ARTICLE **Communicated by Stephen Lisberger**

A Cerebellar Model of Timing and Prediction in the Control of Reaching

Andrew G. Barto
Andrew H. Fagg
Nathan Sitkoff
Department of Computer Science, University of Massachusetts, Amherst, MA 01003, U.S.A.

James C. Houk
Department of Physiology, Northwestern University Medical School, Chicago, IL 60611, U.S.A.

A simplified model of the cerebellum was developed to explore its potential for adaptive, predictive control based on delayed feedback information. An abstract representation of a single Purkinje cell with multistable properties was interfaced, using a formalized premotor network, with a simulated single degree-of-freedom limb. The limb actuator was a nonlinear spring-mass system based on the nonlinear velocity dependence of the stretch reflex. By including realistic mossy fiber signals, as well as realistic conduction delays in afferent and efferent pathways, the model allowed the investigation of timing and predictive processes relevant to cerebellar involvement in the control of movement. The model regulates movement by learning to react in an anticipatory fashion to sensory feedback. Learning depends on training information generated from corrective movements and uses a temporally asymmetric form of plasticity for the parallel fiber synapses on Purkinje cells.

1 Introduction

The neural commands that control rapid limb movements appear to comprise pulse components followed by smaller-step components (Ghez, 1979; Ghez & Martin, 1982), analogous to the pulse-step commands that control rapid eye movements (Robinson, 1975). In the case of eye movements, the pulse component serves to overcome the internal viscosity of the muscles, thus moving the eye rapidly to the target, whereupon the step component holds the eye at its final position. Limb movements involve more inertia than eye movements, so the pulse activation of the agonist muscle must end partway through the movement, and a braking pulse in the antagonist muscle is needed to decelerate the mass of the limb. Ghez and Martin (1982) showed that the braking pulse is produced by a stretch reflex in the antagonist muscle. The central control problem, therefore, is to terminate the pulse phase of the command sent to the agonist muscle at an appropriate time during the movement. The dynamics of the stretch reflex should then bring the movement to a halt at a desired end point. Since the pulse must terminate well in advance of the achievement of the desired end point, this is a problem of timing and prediction in control. In this article, we present a model of how the cerebellum may contribute to the predictive control of limb movements.

The model is a simplified version of the adjustable pattern generator (APG) model being developed by Houk and colleagues (Berthier, Singh, Barto, & Houk, 1993; Houk, Singh, Fisher, & Barto, 1990; Sinkjær, Wu, Barto, & Houk, 1990) to test the computational competence of a conceptual framework for understanding the brain mechanisms of motor control (Houk, 1989; Houk & Barto, 1992; Houk, Keifer, & Barto, 1993; Houk & Wise, 1995; Houk, Buckingham, & Barto, 1996). The model has a modular architecture in which single modules generate elemental motor commands with adjustable time courses, and multiple modules cooperatively produce more complex commands. The APG model is constrained by the modular anatomy of the cerebellar cortex and its connections with the limb premotor network, by the physiology of the neurons comprising this network, and by properties of cerebellar Purkinje cells (PCs). However, it is purposefully abstract to allow us to explore control and learning issues in a computationally feasible manner. The model presented here corresponds to a single module of the APG model consisting of a single unit representing a PC. This unit is modeled as a collection of nonlinear switching elements, which we call dendritic zones, representing segments of a PC dendritic tree.

Our previous modeling studies dealt mainly with two issues: (1) demonstration that a single module can learn to generate appropriate one-dimensional, variable-duration velocity commands (Houk et al., 1990) and (2) a preliminary demonstration that an array of 48 modules can learn to function cooperatively in the control of a simulated nondynamic, two-joint planar limb (Berthier et al., 1993). In these previous simulations, the input layer of the cerebellum, the representation of PCs, and the complexity of the learning problem were greatly simplified. In this article, we employ a more realistic input representation based on what is known about movement-related mossy fiber (MF) signals in the intermediate cerebellum of the monkey (Van Kan, Gibson, & Houk, 1993a) and the Marr-Albus architecture of the granular layer (Tyrrell & Willshaw, 1992). In addition, we use a more complex dynamic spring-mass system (although it is still one-dimensional), and we include realistic conduction delays in the relevant signal pathways. The model also makes use of a trace mechanism in its learning rule. Preliminary results appear in Buckingham, Barto, and Houk (1995) and Barto, Buckingham, and Houk (1996).

We first describe the nonlinear spring-mass system and discuss some of its properties from a control point of view. The following section presents the details of the model. We then present simulation results demonstrating the learning and control abilities of a single dendritic zone, followed by similar results for a model with multiple dendritic zones. We conclude with a discussion of these results.

2 Pulse-Step Control of a Nonlinear Plant

The limb motor plant has prominent nonlinearities that have a strong influence on movement and its control. The plant model used in this study is a spring-mass system with a form of nonlinear damping based on studies of human wrist movement (Gielen & Houk, 1984; Wu, Houk, Young, & Miller, 1990):

$$M\ddot{x} + B(\dot{x})^{\frac{1}{5}} + K(x - x_{eq}) = 0, \tag{2.1}$$

where x is the position (in meters) of an object of mass M (kg) attached to the spring, x_{eq} is the resting, or equilibrium, position, B is the damping coefficient, and K is the spring stiffness (see Figure 1a). This fractional power form of nonlinear damping is derived from a combination of nonlinear muscle properties and spinal reflex mechanisms, the latter driven mainly by feedback from muscle spindle receptors (Gielen & Houk, 1987). Setting $M = 1, B = 3$, and $K = 30$ produces trajectories that are qualitatively similar to those observed in human wrist movement (Wu et al., 1990).

Nonlinear damping of this kind enables fast movements that terminate with little oscillation. Figure 1b is a graph of the damping force as a function of velocity. As velocity decreases, the effective damping coefficient (the curve's slope) increases radically when the velocity gets sufficiently close to zero. This causes a decelerating mass generally to "stick" at a nonequilibrium position, thereafter drifting extremely slowly toward x_{eq}. We call the position at which the mass sticks (defined here as the position at which the absolute value of its velocity falls and remains below 0.9 cm/sec) the end point of a movement, denoted x_e. For all practical purposes, this is where the movement stops.

The control signal in our model sets the equilibrium value x_{eq}, which represents a central motor command setting the threshold of the stretch reflex (Feldman, 1966; Houk & Rymer, 1981). Pulse-step control is effective in producing rapid and well-controlled positioning of the mass in this system. As shown in Figure 1c, the control signal switches from a pulse level, $x_{eq} = x_p$, to a smaller step level, $x_{eq} = x_s$. Also shown are the time courses of the velocity (see Figure 1c, middle) and position (see Figure 1c, bottom) for the resulting movement. Inserting a low-pass filter in the command pathway, a common feature of muscle models, would produce velocity profiles more closely matching those of actual movements, but we have not been concerned with this issue.

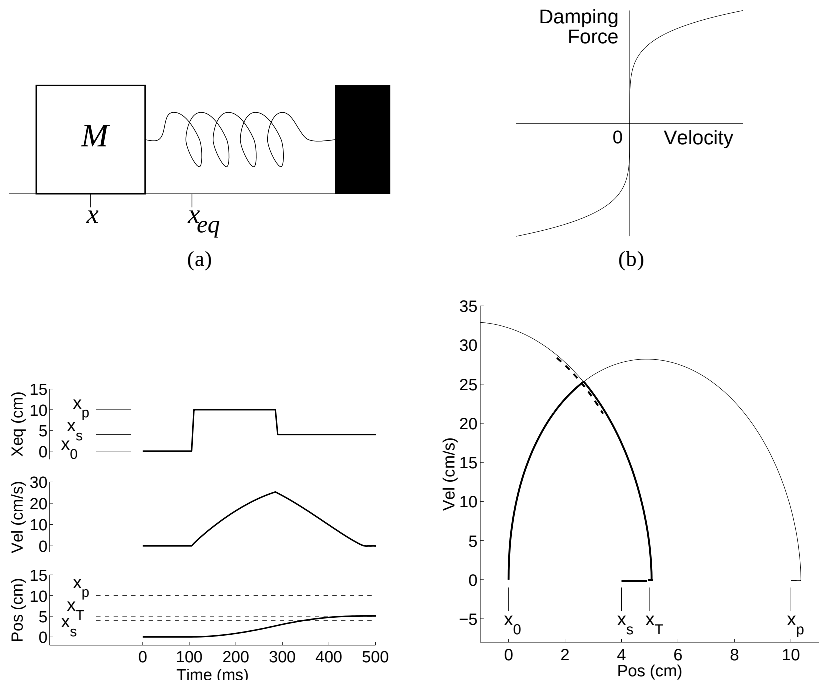


Figure 1: Pulse-step control of a simplified motor plant. (a) Spring-mass system. M, mass; x, position; x_{eq}, resting, or equilibrium, position. (b) Nonlinear damping force as a function of velocity. The plant's effective damping coefficient (the graph's slope) increases rapidly as the velocity magnitude decreases to zero. (c) Pulse-step control. Control of a movement from initial position $x_0 = 0$ to target end point $x_T = 5$ cm. *Top*: The pulse-step command. *Middle*: Velocity as a function of time. *Bottom*: Position as a function of time. (d) Phase-plane trajectory. The bold line is the phase-plane trajectory of the movement of panel c. The dashed line is a plot of the states of the spring-mass system at which the command should switch from pulse to step so that the mass will stick at the end point $x_T = 5$ cm starting from a variety of different initial states.

Figure 1d shows the phase-plane trajectory (velocity plotted against position) followed by the state of the spring-mass system during pulse-step control. When the pulse is being applied, the state follows a trajectory that would end at the equilibrium position $x_p = 10$ cm if the pulse were to continue. When the step begins, the state switches to the trajectory that ends at the equilibrium position $x_s = 4$ cm, but the mass sticks at the target end point, $x_T = 5$ cm, before reaching this equilibrium position. Thus, simply setting the equilibrium position to the target end point as suggested by the equilibrium-point hypothesis (Bizzi, Hogan, Mussa-Ivaldi, & Gister, 1992; Feldman, 1966, 1974) is not a practical solution to the end point positioning task for this system. The dashed line in Figure 1d is an approximate plot of the states at which the switch from pulse to step should occur so that movements starting from a variety of initial states will stick at $x_T = 5$ cm. This switching curve has to vary as a function of the target end point. If the switch from pulse to step occurs too soon (late), the mass will undershoot (overshoot) x_T.

