HHS Public Access

Author manuscript

Nat Rev Neurosci. Author manuscript; available in PMC 2025 November 24.

Published in final edited form as:

Nat Rev Neurosci. 2025 September ; 26(9): 538–553. doi:10.1038/s41583-025-00936-z.

Cerebellar circuit computations for predictive motor control

Katrina P. Nguyen, **Abigail L. Person**[†]

Department of Physiology and Biophysics, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA

Abstract

The rise of the deep neural network as the workhorse of artificial intelligence has brought increased attention to how network architectures serve specialized functions. The cerebellum, with its largely shallow, feedforward architecture, provides a curious example of such a specialized network. Within the cerebellum, tiny supernumerary granule cells project to a monolayer of giant Purkinje neurons that reweight synaptic inputs under the instructive influence of a unitary synaptic input from climbing fibres. What might this predominantly feedforward organization confer computationally? Here we review evidence for and against the hypothesis that the cerebellum learns basic associative feedforward control policies to speed up motor control and learning. We contrast and link this feedforward control framework with another prominent set of theories proposing that the cerebellum computes internal models. Ultimately, we suggest that the cerebellum may implement control through mechanisms that resemble internal models but involve model-free implicit mappings of high dimensional sensorimotor contexts to motor output.

Introduction

The cerebellum is a 3-layered structure emanating from the brainstem that, among other roles, makes movements fast and accurate. The cerebellum receives inputs from across the brain and reformats this information within its vast granule cell layer. Cerebellar granule cells then project, via parallel fibres, to the principal neurons of the cerebellar cortex, Purkinje cells (PCs), famous for their capacity to learn. The strength of the synapses that granule cells make onto PCs is under the control of instructive signals conveyed by unitary climbing fibre inputs from the inferior olive in the brainstem, which ultimately sculpt PC firing. PC activity then impinges on output neurons in the cerebellar nuclei to influence brain circuits outside the cerebellum[1].

While often considered a particularly ancient structure, the cerebellum proper **[G]** first appeared in jawed vertebrates, well after pallial (cortex), basal ganglia, brainstem and spinal circuits evolved[2]. According to current understanding, its emergence coincided evolutionarily with a massive expansion of vertebrate body size and motor repertoires. The

[†] abigail.person@cuanschutz.edu .

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

evolution of larger bodies led to two novel motor control problems that the cerebellum has been proposed to solve. First, it made sensory feedback delays longer since larger bodies increase axonal conductance times between the periphery and the brain. Without the emergence of novel computations to handle such delays, movement corrections reliant on sensory feedback would therefore be based on out-of-date information. Second, larger bodies resulted in more complex self-generated sensory feedback and interaction torques – forces that arise from the action of one body part on another --that the brain must mitigate for effective control. Interaction torques scale multiplicatively with the weight and length of the effector: thus, from a control theory **[G]** standpoint, the sharp increase in self-generated forces as bodies grew larger necessitated a neural system with the computational capacity to compensate for (through predictive computations) and interpret (as self-generated forces rather than non-self generated) these forces[3]. Interestingly, diseases of the cerebellum cause symptoms that resemble problems that can emerge in engineered control systems that do not address sensory delays and multijoint interactions. Engineered control systems that have long sensory feedback delays relative to their speed tend to oscillate around their targets: for example a room's temperature will oscillate around the thermostat set point (target) if feedback is not damped appropriately (Fig. 1, Box 1). Similarly, cerebellar damage results in an effector (the controlled body part) over- or undershooting its targets and sometimes generating pseudo-oscillatory motions at movement endpoints (the end of the movement at the target) as corrections that steer the effector back to a missed target are repeatedly engaged [4,5]. Further, multijoint robotics movements have historically actuated single joints individually, in order to simplify the dynamic interaction torque compensations that would otherwise cause control problems that resemble the symptoms of cerebellar ataxia **[G]** [6].

