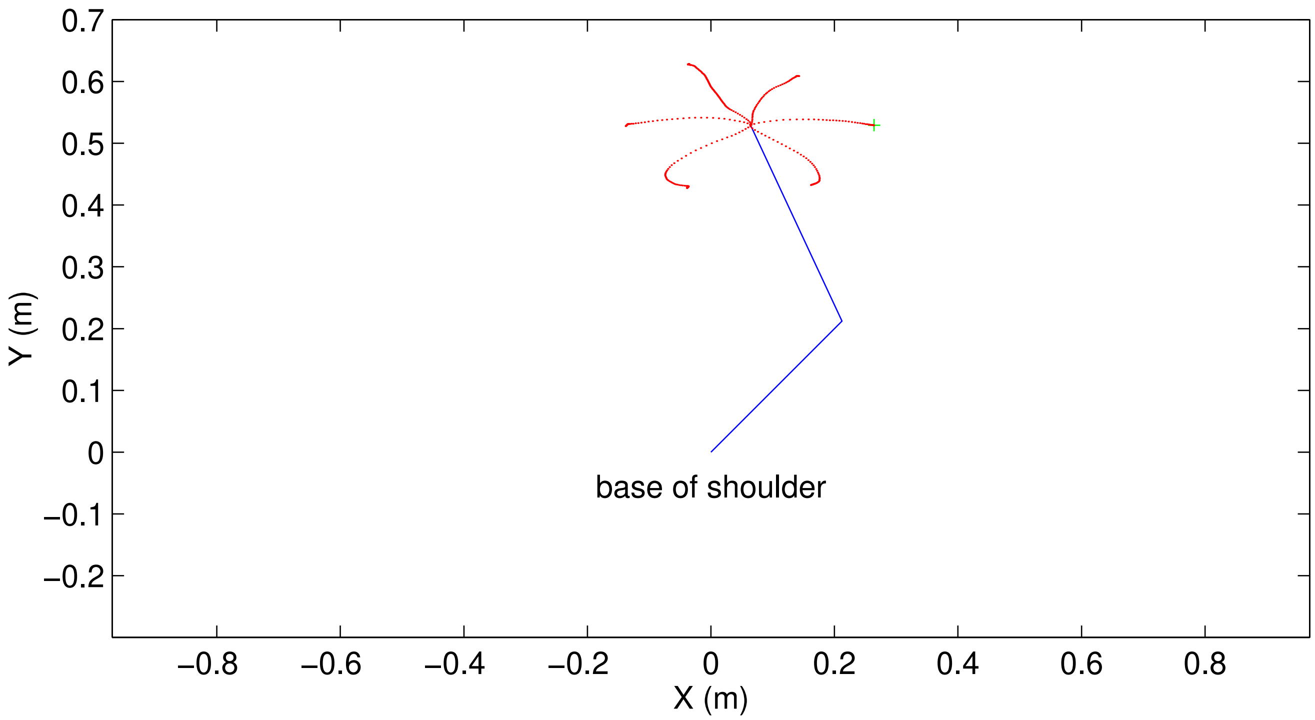


By trial 96, the cerebellar module has learned to generate an appropriate motor pattern for reaching the target, and no longer requires the aid of the EC module. APGs S, E, and SEB are the primary contributors to the final motor program. In addition, note that a braking function is performed by the APGs that correspond to $\bar{S}$ and $\bar{S}\bar{E}\bar{B}$.



Comparing the behavior with that of trial 6, note that even though the EC module generates a trajectory made up of multiple discrete hops, the final trajectory produced by the cerebellar module does not reflect this behavior. This happens because the cerebellum does not simply append corrective movements onto its original movement, but rather blends the EC-generated movement into its own.

The cerebellar controller is capable of learning single target movements within about 50-100 trials of experience. In addition, the cerebellum's ability to store and appropriately generalize over a set of targets has been tested. In these experiments, the start and target positions for a given trial are selected from a uniform distribution over an 80×80 degree range of the joint space. After 10,000 learning trials, the cerebellar controller produces movements with an average of 3.5mm error in the endpoint position (average movement distances are 29 cm). Furthermore, the controller hits the target on 98% of the trials (within 1 cm of the goal), and reaches to within 2 cm of the goal on 99% of the trials.



The above figure demonstrates the behavior of the arm for a single starting location (center of the spider), and six target locations, following this learning session. In all six cases, the cerebellum is able to bring the arm accurately to the target. However, the trajectory taken by the arm is not necessarily the most direct.

5 Conclusions

We have demonstrated a model of cerebellar control that can learn to produce motor signals in a *muscle synergy space* for the purposes of reaching toward a specified target. The IO infers errors in cerebellar-generated movements by observing subsequent corrective movements that are made by a crude, hard-wired extra-cerebellar controller.

The error-corrective teacher concept is one that we feel is fairly general. For example, consider a set of reflexes that are designed to respond to unexpected events (e.g. not reaching a goal position, unexpected muscle pulls, collisions with the environment, or slipping of a grasp). By "eavesdropping" on these reflexes (via proprioceptive or other sensory inputs), the IO can provide the cerebellum with a signed error-corrective signal. With this information, the cerebellum can learn to produce motor outputs similar to those generated by the reflexes, but in an anticipatory manner – correcting the movement before the reflex-generating conditions can occur.