In developing a model of pulse-step control of the limb, one can profit from analogies, where appropriate, with the extensive literature on pulse-step control of saccadic eye movements. However, an important difference between eye and limb control is the absence of a stretch reflex for regulating primate eye muscle activity (Keller & Robinson, 1971). As a consequence, models of the eye motor plant do not contain the nonlinear damping mechanism present in equation 2.1. The stretch reflex is important in generating a braking pulse in the antagonist muscles needed to decelerate the limb (Ghez & Martin, 1982). In fact, the stretch reflex is the predominant mechanism responsible for the entire decelerating portion of the trajectory in Figure 1D. The stretch reflex is also the main mechanism causing the limb to stick at a nonequilibrium position, as witnessed by the drift in limb position that occurs in deafferented patients who lack a stretch reflex (Ghez, Gordon, Ghilardi, Christakos, & Cooper, 1990). For eye movements, the prevention of postsaccadic drift is critically dependent on the precise regulation of the step component of the pulse-step command (Optican & Robinson, 1980). Although it is likely that the step component is also regulated for limb movements, relatively little is known about this mechanism. For the purposes of this article, we assume the presence of a fixed step component and rely on nonlinear damping for causing the limb to stick at an end point.

3 Model Architecture

Both limb and saccadic control systems are highly distributed, involving the cerebral cortex, basal ganglia, cerebellum, tectum, brain stem, and spinal cord. The focus here is on the special role of the cerebellum, which exerts its influence on movement by way of premotor networks. For both limb movements and saccades, there are two levels of premotor network. The upper level is the cortico-rubro-cerebellar network for the limb (Houk et al., 1993) and the tecto-reticulo-cerebellar network for saccades (Houk, Galiana, & Guitton, 1992; Arai, Keller, & Edelman, 1994). These upper-level networks feed control signals to a lower level comprising a propriospinal network for the limb (Alstermark, Lundberg, Pinter, & Sadaki, 1987a) and a brain stem burst network for saccades (Robinson, 1975). Since the emphasis in this article is on the cerebellar cortex, the premotor networks will be given only a formal representation. We assume that the propriospinal network, in analogy with the brain stem burst network, can generate only relatively

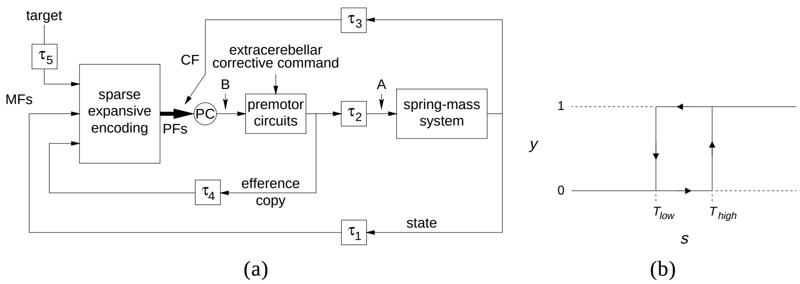


Figure 2: Model architecture. (a) Block diagram. PC, Purkinje cell; MFs, mossy fibers; PFs, parallel fibers; CF, climbing fiber; τ_i, $i = 1, \ldots, 5$, conduction delays. The labels A and B mark places in the feedback loop to which we refer in discussing the model's behavior. (b) Dendritic zone hysteresis. DZ activation, y, switches from 0 to 1 when the input weighted sum, s, exceeds threshold T_{high}, and switches from 1 to 0 when s drops below T_{low}.

crude commands that typically produce dysmetric movements when it operates on its own. However, the system is capable of orthometric control when the cerebellum and upper premotor networks are operative. In order to focus on the critical control functions of PCs in the cerebellar cortex, we will represent the cortico-rubro-cerebellar network as simply an inverting mechanism that converts the inhibitory output of PCs into a positive command signal. For simplicity, we further assume that the output of this cortico-rubro-cerebellar network acts directly on spinal output rather than functioning through the propriospinal network.

The model's main component is a single unit representing a cerebellar PC, whose input is derived from a sparse, expansive encoding of MF signals (see Figure 2a). In defining how the MFs encode information about the spring-mass system, we followed what is known about movement-related MF signals in the intermediate cerebellum of the monkey, where MFs exhibit discharge patterns involving diverse combinations of tonic and phasic components, as well as a variety of onset times relative to the time of movement onset (Van Kan et al., 1993a).

To represent this diversity, the model has a total of 2000 MFs, 800 of which encode information about single variables x, $\dot{x}$, x_{eq}, or x_T (200 MFs devoted to each), with the remaining 1200 MFs encoding information about pair-wise combinations of these variables. Each of the MFs representing a single variable uses a saturated ramp encoding. For example, as the mass's position increases, the firing rate of a pure position-related MF remains zero until a threshold is exceeded and then increases linearly until saturating at a maximum firing rate. The thresholds are distributed uniformly over the relevant variable ranges, and several slopes and saturation levels are used.[1] In addition, the signal conveyed by each pure position and velocity MF is delayed relative to spring-mass movement by an amount chosen uniformly at random from between 15 and 100 ms (τ_1 in Figure 2a). The delay ranges for this and the following types of MFs are within those observed for the intermediate cerebellum of the monkey (Van Kan et al., 1993a). The signals of the efference copy MFs (representing x_{eq}) are delayed between 40 and 150 ms (uniform random) relative to the motor command (τ_4 in Figure 2a). The signal of each target position MF (representing x_T) is delayed between 0 and 100 ms (uniform random) from the start of a trial (τ_5 in Figure 2a). The signal conveyed by each of the 1200 MFs representing pair-wise combinations of the single variables is a weighted sum of the signals of two single-variable MFs: 400 are combinations of pure x and $\dot{x}$ MFs, 400 are combinations of pure x and x_{eq} MFs, and 400 are combinations of pure x_T and $\dot{x}$ MFs. Within these classes, the pairs of MFs were chosen uniformly at random, and the weights, which are positive and sum to one for each MF, were selected uniformly at random. The relative number of MFs in these various classes is consistent with the proportions observed by Van Kan et al., (1993a). The total number of MFs was chosen for computational reasons: we wanted to ensure that the model could accurately represent the transformation required by the control task. We did not rule out the possibility that fewer MFs might also suffice.

We set the efferent delay from the PC to the spring-mass system via premotor circuits to 100 ms (τ_2 in Figure 2a), which is within the range observed for this pathway in the intermediate cerebellum of the monkey (Van Kan, Houk, & Gibson, 1993b), although we experimented with other values as well (see section 5.1). With this delay, the MF delay ranges described imply that the onset of movement-related discharge of the MFs that use efference copy information can lead movement onset by as much as 60 ms or lag it by as much as 50 ms. On the other hand, movement-related discharge of MFs relying on only proprioceptive information always lags movement by between 15 and 100 ms.

Patterns of MF activity are recoded to form sparse activity patterns over 40, 000 binary parallel fibers (PFs), which synapse on the PC. This form of PF state encoding is similar to that used in numerous models of the cerebellum, such as those of Marr (1969) and Albus (1971). We selected this number of PFs to ensure that the model could realize the required transformation. With

[1] Thresholds are distributed at uniform intervals over the ranges of the relevant variables ($[-0.5, 7.5]$ cm for x; $[-25, 25]$ cm/sec for $\dot{x}$; $[0, 1]$ for x_{eq}; and $[3, 7]$ cm for x_T). The slopes were set so that the ramp covers 50%, 25%, or 12.5% of the variable's range. Half of the slopes are negative, so that the MF decreases in activity as its coded variable increases. Saturation levels differ slightly as a function of threshold, with higher thresholds being associated with higher saturation levels. This roughly normalizes the average activity level of the MFs.

as few as 30,000 PFs, learning progresses at a slower rate and asymptotes at a higher average end point error. However, with 60,000 PFs, an improvement in learning performance is not observed. Each PF is the output of a granule unit that sums excitatory input from four randomly chosen MFs. We assumed that local competition takes place among granule units, allowing only 80 of the units to fire (output $= 1$) at the same time. Marr (1969) and Albus (1971) hypothesized that this competition arises from inhibitory interactions through Golgi cells. We implemented this competition by dividing the granule cell population into 80 Golgi-cell receptive fields, each comprising 500 granule units, and allowing only the most active unit in each field to fire at any time step of the simulation (although the model does not explicitly contain units representing Golgi cells). Thus, at each time step, the PF input to the PC is a pattern of 40,000 binary values containing 80 ones.

The PC in the model consists of a number of dendritic zones (DZs) representing segments of the dendritic tree. Our representation of DZs is motivated by observations of plateau potentials in PC dendrites (Llinás & Sugimori, 1980; Ekerot & Oscarsson, 1981; Campbell, Ekerot, Hesslow, & Oscarsson, 1983; Andersson, Campbell, Ederot, Hesslow, & Oscarsson, 1984). These long-lasting potentials (up to several hundred ms in duration) represent a form of bistability, which results from hysteresis produced by the dendritic ion channel system. A number of researchers have suggested that dendritic or neuronal bistability resulting from hysteresis can be computationally useful (Hoffman, 1986; Benson, Bree, Kinahan, & Hoffman, 1987; Kiehn, 1991; Houk et al., 1990; Wang & Ross, 1990; Gutman, 1991, 1994), and Yuen, Hockberger, and Houk (1995) showed how these properties can arise in a biophysical model of the PC dendrite.