How might the cerebellum solve these motor control problems? Some theories propose that the cerebellum computes internal models **[G]**, in which sensory predictions are generated from motor commands to shorten feedback delays through internal simulation. Internal models serve many purposes, and could, in principle, solve both the sensory delay problem and self-vs non-self generated feedback problems described above [5,7–10]. Another class of solutions, however, sidesteps internal simulation and instead posits that the cerebellum serves a more basic function, that of feedforward anticipatory control[11–13]. This latter function is well studied in a canonical cerebellar associative learning **[G]** paradigm, delay eyelid conditioning (DEC), and it has been argued that this fundamental computation could be extrapolated to explain the cerebellum's role in all movements.

This Review explores whether the cerebellum's well understood role in associative learning can account for its contribution to motor control. We discuss the two competing, but not mutually exclusive ideas described above — that the cerebellum computes feedforward control policies, generating predictive motor commands to influence movements or, alternatively, that it computes forward internal models, generating sensory predictions that are used in diverse ways by downstream structures (Fig. 2) — and review empirical data to arbitrate between them. Note that in our use of the term feedforward, we do not mean to imply that the entire motor plan is prespecified before movement onset. Rather, we use feedforward control to mean that information about the current state of a movement — which includes both efference copy **[G]** and sensory feedback — is used as a cue to elicit

a learned command that adjusts the movement in an anticipatory fashion, in the absence of an additional computational step in which the sensorimotor information is used to construct an internal model or is compared to such a model. Our discussion focusses on multiple stages along the cerebellar circuit, piecing together a layer-by-layer algorithm that produces canonical predictive DEC and comparing and contrasting this algorithmic solution with the signals generated during movements. Finally, we end with a discussion of how this control solution interfaces conceptually with feedback control using internal models and evaluate whether, on balance, the current state of the field favors either model. The scope of this review focuses mainly on relating DEC to other discrete movements, thus we do not comprehensively include all cerebellar adaptation paradigms, including vestibular learning and control.

Feedforward control or forward modeling?

In DEC, animals learn to associate a neutral cue (the conditioned stimulus (CS), often a tone) with a second stimulus (the unconditioned stimulus (US), often a periorbital airpuff) that elicits a reflexive eyelid closure: with training they learn to close their eyes just before the second stimulus[14]. This demonstrates a very basic form of a predictive motor policy. Decades of research on the mechanisms of this form of learning have revealed that the cerebellum and brainstem motor nuclei are sufficient to learn and generate these predictive eyelid closures (although variations on this paradigm with extended temporal intervals separating CS and US do require the neocortex)[15,16]. Moreover, this paradigm has revealed how learning is implemented within the cerebellum[17]. According to the textbook view (Fig. 3), during DEC, pauses in PC activity develop that disinhibit the neurons in the cerebellar nuclei that generate the motor command to close the eyelid[18–21]. The timing of these learned pauses is thought to be governed by a reweighting the subset of parallel fibre synaptic inputs onto a PC that are active just before the air puff, with this reweighting being driven by an instructive signal from the climbing fibres that conveys air puff information and drives a complex spike **[G]** in PCs[17,22]. This results in a reduction in the excitatory parallel fibre-to-PC drive that, coupled with stable feedforward inhibition from molecular layer interneurons (MLIs), pauses PC firing around the time of the instructive complex spike[23–26]. At the core of this computation is the translation of an estimate of the elapsed time between the CS and US into a movement, in this case an eyelid closure that occurs before a reflexive blink would otherwise need to happen.

DEC reveals a form of feedforward motor control that links an early cue with a later response through associative learning. This form of control could explain cerebellar contributions to many diverse movements[11,12,17]: for each of which within-movement contextual information could be used in ways analogous to the CS in DEC. For example, the cerebellum may help fast movements land on target by using motor and sensory context to generate predictive stop signals. Coordination deficits observed in individuals with cerebellar disease, such as hypermetric saccades **[G]** [9] and reaches that decelerate too slowly to land on target, are consistent with this idea[27,28]. Furthermore, in nonhuman primate experiments in which the activity of the cerebellar nuclei was reduced via cooling, the anticipatory components of multi-part movements were eliminated, slowing the engagement of the antagonist muscles that mediate braking in reaching movements and saccades and

resulting in hypermetric endpoints[29]. This framework suggests that when a condition, i.e. a spatiotemporal context, becomes associated with a motor error, the cerebellum could use that condition to generate anticipatory control signals that achieve flexible goals faster or avoid errors.