Our model does not make use of *reference trajectories*, as has been assumed in other models of cerebellar control [13, 19]. There is little physiological evidence to support the idea that another subsystem provides to the cerebellum a detailed kinematic plan – a plan which happens to satisfy certain *dynamic* optimality criteria. Rather, we see such a higher-level system as providing a gross motor plan. The actual trajectory taken by the arm, then, arises from the interaction between the controller and the natural dynamics of the arm.

Relying on a crude error-corrector for indirect training information, combined with rejecting the notion of a high-quality reference trajectory, however, comes at a price. As we have demonstrated, our model in its current form fails to produce human-like trajectories. These non-straight-line movements result from the fact that the model is currently only constrained in its behavior to reach the target location at the end of the movement. In continuing work, we are examining how the model may be augmented to incorporate other forms of error correction that will bias the shape of the trajectory, as well as the endpoint of the movement.

References

[1] J. S. Albus. A theory of cerebellar function. ***Mathematical Biosciences***, 10:25–61, 1971.

[2] N. E. Berthier, S. P. Singh, A. G. Barto, and J. C. Houk. Distributed representations of limb motor programs in arrays of adjustable pattern generators. ***Journal of Cognitive Neuroscience***, 5:56–78, 1993.

[3] F. Crepel, N. Hemart, D. Jaillard, and H. Daniel. Cellular mechanisms of long-term depression in the cerebellum. ***Behavioral and Brain Sciences***, 19(3):347, 1996.

[4] E. E. Fetz, P. D. Cheney, and D. C. German. Cortico-motoneuronal connections of precentral cells detected by post-spike averages of emg activity in behaving monkeys. ***Brain Research***, 114:505–510, 1976.

[5] R. Gelman, A. R. Gibson, and J. C. Houk. Inferior olivary neurons in the awake cat: Detection of contact and passive body displacement. ***Journal of Neurophysiology***, 2:185–198, 1985.

[6] J. C. Houk and A. G. Barto. Distributed sensorimotor learning. In G. Stelmach and J. Requin, editors, ***Tutorials in Motor Behavior***, volume 2, pages 71–100. Elsevier, Amsterdam, 1992.

[7] J. C. Houk, J. T. Buckingham, and A. G. Barto. Models of the cerebellum and motor learning. ***Behavioral and Brain Sciences***, 19:368–383, 1996.

[8] J. C. Houk, J. Keifer, and A. G. Barto. Distributed motor commands in the limb premotor network. ***Trends in Neuroscience***, 16:27–33, 1993.

[9] J. C. Houk and S. P. Wise. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral-cortex - their role in planning and controlling action. ***Cerebral Cortex***, 5(2):95–110, 1995.

[10] M. Ito. A new physiological concept on cerebellum. ***Rev Neurol***, 1990.

[11] L. Karachot, R. T. Kado, and M. Ito. Stimulus parameters for induction of long-term depression in in-vitro rat purkinje-cells. ***Neuroscience Research***, 21(2):161–168, 1994.

[12] M. Katayamma and M. Kawato. Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models. ***Biological Cybernetics***, 69:353–362, 1993.

[13] M. Kawato and H. Gomi. A computational model of four regions of the cerebellum based on feedback–error learning. ***Biological Cybernetics***, 68:95–103, 1992.

[14] R. E. Kettner, S. Mahamud, H.-C. Leung, N. Sitkoff, J. C. Houk, B. W. Peterson, and A. G. Barto. Prediction of complex two-dimensional trajectories by a cerebellar model of smooth pursuit eye movement. ***Journal of Neurophysiology***, 1997.

[15] A. H. Klopf. Brain function and adaptive systems - a heterostatic theory. In ***Proceedings of the International Conference on Systems, Man, and Cybernetics***, 1974.

[16] L. E. Miller and J. C. Houk. Motor coordinates in primate red nucleus: Preferential relation to muscle activation versus kinematic variables. ***Journal of Physiology - London***, 488(2):533–548, 1995.

[17] W. T. Miller, F. H. Glanz, and L. G. Kraft. CMAC: An associative neural network alternative to backpropagation. In C. Lau, editor, ***Neural Networks. Theoretical Foundations and Analysis***, pages 233–240. IEEE Press, New York, NY, 1992.

[18] M. Sakurai. Synaptic modification of parallel fibre–purkinje cell transmission in in vitro guinea-pig cerebellar slices. ***J. Physiol. London***, 394:463–480, 1987.

[19] N. Schweighofer. ***Computational Models of the Cerebellum in the Adaptive Control of Movements***. PhD thesis, University of Southern California, 1995.

[20] R. S. Sutton and A. G. Barto. Time-derivative models of pavlovian reinforcement. In M. Gabriel and J. Moore, editors, ***Learning and Computational Neuroscience: Foundations of Adaptive Networks***, pages 497–537. MIT Press, 1990.

[21] W. T. Thach, H. Goodkin, and J. Keating. The cerebellum and the adaptive coordination of movement. ***Annual Review of Neuroscience***, pages 403–442, 1992.