Each DZ in the model is a linear threshold unit with hysteresis. Let $s(t) = \sum_i w_i(t)\phi_i(t)$, where $\phi_i(t)$ denotes the activity of PF i at time t and $w_i(t)$ is the efficacy, or weight, at time step t of the synapse by which PF i influences the PC dendritic segment comprising the DZ. The activity of the DZ at time t, denoted $y(t)$, is either 1 or 0, respectively, representing a state of high or low activity. DZ activity depends on two thresholds: T_{low} and T_{high}, where $T_{low} < T_{high}$. The activity state switches from 0 to 1 when $s(t) > T_{high}$, and it switches from 1 to 0 when $s(t) < T_{low}$ (see Figure 2b). If $T_{high} = T_{low}$, the DZ is the usual linear threshold unit. Unlike plateau potentials, which tend to reset spontaneously after a few hundred milliseconds (Llinás & Sugimori, 1980; Campbell et al., 1983; Andersson et al., 1984), the state of a DZ remains constant until actively switched by input. We have not yet explored the consequences of spontaneous resetting in our model. In the simulations reported below, we investigated the effects of several settings of T_{high} and T_{low}.

The PC's overall activity level at any time is equal to the fraction, f, of its DZs that are in state 1 at that time. In a more detailed model, the PC would inhibit nuclear cells, thereby regulating the buildup of activity in cortico-rubro-cerebellar loops from which motor commands are derived

(Houk et al., 1993). The simpler model described here does not include an explicit representation of these premotor circuits. The motor command, x_{eq}, is simply defined to be $4f + 10(1 - f)$, which means that when the PC is maximally active ($f = 1$), the equilibrium position is the "near" position of 4 cm, and when it is minimally active ($f = 0$), the equilibrium position is the "far" position of 10 cm. These values determine the range of target end points for which the model is able learn accurate positioning commands, but the model is not otherwise sensitive to the specific values. This definition of the motor command reflects the inhibitory effect the PC would have on cortico-rubro-cerebellar loops. As a result, pauses in the PC's activity would disinhibit activity of premotor circuits, which activate an agonist muscle for rightward movement.

We studied three versions of the model that differ in the number of DZs and how the PF input is distributed among them. In the simplest version, a single DZ receives input from all 40,000 PFs. In the other versions, the PC consists of 8 DZs. In one of these, each DZ receives input from all of the PFs; in the other, each DZ receives input from a separate subfield of 5000 PFs. The latter version of the model, which is more realistic due to the orthogonal relationship between PFs and the flattened dendritic trees of PCs in the cerebellum, learned somewhat slower than the other 8-DZ model, but its behavior was similar in other respects (see sec. 5.2).

4 Learning

All the DZs comprising the PC in the model receive training information from a signal representing discharge of a climbing fiber (CF). This signal provides information about the spring-mass system with a delay of 20 ms (τ_3 in Figure 2a), which is within the physiological range for CF signals in cats (Gellman, Gibson, & Houk, 1983). The nature of the training information supplied by the model's CF is an extrapolation of what is known about the responsiveness of proprioceptive CFs, which respond to particular directions of limb movement and appear to signal "unexpected" passive movements, being suppressed during active (hence expected) movements (Gellman, Gibson, & Houk, 1985). We hypothesize that by monitoring the proprioceptive consequences of corrective movements generated by other structures, modules of the cerebellum can learn to regulate motor commands so that they produce more efficient and accurate movement. We follow Berthier et al. (1993) in assuming that the propriospinal premotor network generates simple corrective movements when a movement is inaccurate. These corrective movements do not have to be particularly accurate themselves; they only need to reduce the end point error. The literature on which these assumptions are based is reviewed in some detail in an earlier work (Berthier et al., 1993; see section 6). Although a sequence of such corrective movements alone can produce small final end point error, the sequence would be slow and dynamically erratic. The role of the cerebellum, we hypothesize, is to eliminate the need for corrective movements by learning to suitably regulate the initial movement.

In the model, whenever the mass is coming to rest at a point not near the target end point, an extracerebellar motor command in the form of a single rectangular pulse is generated, causing movement in the correct direction.[2] In response to each rightward corrective movement, the model's directionally-sensitive CF produces a single discharge. The CF is silent during leftward corrective movements (although the CF to a module activating a muscle for leftward movement, if one were present in the model, would discharge in this case). This follows a key assumption that the responsiveness of a PC's CF to movement in a given direction is matched to the degree to which that PC's module is capable of contributing to movement in that direction (see Berthier et al., 1993, for additional details). We also assume a low background firing rate for the CF in the absence of corrective movements. Letting $c(t)$ denote CF activity at time step t, the model implements these assumptions by setting $c(t) = 1$ at the initiation of each rightward corrective movement, $c(t) = 0$ for the remainder of the rightward movement, $c(t) = 0$ during leftward corrective movements, and otherwise $c(t) = \beta = 0.025$, which represents a low background firing rate.[3]

As a result of a corrective movement, the weights of each DZ should change so that the PC contributes to a more accurate motor command. In response to a rightward corrective movement, the weights should change so as to increase the duration of the pulse phase of the command (since the movement stopped short of the target end point), and in response to a leftward corrective movement, the weights should change so as to decrease the duration of the pulse phase of the command (since the movement overshot the target end point). Accomplishing this with a simple learning rule is difficult because the training information in the form of CF activity is significantly delayed with respect to the relevant DZ activity due to the combined effects of movement duration and conduction latencies. To learn under these conditions, the model adopts Klopf's (1972, 1982) hypothesis of synaptic "eligibility traces." Appropriate activity at a synapse is hypothesized to set up a synaptically local memory trace that makes the synapse "eligible" for modification if and when the appropriate training information arrives within a short time period. This allows the learning rule to modify synaptic weights based on synaptic actions that occurred prior to the avail-

[2] Whenever the mass has been "stuck" for 150 ms more than 0.1 cm from the target end point, the motor command, x_{eq}, is set to a value that causes movement toward the target. Specifically, $x_{eq} = x_T + a$ for undershoot and $x_{eq} = x_T - a$ for overshoot, where $a > 0$ was chosen to be sufficiently large to overcome the high low-velocity viscosity. Here, we used $a = 5$ cm.

[3] We experimented with a more realistic representation of background activity in which $c(t) = 1$ with probability β for each background time step t. The results were essentially the same, except that the learning process required about 2.5 times as many trials.

ability of the relevant CF training information. An example eligibility trace for one PF-to-PC synapse is shown in the bottom plot of Figure 3a. This eligibility trace spans the interval from the time of the presynaptic PF's activity until later CF discharges occur (plot CF in the figure), when this synapse's weight is modified.

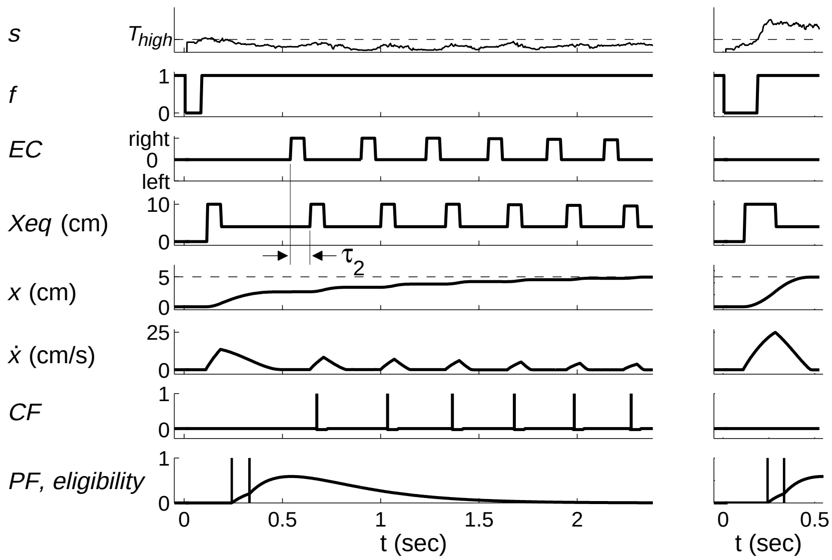
To define the learning rule, we have to specify how synapses become eligible for modification and how CF activity alters the synaptic weights based on the eligibility of synapses. We first describe the eligibility process. The idea is that a synapse becomes eligible for modification if its presynaptic PF was active in the recent past at the same time that the synapse's DZ was in state 1. Eligibility then persists as a graded quantity—a trace—that reflects both how frequently and how long in the past this eligibility-triggering condition was satisfied for that synapse. Although learning is not sensitive to the exact time course of eligibility traces, a synapse should reach peak eligibility at roughly the time at which a relevant CF discharge would reach the PC. By a relevant CF discharge, we mean one produced by a correction to a movement that was influenced by the eligibility-triggering activity at the given synapse.

One of the simplest methods for computing eligibility is to simulate a second-order linear filter whose input is 1 whenever the triggering condition is satisfied and 0 otherwise. The filter's parameters were set so that its impulse response rises to a peak about 255 ms after the triggering event, and then decays asymptotically to zero with a time constant of approximately 600 ms. A synapse is therefore maximally eligible 255 ms after the triggering event and becomes effectively ineligible approximately 2 sec later, assuming no additional triggering events occur (see the bottom plot of Figure 3a). This time course is appropriate for the movement durations and conduction delays in this model. An intracellular signal transduction mechanism for producing this kind of eligibility trace was proposed in Kettner et al. (1997). We also found it useful to limit the magnitude of eligibility so that prolonged periods during which the triggering condition is satisfied do not lead to excessively high eligibility, and hence to large weight changes. In the discussion, we comment on the biological realism of the eligibility idea.

Letting $e_i(t)$ denote the eligibility of synapse i at time t, the model generates the eligibility trace for each synapse i by the following difference equations involving the intermediate variables $\bar{e}_i$ and $\hat{e}_i$:

$$\bar{e}_i(t) = .98\bar{e}_i(t-1) + .02y(t)\phi_i(t),$$
$$\hat{e}_i(t) = .98\hat{e}_i(t-1) + .02\bar{e}_i(t-1),$$
$$e_i(t) = \min\{\hat{e}_i(t), 0.1\},$$

where $y(t)$ is the binary activity state of the synapse's DZ at time step t, and $\phi_i(t)$ is the activity of the presynaptic PF. Each time step in the model represents 5 ms of real time.