While DEC provides an example of learned anticipatory feedforward control by the cerebellum, the alternative hypothesis that the cerebellum computes internal models used for rapid error correction is the focus of another dynamic subfield of cerebellar research[5,8,9,30]. Forward internal models, which predict sensory consequences of actions from motor efference copies, could be used to generate anticipatory corrections[5,9,31–33], but to do so they require a secondary calculation, sensory prediction, that is fully bypassed in the feedforward control framework. A second form of internal model, inverse models, in which motor commands are computed to accomplish a motor goal, also require extensive computation as they must model joint torques that vary wildly under different conditions[8,34]. By contrast, feedforward control policies simply associate the sensorimotor conditions (efference copy, sensory feedback and goal) with the adjustments that the brain had to make in the past under such conditions, and generate those adjustments preemptively as a learned anticipatory command[11,35,36].

How well do the empirical data support or challenge the feedforward motor control model of cerebellar function? The recent explosion of studies on cerebellar physiology and anatomy, and its proposed involvement in many behaviors, allows a revisiting of this question. Here, we use four core mechanistic principles of DEC to guide interpretation of data collected during other behaviors, focusing on the motor control while acknowledging that these principles can be extended to non-motor domains. The first principle is that the output nuclei of the cerebellum generate anticipatory motor commands that actuate directional control of parts of the body, pushing and pulling along ego-centric coordinates [21,37]. The second is that this actuation is caused by learned pauses in PC activity[18,20]. The third principle is that the timing of these pauses is determined by plasticity at the synapses made by temporally sparse parallel fibres onto PCs, guided by cues from instructive climbing fibres[38–40]. Finally, the fourth principle is that information entering the cerebellum is reformatted by the granule cell layer to augment the spatiotemporal state representations, including time, that cue the anticipatory responses eventually read out in the nuclei[22,40]. Below, we consider these mechanistic principles and contrast them with predictions associated with forward model hypotheses.

Cerebellar nuclei control signals

In DEC, cerebellar nuclear output neurons drive predictive eyelid closure[21]. In keeping with this theme, cerebellar nuclear neurons involved in other motor domains, including those controlling the limbs, eyes and whiskers, may also generate predictive control signals, explored below. We will also see that these signals do not account for all observations in cerebellar output signaling, with some cerebellar regions showing reduced, rather than increased, rate modulation during well learned movements, and increased activity selectively when the sensory-motor context is altered [41–51].

Studies of the oculomotor system in monkeys that have focused on the cerebellar caudal fastigial nucleus have found activity in neurons that are tuned to saccade direction that precedes the acceleration of contraversive saccades and the deceleration of ipsiversive saccades[31,45,46,52]. These activity patterns could provide a simple push-pull system for accelerating or braking eye movements via timed activation of the eye muscles [32] and could mediate the precise timing required to end a saccade on target. These observations are consistent with the hypothesis that cerebellar output channels are controllers[46,48]. Reaching movements have also received considerable attention with respect to mechanistic control by the cerebellum. In reaching mice, many neurons in the anterior interposed nucleus exhibit a brief increase in firing rates just before the limb reaches its endpoint. The magnitude of this burst causally influences the reach deceleration rate by exerting an inward pull on the limb[41,53,54]. Thus braking activity ascribed to the cerebellar nuclei is akin to slowing a car down by throwing it into reverse. Like the conditioned bursts observed in the cerebellar nuclei that drive DEC, this activity is temporally aligned to the movement. However, instead of generating an anticipatory eyelid closure, these cells may drive anticipatory deceleration, helping the limb land on target. Push modules with opposing effects on the limb likely also exist: in both mice and primates, reaching movements can be accelerated by altering cerebellar output [55,56]. These studies align with early work in primates that used electrical stimulation coupled with electromyograms and movement monitoring to identify a somatotopic map of limb muscles along the medio-lateral extent of the cerebellar nuclei [28,48,49,57–59] as well as recordings from these nuclei during center-out reaching tasks that suggested directional tuning. Thus, the cerebellar nuclei contain cell groups closely linked to the effectors of limb control, and these cells can be recruited to exert short latency adjustments to muscle output[36,60].

Other push-pull actuation modules have been identified for effectors of orofacial behaviors, including whisking[37,61–63], licking [64,65], and even eyelid opening[21]. Importantly, the principle of push-pull actuation need not be limited to the motor domain. Working memory, anxiety, navigation, fear and time-evolving cortical population activity patterns are also associated with activity in cerebellar output domains [42,65–71], suggesting that they may also be under flexible learned control by the cerebellum[72]. Together, these data suggest that segregated cerebellar output channels actuate individual body effectors, emotional effectors or neural dynamics.

In contrast to this premotor bursting activity, there is also evidence that some cerebellar nuclear regions, such as the rostral fastigial nucleus, become demodulated under previously experienced conditions, such that their firing rates change only under conditions in which there is unexpected sensory feedback (as would be expected for a sensory prediction error signal)[51]. This does not refute the feedforward control hypothesis, but suggests that other computations may exist alongside predictive control to drive cerebellar nucleus activity. These may include sensory prediction, as expected for forward models. Importantly, multiple cell types exist within the cerebellar nuclei and thus far we have little understanding of whether diverse tuning can be ascribed to output class heterogeneity[73].

On balance, while many reports are consistent with the existence of anticipatory control signals in the cerebellar nuclei, studies of nuclear activity have received relatively little

attention compared to the rest of the cerebellum. Much work therefore remains to be done to clarify whether cerebellar output represents a consistent signal, be it control, sensory prediction or otherwise.

Timed PC suppression

Cerebellar PCs are spontaneously active inhibitory neurons that emit two types of spikes – 'simple spikes'**[G]** which propagate and convey information to downstream targets, and 'complex spikes' which are an extracellular readout of dendritic spikes driven by climbing fibres. During acquisition of conditioned eyelid responses, many PCs that exhibit complex spike responses to the UC develop simple spike suppression patterns, such that simple spike firing is inversely aligned with burst firing in the cerebellar nuclei and the learned conditioned responses [18,20,21,40,74–76]. This learned linkage between the CS and the generation of a predictive eyelid closure through timed PC spike suppression is the canonical mechanistic basis of a feedforward motor control policy [11]. Many studies, primarily in monkeys, have rigorously examined PC firing patterns during a variety of other motor behaviors and, at first glance, the responses that have been observed are totally disparate from patterning observed during DEC. Below we review a number of interpretations of this encoding and then return to newer data collected during DEC that may reconcile PC patterning across paradigms.

In immediate contrast to the idea that conditioned responses result from learned pauses in PC firing, many studies of PCs have revealed increases in simple spike rates during a wide variety of movements (Fig. 4) [30,61,66,77–84]. Furthermore, this facilitation of simple spike firing is not always oppositely tuned to complex spike tuning, as might be assumed [80,83]. These facilitating PC firing patterns have been proposed to implement various computations, the most notable of which are internal models. Both flavors of internal models (forward and inverse) have been tested in PCs. Elegant experiments in which forces required to move a manipulandum were systematically varied while monkeys generated movements with identical kinematics strongly favored the forward model hypothesis, finding that PCs encode movement kinematics rather than the muscle commands required to perform the movements[30,80].

Despite these findings, individual PC simple spike patterns remain puzzling. Their response patterns are highly diverse, even among PCs with shared complex spike tuning (i.e. respond to errors of a common direction), and the activity of single neurons accounts for relatively little behavioral variance[80,85]. This raised the question of whether additional types of coding by PCs contribute to cerebellar output[80]. A breakthrough came from studies showing that PC tuning diversity becomes more interpretable when these cells are considered at the population level, with behavioral tuning strengthening as analyses incorporate larger PC ensembles[80,85]. Investigators initially inferred that PCs with shared complex spike tuning are more likely to converge anatomically within the cerebellar nuclei and therefore pooled these PCs to determine their encoding as a population. This pooling significantly improved behavioral encoding accuracy[80,85]. Notably, similar results have been observed even in studies in which complex spike tuning was not directly assessed — that is, when convergence was instead inferred through the topographic projections of PCs onto the

cerebellar nuclei [61,84]. Summarizing, groups of PCs more closely align with kinematic variables, consistent with population coding hypotheses.