(a) (b)

Figure 3: Single DZ behavior. The target end point, x_T, was switched from 0 to 5 cm at time 0. Shown are the time courses of the DZ's summed input, s; activation state, f; extracerebellar corrective command, EC; motor command, x_{eq} (after the 100 ms efferent delay τ_2); and the position, x, and velocity, $\dot{x}$, of the mass for a movement that started at initial position $x_0 = 0$. Plot CF shows climbing fiber activity, and the bottom plot shows the binary activity of an arbitrarily selected PF together with the eligibility trace of its synapse onto the DZ. (The eligibility trace's amplitude is scaled up to make it easily visible; peak eligibility here is 0.029.) $T_{low} = 0.8$ and $T_{high} = 1$. (a) Early in learning (four trials). DZ state switched to 1 too soon, which caused the mass to undershoot the target. A sequence of six rightward corrective movements was generated by the extracerebellar system (EC) because all but the last failed to bring the mass close enough to the target. Each corrective movement caused a CF discharge. Each discharge of the selected PF contributed to the eligibility trace because the DZ was in state 1 at these times. The weight of this PF's synapse (not shown) decreased when the CF discharge coincided with nonzero eligibility. (b) Late in learning (1000 trials). The model consistently produced accurate reaching with fast, smooth movements requiring no corrections (and hence with no CF discharges). To accomplish this, the DZ learned to switch to 1 well before (about 300 ms) the end point was reached.

The remainder of the model's learning mechanism is a rule determining how the weights of eligible synapses are altered by CF activity. The logic of this learning rule is a result of the following reasoning. When the weights of a DZ's eligible synapses decrease, that DZ becomes less likely to switch to state 1 in the future when a situation (represented by a pattern of PF activity) is encountered that is similar to the one that was present when the eligibility trace was initiated. This tends to prolong the pulse phase of the motor command by delaying the DZ's contribution to PC inhibition, which increases movement duration and moves the initial movement's end point to the right. Thus, the weights of eligible synapses should decrease as a result of each rightward corrective movement. Since the CF produces a discharge on each rightward corrective movement, CF discharge should cause depression of the eligible synapses. On the other hand, increasing the weights of a DZ's eligible synapses makes that DZ more likely to switch to state 1 under similar circumstances in the future, which tends to shorten the pulse phase of the motor command, thus decreasing movement duration and moving the end point leftward. Therefore, leftward corrective movements should cause potentiation of the eligible synapses. In the model, this is accomplished by letting the CF signal drop below its background rate during leftward corrective movements.

The following learning rule implements this logic:

$$\Delta w_i(t) = -\alpha e_i(t)[c(t-\tau_3) - \beta],$$
$$w_i(t) = \max\{w_i(t-1) + \Delta w_i(t), 0\}, \tag{4.1}$$

where $\alpha > 0$ is a parameter influencing the rate of learning that was set to 2×10^{-3} in the simulations described below.[4] The term β implies that weights do not change during background CF activity and that eligible weights increase during leftward corrective movements when CF activity drops below its background rate. Note that since $\beta \ll 1$, weight increases are much smaller than weight decreases. The term $c(t-\tau_3)$, where $\tau_3 = 20$ ms is the CF conduction delay, is the CF signal that reaches the synapse at time t. Since eligibility, $e_i(t)$, is a multiplicative factor, weights change in proportion to their degree of eligibility. All the DZs comprising the model's PC learn independently according to this rule.

To summarize the model's learning mechanism, training information is supplied by CF responses to corrective movements. The CF for the single module described here discharges reliably in response to rightward corrective movements. This follows from the specificity of the CF system and the assumption that this module controls an agonist for rightward movement. Rightward corrective movements therefore raise the CF's activity above its background rate. For leftward corrective movements, the CF's activity decreases slightly below its background rate. The weights of the synapses from PFs to the DZs comprising the model's PC change in response to CF activity so that the duration of the pulse phase of the motor command is increased

[4] This was chosen to be such a small value because the resultant change in PC activation due to each learning step could be 80 times larger since 80 of the PF inputs are 1 at each time step.

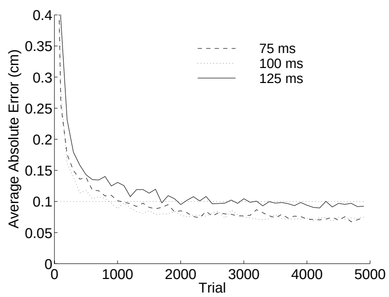
in the case of rightward corrective movements and decreased in the case of leftward corrective movements. The model uses eligibility traces to bridge the time interval between the activity of the DZs and the relevant later CF activity. A synapse becomes eligible for modification when presynaptic activity coincides with the postsynaptic DZ being in state 1. Eligibility is realized as a synaptically local trace that persists for several seconds after the coincidence of pre- and postsynaptic activity. When CF activity rises above its background level, the weights of the synapses are depressed in proportion to their current degree of eligibility, which tends to lengthen the pulse phase of the command. When CF activity falls below its background level, synapses are facilitated in proportion to their eligibility, which tends to shorten the pulse phase of the command.

5 Simulations

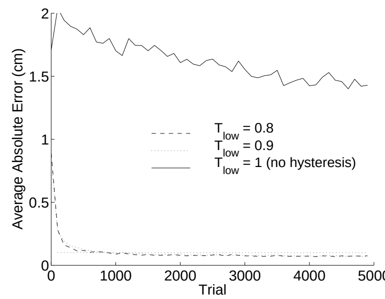
5.1 Single Dendritic Zone. We performed a number of simulations of a single DZ learning to control the nonlinear spring-mass system. We trained the DZ to move the mass from initial positions selected randomly from the interval [0, 2 cm] to a target position randomly set to 3, 4, or 5 cm. DZ state 0 corresponded to the pulse phase of a motor command, which set a "far" equilibrium position of 10 cm; DZ state 1 corresponded to the step phase, which set a "near" equilibrium position of 4 cm (see section 2). Each simulation consisted of a series of trial movements. At the beginning of the first trial movement, we randomly initialized all 40,000 weights so that the weighted sum, s, fell uniformly between 0.68 and 1.48 for any initial pattern of PF activity. Each trial began when the state of the DZ was set to 0. We initialized each eligibility trace, $e_i(t)$, to 0 ($\bar{e}_i(t)$ and $\hat{e}_i(t)$ were also set to 0). We also set the pattern of MF activity to be consistent with the initial state of the spring-mass system.

To study the influence of loop delay on learning and performance, we conducted simulations in which the loop delay was varied by setting the efferent delay (τ_2 in Figure 2a) to 75, 100, or 125 ms. Figure 4a shows how the end point error decreased with trials for the various efferent delays (with $T_{low} = 0.8$ and $T_{high} = 1$). The DZ's behavior is largely insensitive to this range of delays. In each case, the average absolute error rapidly dropped below 0.1 cm (see the dotted line in Figure 4a), the trigger criterion for the extracerebellar corrective movement. In all the simulations reported below, we set $\tau_2 = 100$ ms. However, the model's behavior is sensitive to the amount of DZ hysteresis. Figure 4B shows how end point error decreased over trials for several different values of T_{low}, with T_{high} fixed at 1. Learning was seriously disrupted when there was no hysteresis ($T_{low} = T_{high} = 1$). In all simulations reported below, $T_{low} = 0.8$ and $T_{high} = 1$, unless noted otherwise.

Figure 3 shows time courses of relevant variables at different stages in learning to move to target end point $x_T = 5$ cm from initial position $x_0 = 0$.



(a)



(b)

Figure 4: End point error ($|x_e - x_T|$) as a function of trial number for single DZ learning. Each plotted point is an average over a bin of 50 trials of 10 learning runs. The dotted horizontal line shows the minimum threshold above which corrective movements were generated. (a) Effect of loop delay. Plots for efferent delays (τ_2) of 75, 100, and 125 ms. Here, $T_{low} = 0.8$ and $T_{high} = 1$. (b) Effect of hysteresis. Plots for $T_{high} = 1$ and $T_{low} = .8, .9,$ and 1 (no hysteresis). Efferent delay (τ_2) was 100 ms.

Early in learning (four trials, panel A), the DZ switched back to state 1 too soon (plot f), which caused the mass to undershoot the target. Because of this undershoot, the extracerebellar system (EC) generated a corrective movement. In fact, a sequence of six corrective movements was generated because all but the last failed to bring the mass close enough to the target. Each corrective movement caused a CF discharge. The resulting movement accurately reached the target, but along a slow and irregular trajectory.

Plotted at the bottom of Figure 3 is the binary activity of an arbitrarily selected PF and the eligibility trace of its synapse onto the PC. Note that each discharge of the PF contributed to the trace because the DZ was in state 1 at these times. The weight of this PF's synapse (not shown) decreased when the CF discharge coincided with nonzero eligibility. The decrease of this weight, along with decreases of many others, tended to prolong the pulse phase of the motor command by delaying the DZ's switch to state 1. None of the synaptic weights increased during this trial because there was no leftward corrective movement (see Figure 7a for an example of a trial with a leftward corrective movement). Later in learning (after 1000 trials, Figure 3b), the model consistently produced accurate reaching with fast, smooth movements requiring no corrections (and hence causing no CF discharges). To accomplish this, the DZ learned to switch to state 1 well before (about 300 ms) the end point was reached.