Viewing PC activity through the lens of population coding **[G]** has revealed additional encoding features. For example, oppositely tuned PCs — those with firing rates that are either suppressed or facilitated during a specific movement, hereafter referred to as suppressing and facilitating PCs, respectively— can be approximately equally represented at the population level in terms of the numbers of each type[84]. If these two PC populations converge onto the same cerebellar nuclear neurons, the net inhibitory postsynaptic current that they would generate would therefore remain relatively stable, with only a divergence in the balance between the two populations leading to a net change in inhibition of the cerebellar nuclear target and control of that actuation channel (Fig. 4). This pattern is indeed observed in PC populations that control saccades, reaches, tongue and whisker movements, with net suppression of the PC population occurring at discrete moments during movements[61,84,86–88]. This net suppression could disinhibit the cerebellar nuclei, generating a control signal that is in general alignment with basic feedforward predictive control signals seen in DEC.

On the other hand, balanced PC populations could also implement a form of cancellation, a hypothesized function of internal models[5,7,10,87,89]. This interpretation suggests that facilitating and suppressing PCs may encode distinct variables that are subsequently compared, such as sensory feedback and sensory prediction or self- versus externally generated sensory feedback, respectively[89,90]. Notably, studies show that sensory feedback selectively engages MLIs to suppress PCs[61,63] whereas efference copy pathways do not, establishing a mechanistic basis for privileged access of sensory feedback to PC suppression.

Sensory reafference cancellation has also been proposed to occur at the level of the cerebellar nuclei, where inhibitory PCs may cancel excitatory mossy fibre input from brainstem nuceli[51,61]. An illuminating study on this topic found that neurons in cerebellar nuclei responded strongly to sensory-evoked whisking but not to spontaneous whisking, despite similar kinematics[61]. Recordings from PCs provided mechanistic insight into this dichotomy, revealing that the temporal coherence of PC activity varied between these conditions. Nuclear activity was preferentially evoked when PC activity was temporally locked to sensory responses; during spontaneous whisking, PC firing rate increases lacked the temporal coherence necessary to drive nuclear activation. This finding aligns with the synaptic mechanisms by which PC population activity can disinhibit the cerebellar nuclei. The rapid kinetics of PC–nuclear interactions allow nuclear neurons to respond within milliseconds to population disinhibition, with this disinhibition driven by the precise temporal alignment of PC spikes and pauses [91,92].

The previous discussion assumes that facilitating and suppressing PCs converge in the cerebellar nuclei; however, they may instead function as independent modules, each controlling independent actuation channels (Fig. 4). A recent synthesis of this idea proposed the existence of actuation channels linked to parallel ‘upbound’ and ‘downbound’ modules in the cerebellar cortex[93]. These types of modules are distinguished by the expression

or absence of Aldolase C (also known as zebrin) and, in turn, may map approximately onto facilitating and suppressing PCs. Some reports suggest that facilitating PCs belong to upbound modules, which are zebrin-positive zones that receive instructive input from specific regions of the inferior olive, while suppressing PCs correspond to downbound, zebrin-negative zones[94]. A recent study that modulated two of these channels (one upbound and one downbound) independently showed that they clearly actuate distinct types of eye movements[94]. This suggested that at least these neighboring modules did not obviously converge and collaborate[94], which calls into question the interpretations around population codes based on merged facilitating and suppressing PCs discussed above[80,84]. These questions, therefore, remain under active investigation.

Among the studies described above examining PC encoding, none fully align with the original premise that learned PC pauses that mediate DEC form a canonical PC response type. However, PC signals during DEC may also be more complex than initially characterized, challenging the assumptions with which we started. In DEC, a consensus exists that PCs for which the US drives complex spikes, develop pauses. Yet, recent work suggests that while this may be true, it is incomplete. For example, a recent preprint has reported that a sizable population of PCs show simple spike facilitation that dependent on pausing PCs disinhibiting downstream cerebellar outputs that in turn engage feedback networks to excite PCs[95]. These facilitating PCs may contribute to agonist/antagonist muscle coordination and timing[95]. In conceptually related work, recursive network connections, in which outputs feedback as inputs, have been shown to facilitate conditioned sequence learning, via a mechanisms in which learned cerebellar conditioned responses are used as CS for subsequent conditioned responses[96]. Finally, high levels of PC functional diversity have been noted during trace eyelid conditioning **[G]**, allowing the neurons to be classified into 8 functional groups with complex time-varying firing rates that go beyond simple spike suppression or facilitation. While complex spike tuning may be diverse among these groups, future studies will be required to decipher the extent to which this PC diversity enhances or expands learned behavioral control by the cerebellum [97,98]. Each of these studies emphasizes the role of recurrent neural pathways within the cerebellum as essential for the formation of such heterogeneous patterns of activity, adding computational depth to the otherwise purely feedforward view of cerebellar computations.