Figure 5a shows the paths of a number of movements controlled by a well-trained DZ. The initial position of the mass for each movement is indicated by the circle at the left end of each line, and the target end points are indicated by the vertical dashed lines. The asterisk on each path marks

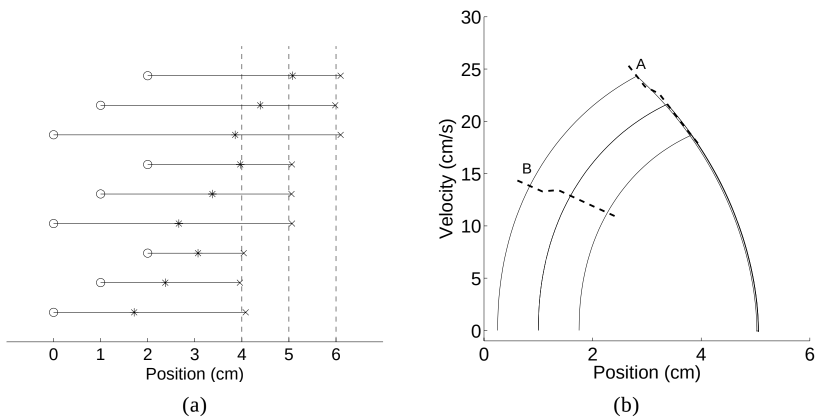


Figure 5: (a) Paths of a number of movements controlled by a single well-trained DZ. The path of each movement is shown by a horizontal line. The initial position of the mass for each movement is indicated by the circle at the left end of each line, and the target end points are indicated by the vertical dashed lines. The asterisk on each path marks the position of the mass when the PC state switched from 0 to 1. The actual end point of the movement is indicated by the $\times$ at the right end of each line. The DZ switches state well before a movement ends. The model used a motor efference delay of 100 ms. (b) Switching curves. Phase-plane portraits of switching curves for target $x_T = 5$ cm learned by the model. Two switching curves and three example movement trajectories are shown. See the text for an explanation.

the position of the mass when the DZ switched state from 0 to 1. The end point of each movement is indicated by the $\times$ at the right end of each line. One can see that the movements were accurate across a range of initial positions and target end points. It is apparent that the DZ switched state well before the end of each movement.

Figure 5b shows two representations (the dashed lines labeled A and B) of the switching curve learned by the DZ for target $x_T = 5$ cm, together with three sample phase-plane trajectories. Switching curve A is the switching curve as it appears after the efferent delay τ_2, that is, as seen from the point marked A in Figure 2a. When the spring-mass system's state crosses this curve, the command input to the spring switches from pulse to step. Clearly, it is positioned correctly to cause the mass to stick close to the desired end point for a range of initial conditions. Switching curve B, on the other hand, is the switching curve as it appears before the efferent delay, that is, as seen from the point marked B in Figure 2a. This curve is crossed 100 ms before switching curve A is crossed due to the 100 ms efferent delay. When the system state crosses this curve, the DZ switches state. One can see that the DZ learned to switch 100 ms before the motor command must switch at the spring itself, appropriately compensating for the 100 ms latency of the efferent pathway. To do this, the DZ effectively learned to “recognize” the patterns of PF activity that were present at its synapses when the system state crossed switching curve B. It is important to note that due to the various delays in the MF pathways, the recognized PF patterns actually encoded information about the spring-mass state as it was between 15 and 100 ms earlier.

5.2 Multiple Dendritic Zones. We simulated two versions of the model in which the PC consists of eight DZs. In each case, the PC's activity level at any time is the fraction of its DZs that are in state 1 at that time (see section 3). In one version, each DZ receives input from all of the PFs (uniform model); in the other, each DZ receives input from a separate subfield of 5000 PFs (subfield model). Figure 6 shows how the end point error decreased with trials for these two variations, as well as for the single DZ model. The uniform model learned significantly faster than the others and reached a smaller final error. We believe this is due to the fact that the uniform model has many more adjustable parameters than the others so that there are many different potential solutions: the algorithm is constrained only to reduce the end point error, which can be accomplished in many ways. To save computer time, we restricted further simulations to the uniform model, but it is likely that the subfield model would have produced similar results.

Figure 7 illustrates some details of the behavior of the uniform model. Panel a is analogous to Figure 4a except that it shows a trial in which there was a single leftward corrective movement instead of multiple rightward corrective movements. Note that the leftward corrective movement did not generate CF discharges but instead slightly depressed CF background rate. Unlike the single DZ case, here the motor command was graded due to the varying contributions of the eight DZs. This variety was due to the differing initial weights of the DZs. Later in learning (1500 trials, panel b), fast, accurate, and smooth movements were accomplished, although the motor command was not a pure pulse.

We also investigated the effects of different levels of hysteresis on the uniform model by fixing T_{high} to 1 and varying T_{low}, as we did for the single DZ system (Figure 4b). Unlike the single DZ case, hysteresis had no significant effect on the learning rate of the uniform model. However, we did note that without hysteresis ($T_{low} = 1$; see Figure 8), the final motor command was more irregular than it was with hysteresis. This was the result of multiple switching by approximately half of the DZs.

6 Discussion

By including realistic conduction delays in afferent and efferent pathways, the model described here allowed the investigation of timing and predictive processes relevant to cerebellar involvement in the control of movement.

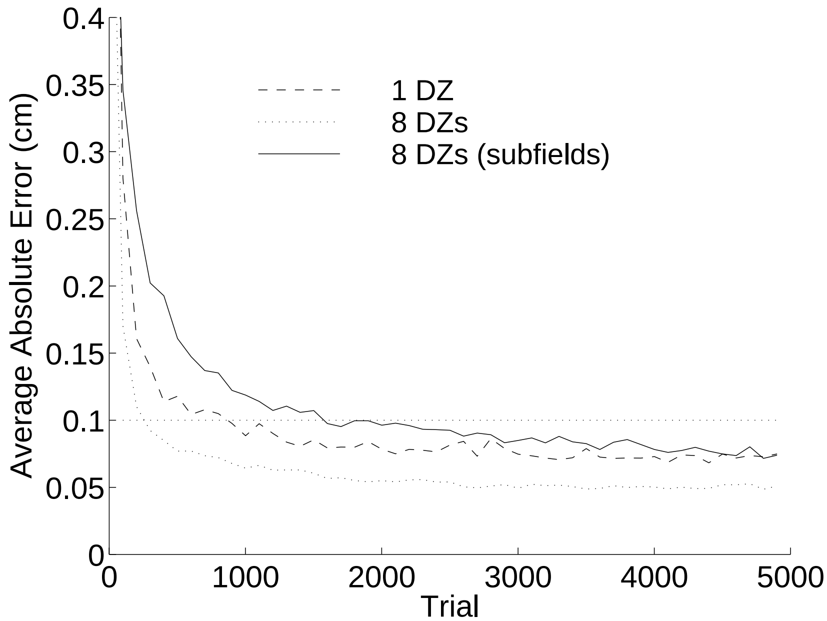


Figure 6: End point error ($|x_e - x_T|$) as a function of trial number for multi-DZ learning. In the uniform model, each DZ receives input from all of the PFs; in the subfield model, each DZ receives input from a separate subfield of 5000 PFs. Also shown is a plot for the single DZ model. The uniform model learned significantly faster than the others and reached a smaller final error. Each plotted point is an average over a bin of 50 trials of 10 learning runs. The dotted horizontal line represents the minimum threshold above which corrective movements were generated. $T_{low} = 0.8$, $T_{high} = 1$, and $\tau_2 = 100$ ms.

Moreover, the nonlinearity of the simple motor plant, which is based on muscle mechanical and spinal reflex properties, makes the control problem reflect properties of skeletomotor control better than would a simpler linear plant. While making the control problem more difficult from a conventional control perspective, the nonlinear damping has the advantage of allowing fast movements to be made with little or no oscillation, effectively solving the stability problem, at least for the one-degree-of-freedom positioning task studied here.

Key to the model's ability to perform accurate end point positioning is its ability to learn predictive control. This is illustrated most clearly in the case of a single DZ, for which clear switching curves could be derived and related to plant dynamics (see Figure 5b). The model's relative insensitivity to loop delay is due to its predictive use of a rich array of afferent and efference-copy

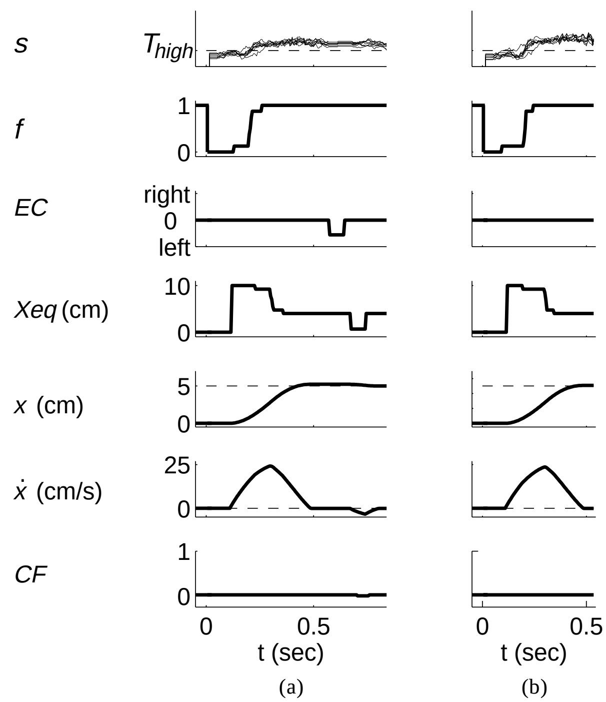


Figure 7: Multiple DZ behavior. This figure is analogous to Figure 3 but for a PC consisting of eight DZs, each receiving input from all the PFs (uniform model). The target end point, x_T, was switched from 0 to 5 cm at time 0. Shown are the time courses of the eight DZs' summed inputs, s; the PC's activation state, f; extracerebellar corrective command, EC; motor command, x_{eq} (after the 100 ms efferent delay τ_2); and the position, x, and velocity, $\dot{x}$, of the mass for a movement that started at initial position $x_0 = 0$. The bottom plot shows CF activity. $T_{low} = 0.8$ and $T_{high} = 1$. (a) Early in learning (250 trials). One of the DZs switched back to 1 too late, which caused the mass to overshoot the target slightly. The extracerebellar (EC) system generated a leftward corrective movement, which decreased CF activity below its low background level. (b) Late in learning (1500 trials). The model consistently produced accurate reaching with fast, smooth movements requiring no corrections. Note that the command is still basically a pulse-step, although it is no longer binary.

signals. The model does not explicitly predict the motor plant's behavior; that is, it does not use a forward model of the motor plant, a role suggested for the cerebellum by several researchers (Ito, 1984; Keeler, 1990; Miall, Weir, Wolpert, & Stein, 1993). In fact, the model makes no explicit predictions of any kind, if this is taken to mean the creation of representations of future events. Instead, it learns to generate motor commands in a manner that

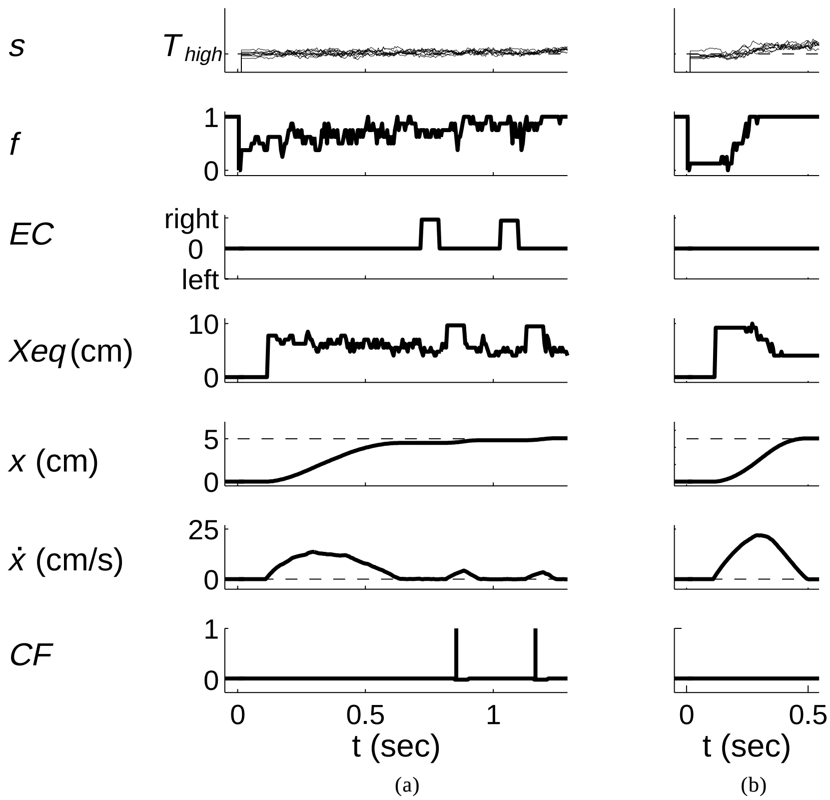


Figure 8: Multiple DZ behavior without hysteresis. This figure is analogous to Figure 7 except that there was no hysteresis ($T_{low} = 1$). (a) Early in learning (175 trials). (b) Late in learning (1,000 trials). The motor command, x_{eq}, was more irregular than it was with hysteresis. This was the result of multiple switching by approximately half of the DZs.

causes desired future behavior. The model is a kind of direct adaptive controller (e.g., Goodwin & Sin, 1984), where the term *direct* refers to the lack of a model of the controlled system.

Our model adopts the hypothesis of Marr (1969) and Albus (1971) that the granular layer provides a sparse expansive encoding that increases the ease with which a large number of associations can be formed (Buckingham & Willshaw, 1992; Tyrrell & Willshaw, 1992). We combined this hypothesis with a more realistic representation of movement-related MF signals (Van Kan et al., 1993a). Although the model's use of such a large number of PFs, and hence adjustable parameters (the PF-to-PC synaptic weights), for such a simple task is a defect from a purely engineering perspective, it is a result of our attempt to represent faithfully what is known about how information is encoded in the MF signals, coupled with our use of a random selection of MF inputs to granule units. Careful design of the latter connection pattern would decrease the number of PFs required. However, simulations show that the current model's behavior degrades if the number of PFs is significantly decreased. We did not investigate the generalization capabilities of the model, which would also be influenced by the input encoding and the number of adjustable parameters.

Experimental data on the effects of CF discharge on PF-to-PC synapses suggest an instructive role for CF signals, as adopted by the model. There now seems to be good, though not universal, agreement that CF activity, when coupled with other factors, produces long-term depression (LTD) of the action of PF-to-PC synapses (e.g., Crepel et al., 1996), as postulated by Albus (1971). Less is known about possible long-term potentiation (LTP) at these synapses, which the model also uses, although LTP has been induced in brain slices by stimulating PFs in the absence of CF activity (Sakurai, 1987), which is consistent with our model's learning rule.

An essential feature of the model's learning rule is its use of synaptically local eligibility traces for learning with delayed training information. Eligibility traces are key components of many reinforcement learning systems (e.g., Sutton & Barto, 1998) as well as models of classical conditioning (Sutton & Barto, 1981, 1990; Klopf, 1988), where they address the sensitivity of conditioning to the time interval between the conditioned and the unconditioned stimuli and the anticipatory nature of the conditioned response. Eligibility traces play the same role in this model, whose learning mechanism is much like classical conditioning, with corrective movements playing the role of unconditioned responses.[5] Our model is therefore in accord with the view that general principles of cerebellar-dependent learning may be involved in adaptation of the vestibulo-ocular reflex, classical conditioning of the eyelid response, as well as learning in saccadic eye movements and limb movements (Houk et al., 1996; Raymond, Lisberger, & Mauk, 1996). We hypothesize that for reaching, the role of the cerebellum is to eliminate corrective movements by suitably tuning the initial movement.

Only a few studies of cerebellar plasticity have attempted to manipulate the relative timing of the experimental variables used to elicit LTD. In several studies, LTD occurred only if CF stimulation preceded PF stimulation (Ekerot & Kano, 1989; Schreurs & Alkon, 1993). Recently, however, Chen and Thompson (1995) demonstrated that delaying CF activation by 250 ms after a PF volley facilitates the appearance of LTD, suggesting that there may be a cellular mechanism that compensates for the time interval. Schreurs, Oh,

[5] The present model lacks the ability to produce an analog of higher-order conditioning, one of the key features of the classical conditioning models. We know of no studies of the cerebellum's involvement in higher-order classical conditioning.

and Alkon (1996) showed that a form of LTD, which they call pairing-specific LTD, results only when PF stimulation precedes CF stimulation. Although these studies were motivated by the timing parameters required for classical conditioning of the rabbit nictitating membrane response, their results are relevant to other aspects of motor learning as well. Houk and Alford (1996) presented a model suggesting how intracellular signal transduction mechanisms that mediate LTD could give rise to an eligibility trace. Recent results in which the timing of intracellular signals was controlled photolytically appear to suggest that CF activity should precede PF activity in order to produce LTD (Lev-Ram, Jiang, Wood, Lawrence, & Tsien, 1997). However, this conclusion depends critically on the interpretation given to the various intracellular signals. We hope that the computational importance of a trace mechanism will stimulate additional cellular studies to explore this critical issue.

The nature of the training information provided by climbing fibers is incompletely understood. In oculomotor regions of the cerebellum, CFs are sensitive to retinal slip and thus are well suited to detect errors in the stabilization of visual input. By analogy, one presumes that the somatosensory sensitivity of CFs in limb regions has an analogous error-detection function, although this has been difficult to specify in detail (Fu, Mason, Flament, Coltz, & Ebner, 1997; Houk et al., 1996; Kitazawa, Kimura, & Yin, 1998; Simpson, Wylie, & de Zeeuw, 1996). In this model, we adopted our earlier working hypothesis that CFs detect hypometria by responding to corrective movements in the same direction as the primary movement (Berthier et al., 1993). This was rationalized from the finding that CFs with directional sensitivity to passive limb movements (units located in the rostral medial accessory olive) are inhibited during self-generated movements (Gellman et al., 1985) but fire when perturbations occur during or at the end of the movement (Andersson & Armstrong, 1987; Horn, Van Kan, & Gibson, 1996). We assume that corrective movements occur near the end of inaccurate movements and that they function like perturbations to fire CFs in a directionally selective manner.

Reaching movements are known to consist of a primary movement, which is often succeeded by one or more secondary movements, the latter being corrective in nature (Prablanc & Martin, 1992). Lesion studies have demonstrated the involvement of several neural pathways in the generation of both the primary movements and the corrections (Pettersson, Lundberg, Alstermark, Isa, & Tantisira, 1997). Small corrections do not require vision of the arm and are often made without subject awareness and at shorter latencies than the primary movements (Goodale, Pélisson, & Prablanc, 1986). These findings suggest the involvement of a simple, automatic mechanism such as the propriospinal network (Alstermark, Eide, Górska, Lundberg, & Pettersson, 1984; Alstermark, Górska, Pettersson, & Walkowska, 1987b). Major corrections, such as reversals in direction, engage the corticospinal system (Georgopoulos, Kalaska, Caminiti, & Massey, 1983). In this model, we assumed that all corrections following primary movements are made by a simple, extracerebellar process presumed to be mediated by the propriospinal system. This was meant to be a minimalistic assumption; the model could have made use of training information derived from more accurate corrective movements generated, for example, by the corticospinal system. In fact, since training information is derived from the proprioceptive consequences of corrective movements, the model is capable of learning from corrections generated by any system or combination of systems.

We also used the model to experiment with possible computational roles for plateau potentials in PC dendrites (Llinás & Sugimori, 1980; Ekerot & Oscarsson, 1981; Campbell et al., 1983; Andersson et al., 1984). Our representation of DZs as linear threshold elements with hysteresis allows them to produce abstract analogs of plateau potentials. Hysteresis is sometimes used in two-action control systems to reduce "chattering" caused by repeated crossing of the switching curve. It has the same effect here in making the DZs switch state less frequently, which makes the model's motor commands less erratic. Hysteresis greatly facilitated learning in the single-DZ case, presumably because it prevented chattering in motor commands, thereby making them closer to the pulse-step form and reducing the amount of learning required. In the multiple-DZ case, hysteresis had little influence on learning, perhaps because the motor commands were relatively smooth without hysteresis since they resulted from the activity of multiple DZs. We did, however, observe increased chatter in the pulse-step command when hysteresis was removed (see Figure 8), suggesting that hysteresis could have a role in facilitating the generation of well-formed motor commands. More study is needed to explore possible computational roles of nonlinear properties of PC dendrites.