Taken together, do these reports support the idea that simple feedforward control principles can account for cerebellar contributions to learned movement or lead to a refutation of the hypothesis? The presence of complex time-varying signals even within a low-dimensional behavior (DEC) seems to challenge the former notion immediately. However, network models that incorporate anatomically realistic recurrent connectivity show that a kernel of learned PC spike suppression ripples through these recurrent connections to give rise to much richer PC firing patterns that nevertheless conform to feedforward control principles[99]. If recurrent connectivity facilitates cascades of PC modulation patterns to diversify PC responses, it both provides a mechanistic grounding for understanding the source of this diversity and a potential role for such heterogeneity in linking together modules of control for behaviors involving more muscle groups across time. It may also account for the emergence of PC patterns that resemble internal models, since the learning rules

like LTD applied to recurrent activity could implement control systems-like engineering solutions to biological motor control problems. Moreover, our emerging appreciation of the molecular diversity of PCs, which cluster into multiple groups based on transcriptome expression, could lead to identification of distinct computational modules that intermingle within the cerebellar cortex[94,100]. Already, studies support the idea that there are links between molecular and physiological diversity, some of which map onto the traditionally appreciated zebrin positive and negative zones[94,101]. Thus, future work incorporating connectivity, molecular and physiological diversity and computational hypotheses of cerebellar contributions to motor control, promises to yield rapid progress in resolving these debates in upcoming years.

Instructive signals for plasticity

Both feedforward control and internal models must be learned. Mechanistic hypotheses of how these models are learned assume learning is under the control of error-driven climbing fibre-driven complex spike responses. Many observed complex spike patterns are largely consistent with the predictions of both hypotheses. During DEC, climbing fibres are activated by the airpuff, and their activity is sufficient to drive associative plasticity[23,102–104], implementing a form of supervised learning **[G]**. Climbing fibres also signal sensorimotor error in a variety of motor adaptation paradigms, such as vestibulo-ocular reflex (VOR) gain and smooth pursuit adaptation, in which climbing fibres signal visual motion detected on the retina when the visual target is supposed to be stabilized by eye movement itself[105–107], consistent with a role for these signals in tuning internal models. Similarly, saccades that land too far to one side of a target will elicit complex spikes [81,108,109] that are tuned to the direction of the error but in sensory coordinates, such as the location of a visual target relative to the fovea[109], as predicted by the forward model hypothesis[109]. The PCs that receive this signal are those whose simple spike firing rates must be suppressed to pull the eyes to the required position. Thus, it is proposed that the climbing fibres provide an instructive signal that will induce plasticity that serves to reduce the firing of this population of PCs, leading to an adjustment in saccade endpoint in subsequent saccades[81]. However, because the learning of putative feedforward controllers or internal models are mechanistically analogous, these data do not distinguish between these hypotheses[12].

Despite the dominance of the error signal hypothesis, the canonical error signals that it predicts do not align with many observations of complex spikes in the literature. Interested readers are referred to excellent recent reviews focused on complex spike tuning diversity[110,111]. Below, we summarize a handful of key results, with the aim of resolving what may seem a dizzying array of observations. Briefly, we show how population coding, newly discovered circuitry, and diverse associative plasticity rules suggest that there are multiple parallel modules that accomplish cerebellar computations.

Unlike saccade-associated complex spikes, the tuning of complex spikes during limb movements has historically been unclear. The most salient complex spike signals during monkey reaches occur before the movement[112] and complex spike tuning has been shown to predict movement endpoint, rather than responding to endpoint errors[112]. Similarly, in

monkeys using a manipulandum to track a randomly moving dot, complex spikes predicted errors, rather than responding to them[113,114]. This experimental structure deprives the subject from reliable sensory predictors with which to generate an associative correction, thus the finding that these complex spikes predict errors is equivalent to saying that they lead errors, probably by perturbing simple spiking to causally influence movements. Indeed, the view that complex spikes can drive movement adjustments, either in advance of or in response to sensory events has been proposed for other behaviors, including saccades[108]. This perturbation of simple spikes by complex spikes has also been suggested to serve an important role in learning, through an instruct-then-reinforce role[115].

There are many other examples of a misalignment between the straightforward error signal theory of complex spike tuning and the empirical data. Complex spikes that occur following the airpuff in DEC shift their timing after learning, so that they occur after the CS[24], a pattern consistent with temporal difference learning **[G]** algorithms. Moreover, during a classical conditioning **[G]** paradigm in which conditioned signals cue reward rather than an aversive stimulus, complex spikes respond first to the reward, but shift to respond to the reward-predictive cue after pairing[116]. This resembles the reward-prediction error signals discovered first in midbrain dopamine neurons and is an obvious departure from the canonical sensory prediction error signal hypothesis of complex spikes[111,116]. One proposed explanation for these observations is provided by the idea that complex spike tuning shifts to encode subjective salience, which prepares action states[117,118].

Complex spikes are very amenable to population imaging because they are characterized by massive calcium signals in PC dendrites. Experiments employing wide-scale calcium imaging are therefore well positioned to address whether the tuning properties identified via electrophysiological recordings are common or are restricted to subsets of cerebellar cortical zones. Indeed, wide-scale imaging reveals both reward and sensorimotor error tuned complex spikes that are spatially distributed[119–121]. Additionally, complex spikes encoding classic directionally-tuned error signals have been reported in population activity during gain-change tasks involving mouse and monkey forelimb movements [77,78,122,123]. Interestingly, a recent preprint has reported that these signals disappear quickly with rapid behavioral adaptation[122]. These signals therefore closely track sensory prediction errors, consistent with the long-held hypotheses about cerebellar signals. However, population imaging has also revealed nuances in the coherence of complex spike signaling across task learning. Complex spikes across PC populations can be highly synchronous as a result of gap junctional coupling in the inferior olive[124]. Inhibitory feedback from the cerebellar nuclei modulates this coupling, resulting in a shift in complex spike synchrony across learning[23,118,125–127]. The relationship between complex spike synchrony and learning state suggest that inferior olivary coupling could be important for the plasticity induction roles of these signals[118,128,129]. Moreover, olivary coupling could influence complex spike tuning, with either highly shared or highly heterogeneous signaling across the population. Thus, olivary coupling along with differences in the details of behavioral tasks (such as adaptation time course and reward or error contingencies), the cerebellar cortical module or microzone **[G]** examined, and the species studied may all contribute to the challenge of identifying a single coherent coding scheme for cerebellar instructive signals.

Of course, the tuning of complex spikes is of interest mainly because these signals drive plasticity at the parallel-fibre-to-PC synapse, resulting in motor learning[38,106,107,130–134]. Coincident signaling of climbing fibres with parallel fibres leads to long-term depression (LTD) **[G]** of these synapses while parallel fibre firing alone leads to long-term potentiation (LTP) **[G]** [115,135]. This elegant, bidirectional learning rule forms the basis of powerful computations underlying the cerebellum's proposed role as an adaptive filter **[G]** [36,136]. For example, if PC simple spike rates are monitored over multiple trials in a motor learning task, it is seen that rates are lower on trials that follow a trial in which a complex spike was generated and higher on trials that follow a trial in which there was no complex spike, demonstrating a rapid form of learning [81,84,106,107,133]. Moreover, pharmacological or optogenetic manipulation of the inferior olive to drive up or down complex spike rates can incept associations or block the extinction of associations, respectively, observations that are also remarkably consistent with these plasticity rules[102–104,132,134]. However, nuances to these rules exist. For example, under some conditions, it has been shown that complex spikes that occur in doublets selectively induce parallel fibre synaptic depression, while coincident activation of singlet complex spikes and with parallel fibre activity induces potentiation[115]. An intriguing insight that could potentially explain this heterogeneity comes from the discovery that regional diversity exists in the temporal associativity rule inducing parallel fibre LTD: while coincident activation of climbing fibres and parallel fibres drove plasticity in vermal regions, it failed to induce synaptic depression in a distinct cerebellar region, the flocculus[137]. There, synaptic depression was instead reliably induced when parallel fibre activity preceded climbing fibre activity by 120 ms , which is also the predicted latency of visual feedback to the structure. A recent preprint has reported that the length of the instructive window is under learned control itself[138], responsive to naturalistic error feedback delays for different cerebellar control modules. Thus, the timing of feedback signals may be linked to the eligibility trace **[G]** for synaptic learning rules.