Several previous cerebellar models dealing with eye movement are closely related to the model of limb control presented here. Like our model, the model of adaptive control of saccades due to Schweighofer, Arbib, and Dominey (1996a, b) follows Berthier et al. (1993) and Houk et al. (1990) in making use of corrective movements as sources of training information. Schweighofer et al. also use eligibility traces following the classical conditioning models of Sutton and Barto (1981, 1990). Unlike the monotonically decaying traces in these models, however, the eligibility traces of Schweighofer et al. reach peaks sometime after being initiated. This is in accord with Klopf's (1972) original conception that peak eligibility occurs at the optimal interstimulus interval for learning (see also Klopf, 1988). Our model also adopts this type of eligibility trace. The key differences between our model and that of Schweighofer et al. are due to differences in the dynamics of the motor plant and the degree of attention paid to system delays and afferent encoding. Because we are concerned with limb movement, our motor plant has significant inertia, which, together with nontrivial delays in various conduction channels, requires significant anticipatory control as illustrated by our simulations. The eye plant of the Schweighofer et al.

model lacks significant dynamics (the plant is essentially inertia-less), and it is not apparent that conduction delays are included. The ramp encodings we use for most of the MF signals are also more faithful representations of experimentally observed MF encodings.

Our model also shares features with the model of predictive smooth-pursuit eye movements due to Kettner et al. (1997). Like ours, this model includes MF inputs with diverse response properties and delays, a granular layer that expansively recodes this input, and a similar learning rule using eligibility traces generated by a second-order linear system. The PCs of that model, however, are continuous elements as opposed to the multistable ones used in our model (although as the number of DZs is increased, our model more closely approximates a continuous system). Additionally, training information in the Kettner model is provided by CFs that detect failures of image stabilization (retinal slip) instead of corrective eye movements.

A model of limb movement related to ours is the feedback-error learning model of Kawato and Gomi (1992, 1993) in which the cerebellum learns to act as an inverse dynamic model of the motor plant, being trained by feedback generated from movement caused by an extracerebellar system. This is similar to what we have done in the our model, with two exceptions. First, our training information is intermittent feedback from discrete corrective movements instead of a continuous feedback signal. Second, unlike feedback-error learning models, as well as the limb control model of Schweighofer (1995), we do not assume that reference trajectories specifying the complete kinematic details of the desired movement are supplied to the cerebellum by another brain region. Therefore, we do not hypothesize that the cerebellum becomes an inverse dynamic model of the plant in the sense of associating a reference trajectory to appropriate control signals. Target signals in our model do not convey this kind of detailed information about the desired trajectory. Instead, through learning, target signals become associated with movements whose kinematic details are determined by the properties of the motor plant. Our model therefore has elements in common with the equilibrium-point hypothesis (Bizzi et al., 1992; Feldman, 1966, 1974) in that muscles and spinal reflexes play essential roles in trajectory formation. Unlike that hypothesis, however, movement end points are generally not equilibrium positions.

The model presented in this article has a number of limitations. It lacks representations of many of the components of the full APG model on which it is based. In that model, movement would be the result of the combined effects of the elemental commands of a number of cerebellar APG modules that operate simultaneously. Here, we described only a single module consisting of a single PC and included no explicit representation of premotor circuits. Because the model presented here consists of a single PC controlling a single agonist actuator, it does not illustrate critical features of the full model. It does not show, for example, that during a movement, most PCs would have to increase activity to inhibit muscle synergies that should not fully participate in the movement. In the model presented here, the single PC always has to decrease activity to generate a motor command. Our model also suggests that after learning, the extracerebellar source of corrective movements no longer plays a role in limb movement. This is consistent with the feedback-error learning model but at variance with models of saccade generation in which cerebellar control augments, rather than replaces, the control provided by the brain stem burst generator (Dean, 1995; Arai et al., 1994; Optican, 1995). Our model does not adopt this approach because much less is known about the propriospinal network than is known about the brain stem pulse generator. However, this would be worthwhile to pursue in future research.

Finally, nothing in this article suggests how the model presented here might extend to more complex control problems involving multi-degree-of-freedom limbs. One of the objectives of the full APG model is to explore how the collective behavior of multiple APG modules can accomplish pulse-step control of a more complex motor plant without resorting to preplanned reference trajectories. Our research is continuing in this direction (Fagg, Sitkoff, Barto, & Houk, 1997a, b).

Acknowledgments

This work was supported by NIH 1-50 MH 48185. We thank Jay Buckingham for his contributions to an earlier version of this model and Sascha Engelbrecht for helpful comments.

References

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences, 10,* 25–61.

Alstermark, B., Eide, E., Górska, T., Lundberg, A., & Pettersson, L. G. (1984). Visually guided switching of forelimb target reaching in cats. *Acta. Physiol. Scand., 120,* 151–153.

Alstermark, B., T. Górska, A. L., Pettersson, L. G., & Walkowska, M. (1987b). Effect of different spinal cord lesions on visually guided switching of target-reaching in cats. *Neuroscience Research, 5,* 63–67.

Alstermark, B., Lundberg, A., Pinter, M. J., & Sadaki, S. (1987a). Long C3-C5 propriospinal neurones in the cat. *Brain Research, 404,* 382–388.

Andersson, G., & Armstrong, D. M. (1987). Complex spikes in Purkinje cells in the lateral vermis (b zone) of the cat cerebellum during locomotion. *Journal of Physiology, London, 385,* 107–134.

Andersson, G., Campbell, N. C., Ekerot, C. F., Hesslow, G., & Oscarsson, O. (1984). Integration of mossy fiber and climbing fiber inputs to Purkinje cells. *Experimental Brain Research (Suppl), 9,* 145–150.

Arai, K., Keller, E. L., & Edelman, J. A. (1994). Two-dimensional neural network model of the primate saccadic system. *Neural Networks, 7,* 1115–1135.

Barto, A. G., Buckingham, J. T., & Houk, J. C. (1996). A predictive switching model of cerebellar movement control. In D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo (Eds.), *Advances in neural information processing systems: Proceedings of the 1995 Conference* (pp. 138–144). Cambridge, MA: MIT Press.

Benson, M. W., Bree, G. M., Kinahan, P. E., & Hoffmann, G. W. (1987). A teachable neural network based on an unorthodox neuron. *Physica D, 22*, 233–246.

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

Bizzi, E., Hogan, N., Mussa-Ivaldi, F. A., & Gister, S. (1992). Does the nervous system use equilibrium-point control to guide single and multiple joint movements? *Behavioral and Brain Sciences, 15*, 603–613.

Buckingham, J. T., Barto, A. G., & Houk, J. C. (1995). Adaptive predictive control with a cerebellar model. In *Proceedings of the 1995 World Congress on Neural Networks* (pp. 373–380). Mahwah, NJ: Erlbaum.

Buckingham, J. T., & Willshaw, D. (1992). A note on the storage capacity of the associative net. *Network: Computation in Neural Systems, 3*(4), 404–414.

Campbell, N. C., Ekerot, C. F., Hesslow, G., & Oscarsson, O. (1983). Dendritic plateau potentials evoked in Purkinje cells by parallel fibre volleys in the cat. *Journal of Physiology (Lond), 340*, 209–223.

Chen, C., & Thompson, R. F. (1995). Temporal specificity of long–term depression in parallel fiber—Purkinje synapses in rat cerebellar slice. *Learning and Memory, 2*, 185–198.

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

Dean, P. (1995). Modelling the role of the cerebellar fastigial nuclei in producing accurate saccades: The importance of burst timing. *Neuroscience, 68*, 1059–1077.

Ekerot, C. F., & Kano, M. (1989). Stimulation parameters influencing climbing fiber induced long-term depression of parallel fiber synapses. *Neuroscience Research, 6*, 264–268.

Ekerot, C. F., & Oscarsson, O. (1981). Prolonged depolarization elicited in Purkinje cell dendrites by climbing fiber impulses in the cat. *Journal of Physiology (Lond), 318*, 207–221.

Fagg, A. H., Sitkoff, N., Barto, A. G., & Houk, J. C. (1997a). Cerebellar learning for control of a two-link arm in muscle space. In *Proceedings of the IEEE Conference on Robotics and Automation* (pp. 2638–2644).

Fagg, A. H., Sitkoff, N., Barto, A. G., & Houk, J. C. (1997b). A model of cerebellar learning for control of arm movements using muscle synergies. In *Proceedings of the IEEE Symposium on Computational Intelligence in Robotics and Automation* (pp. 6–12).

Feldman, A. (1966). Functional tuning of the nervous system with control of movement or maintenance of a steady posture. II Controllable parameters of the muscle. *Biophysics, 11*, 565–578.

Feldman, A. (1974). Change in length of the muscle as a consequence of the shift in equilibrium in the muscle-load system. *Biophysics, 19*, 544–548.

Fu, Q.-G., Mason, C. R., Flament, D., Coltz, J. D., & Ebner, T. J. (1997). Movement kinematics encoded in complex spike discharge of primate cerebellar Purkinje cells. *NeuroReport, 8*, 523–529.

Gellman, R., Gibson, A. R., & Houk, J. C. (1983). Somatosensory properties of the inferior olive of the cat. *Journal of Comparative Neurology, 215*, 228–243.

Gellman, R., Gibson, A. R., & Houk, J. C. (1985). Inferior olivary neurons in the awake cat: Detection of contact and passive body displacement. *Journal of Neurophysiology, 54*, 40–60.

Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1983). Interruption of motor cortical discharge subserving aimed arm movements. *Experimental Brain Research, 49*, 327–340.

Ghez, C. (1979). Contributions of central programs to rapid limb movement in the cat. In H. Asanuma and V. J. Wilson (Eds.), *Integration in the nervous system* (pp. 305–320). Tokyo: Igaku-Shoin.

Ghez, C., Gordon, J., Ghilardi, M. F., Christakos, C. N., & Cooper, S. E. (1990). Roles of proprioceptive input in the programming of arm trajectories. In *The Brain: Cold Spring Harbor Symp. Quant. Biol.* (Vol. 55, pp. 837–847). Cold Spring Harbor, NY: Cold Springs Harbor Laboratory Press.

Ghez, C., & Martin, J. H. (1982). The control of rapid limb movement in the cat. III. Agonist-antagonist coupling. *Experimental Brain Research, 45*, 115–125.

Gielen, C. C. A. M., & Houk, J. C. (1984). Nonlinear viscosity of human wrist. *Journal of Neurophysiology, 52*, 553–569.

Gielen, C. C. A. M., & Houk, J. C. (1987). A model of the motor servo: Incorporating nonlinear spindle receptor and muscle mechanical properties. *Biological Cybernetics, 57*, 217–231.

Goodale, M. A., Pélisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature, 320*, 748–750.

Goodwin, G. C., & Sin, K. S. (1984). *Adaptive filtering prediction and control*. Englewood Cliffs, NJ: Prentice Hall.

Gutman, A. M. (1991). Bistability of dendrites. *International Journal of Neural Systems, 1*, 291–304.

Gutman, A. M. (1994). Gelfand-Tsetlin principle of minimal afferentation and bistability of dendrites. *International Journal of Neural Systems, 5*, 83–86.

Hoffman, G. W. (1986). A neural network model based on the analogy with the immune system. *Journal of Theoretical Biology, 122*, 33–67.

Horn, K. M., Van Kan, P. L. D., & Gibson, A. R. (1996). Reduction of rostral dorsal acessory olive responses during reaching. *Journal of Neurophysiology, 76*, 4140–4151.

Houk, J. C. (1989). Cooperative control of limb movements by the motor cortex. In R. M. J. Cotterill (Ed.), *Models of brain function* (pp. 309–325). Cambridge: Cambridge University Press.

Houk, J. C., & Alford, S. (1996). Computational significance of the cellular mechanisms for synaptic plasticity in Purkinje cells. *Behavioral and Brain Sciences, 19*(3), 457.

Houk, J. C., & Barto, A. G. (1992). Distributed sensorimotor learning. In G. E. Stelmach and J. Requin (Eds.), *Tutorials in motor behavior II* (pp. 71–100). Amsterdam: Elsevier Science Publishers.

Houk, J. C., Buckingham, J. T., & Barto, A. G. (1996). Models of the cerebellum and motor learning. *Behavioral and Brain Sciences, 19,* 368–383.

Houk, J. C., Galiana, H. L., & Guitton, D. (1992). Cooperative control of gaze by the superior colliculus, brainstem and cerebellum. In G. E. Stelmach and J. Requin (Eds.), *Tutorials in motor behavior II* (pp. 443–474). Amsterdam: Elsevier Science Publishers.

Houk, J. C., Keifer, J., & Barto, A. G. (1993). Distributed motor commands in the limb premotor network. *Trends in Neuroscience, 16*, 27–33.

Houk, J. C., & Rymer, W. Z. (1981). Neural control of muscle length and tension. In V. B. Brooks (Ed.), *Handbook of physiology: Sec. 1: Vol. 2. Motor control* (pp. 247–323). Bethesda, MD: American Physiological Society.

Houk, J. C., Singh, S. P., Fisher, C., & Barto, A. G. (1990). An adaptive network inspired by the anatomy and physiology of the cerebellum. In T. Miller, R. S. Sutton, and P. J. Werbos (Eds.), *Neural networks for control* (pp. 301–348). Cambridge, MA: MIT Press.

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

Ito, M. (1984). *The cerebellum and neural control.* New York: Raven Press.

Kawato,M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics, 68,* 95–103.

Kawato, M., & Gomi, H. (1993). Feedback-error-learning model of cerebellar motor control. In N. Mano, I. Hamada, and M. R. DeLong (Eds.), *Role of the cerebellum and basal ganglia in voluntary movement* (pp. 51–61). Amsterdam: Elsevier Science.

Keeler, J. D. (1990). A dynamical system view of cerebellar function. *Physica D, 42,* 396–410.

Keller, E. L., & Robinson, D. A. (1971). Absence of stretch reflex in extraocular muscles of the monkey. *Journal of Neurophysiology, 34,* 908–919.

Kettner, R. E., Mahamud, S., Leung, H.-C., Barto, A. G., Houk, J. C., Peterson, B. W., & Sitkoff, N. (1997). Prediction of complex two-dimensional trajectories by the eye and by a cerebellar model of smooth eye movement. *Journal of Neurophysiology, 77*, 2115–2130.

Kiehn, O. (1991). Plateau potentials and active integration in the "final common pathway" for motor behavior. *Trends in Neuroscience, 14*, 68–73.

Kitazawa, S., Kimura, T., & Yin, P.-B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature, 392*, 494–497.

Klopf, A. H. (1972). *Brain function and adaptive systems—A heterostatic theory* (Tech. Rep. No. AFCRL-72-0164). Bedford, MA: Air Force Cambridge Research Laboratories.

Klopf, A. H. (1982). *The hedonistic neuron: A theory of memory, learning, and intelligence.* Washington, DC: Hemisphere.

Klopf, A. H. (1988). A neuronal model of classical conditioning. *Psychobiology, 16*, 85–125.

Lev-Ram, V., Jiang, T., Wood, J., Lawrence, D. S., & Tsien, R. Y. (1997). Synergies and coincidence requirements between NO, cGMP, and Ca2+ in the induction of cerebellar long-term depression. *Neuron, 18*, 1025–1038.

Llinás, R., & Sugimori, M. (1980). Electrophysiological properties of in vitro Purkinje cell dendrites in mammalian cerebellar slices. *Journal of Physiology (London), 305*, 197–213.

Marr, D. (1969). A theory of cerebellar cortex. *Journal of Physiology (London), 202*, 437–470.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a Smith predictor? *Journal of Motor Behavior, 25*, 203–216.

Optican, L. M. (1995). A field theory of saccade generation: Temporal-to-spatial transform in the superior colliculus. *Vision Research, 35*, 3313–3320.

Optican, L. M., & Robinson, D. A. (1980). Cerebellar-dependent adaptive control of primate saccadic system. *Journal of Neurophysiology, 44*, 1058–1076.

Pettersson, L.-G., Lundberg, A., Alstermark, B., Isa, T., & Tantisira, B. (1997). Effect of spinal cord lesions on forelimb target-reaching and on visually guided switching of target-reaching in the cat. *Neuroscience Research, 29*, 241–256.

Prablanc, C., & Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *Journal of Neurophysiology, 67*, 455–469.

Raymond, J. L., Lisberger, S. G., & Mauk, M. D. (1996). The cerebellum: A neuronal learning machine? *Science, 272*(5265), 1126–1131.

Robinson, D. A. (1975). Oculomotor control signals. In G. Lennerstrand & P. Bach-y-rita (Eds.), *Basic mechanisms of ocular motility and their clinical implications* (pp. 337–374). Oxford: Pergamon Press.

Sakurai, M. (1987). Synaptic modification of parallel fibre–Purkinje cell transmission in in vitro guinea-pig cerebellar slices. *Journal of Physiology (London), 394*, 463–480.

Schreurs, B. G., & Alkon, D. L. (1993). Rabbit cerebellar slice analysis of long-term depression and its role in classical conditioning. *Brain Research, 631*, 235–240.

Schreurs, B. G., Oh, M. M., & Alkon, D. L. (1996). Pairing-specific long-term depression of Purkinje cell excitatory postsynaptic potentials results from a classical conditioning procedure in the rabbit cerebellar slice. *Journal of Physiology, 75*, 1051–1060.

Schweighofer, N. (1995). *Computational models of the cerebellum in the adaptive control of movements*. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Schweighofer, N., Arbib, M. A., & Dominey, P. F. (1996a). A model of the cerebellum in adaptive control of saccadic gain I. The model and its biological substrate. *Biological Cybernetics, 75*, 19–28.

Schweighofer, N., Arbib, M. A., & Dominey, P. F. (1996b). A model of the cerebelllum in adaptive control of saccadic gain II. Simulation results. *Biological Cybernetics, 75*, 29–36.

Simpson, J. I., Wylie, D. R., & de Zeeuw, C. I. (1996). On climbing fiber signals and their consequence(s). *Behavioral and Brain Sciences, 19*, 384–398.

Sinkjaer, T., Wu, C. H., Barto, A. G., & Houk, J. C. (1990). Cerebellar control of endpoint position—A simulation model. In *Proceedings of the 1990 International Joint Conference on Neural Networks* (pp. II705–710). San Diego, CA.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88*, 135–170.

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel and J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 497–537). Cambridge, MA: MIT Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Tyrrell, T., & Willshaw, D. J. (1992). Cerebellar cortex: Its simulation and the relevance of Marr's theory. *Proceedings of the Royal Society of London B, 336*, 239–257.

Van Kan, P. L. E., Gibson, A. R., & Houk, J. C. (1993a). Movement-related inputs to intermediate cerebellum of the monkey. *Journal of Neurophysiology, 69*, 74–94.

Van Kan, P. L. E., Houk, J. C., & Gibson, A. R. (1993b). Output organization of intermediate cerebellum of the monkey. *Journal of Neurophysiology, 69*, 57–73.

Wang, L., & Ross, J. (1990). Synchronous neural networks of nonlinear threshold elements with hysteresis. *Proceedings of the National Academy of Science USA, 87*, 988–992.

Wu, C. H., Houk, J. C., Young, K. Y., & Miller, L. E. (1990). Nonlinear damping of limb motion. In J. M. Winters and S. L.-Y. Woo (Eds.), *Multiple muscle systems: Biomechanics and movement organization* (pp. 214–235). New York: Springer-Verlag.

Yuen, G. L., Hockberger, P. E., & Houk, J. C. (1995). Bistability in cerebellar Purkinje cell dendrites modelled with high-threshold calcium and delayed-rectifier potassium channels. *Biological Cybernetics, 73*, 375–388.

Received April 5, 1996; accepted July 2, 1998.