Another complication in understanding how complex spikes influence synaptic plasticity is that complex spikes occur constantly in PCs, averaging ~1 Hz under baseline conditions[139]. How learned synaptic weights are protected from this seemingly constant assault on their integrity is an open question. One hypothesis suggests that the baseline firing rates of climbing fibres and parallel fibres establish an equilibrium, such that synaptic weights that are stochastically depressed simply re-establish their strength through subsequent stochastic strengthening [140,141]. However, there may be additional circuit mechanisms that protect learned synaptic weights from being over-written by stochastic drift[136]. Recent work has revealed the diversity of another cell type within the cerebellar cortex that is well suited to a read-write role in memory formation and stabilization: the MLI[142–144]. Historically, these cells were known as 'basket cells' and 'stellate cells', based on the position of their cell bodies[1]. However, recent transcriptomics data, coupled with in vitro electrophysiology and spectacular high resolution anatomy, has revised this nomenclature to identify at least two interneuron types, dubbed MLI1 and MLI2[145]. MLI1s resemble the traditional basket cell but form a more specific cell class in several ways. MLI1s extend axons parasagittally and tend to make characteristic inhibitory synaptic endings around the axon initial segment of PCs, forming a structure known as a 'pinceau', which exert both synaptic and ephaptic inhibitory control of PCs[1,146,147]. MLI1s also receive input from parallel fibres and can

therefore be viewed as forming a classic feedforward inhibitory motif in the cerebellar molecular layer. Interestingly, although MLI2s also receive parallel fibre input, they do not make substantial contacts onto PCs[145]. Rather, they selectively innervate MLI1s. This organization resembles disinhibitory interneuron motifs in cerebral cortex and suggests that MLI2s could functionally disinhibit PCs. Because MLI-mediated inhibition of PCs can block associative learning[142,144] and enhanced excitation of PCs can boost associative learning, it will be important to investigate whether the diversity of MLI motifs can explain how learning is gated, protected, and expressed.

Taken together, many unanswered questions remain about the relationships between complex spikes and parallel fibre-to-PC plasticity in vivo. Temporal delays, the role of complex spike synchrony or asynchrony, MLI gating, inter-trial intervals and non-canonical tuning all contribute to the sense in the field that our understanding of complex spike-instructed parallel fibre learning rules is fragmentary. From what we do know, complex spike signaling can clearly elicit short-term reductions in simple spike activity, but unfortunately both signals are often not detected simultaneously in large population recordings. As technologies improve to monitor both complex spikes and simple spikes at scale, enabling trial-by-trial relationships to be identified, it will be interesting to see how diverse rules emerge in different regions, whether highly synchronous or asynchronous complex spike activity differentially evokes plastic changes, and how homeostatic circuit mechanisms protect or erode synaptic plasticity in the cerebellar cortex.

A GCL temporal basis set

A key feature of cerebellar-mediated classical conditioning, including DEC, is that the temporal delay between the CS onset and the US is learned. It has been hypothesized that this explicit form of temporal learning involves the cerebellar granule layer reformatting the time-invariant, neutral cue-related signals into a temporal basis set, a population coding motif in which elapsed time is represented by a cascade of neuronal activation across a population (Fig. 5) [22,40]. This population coding motif solves the problem of learning elapsed time between stimuli by facilitating the temporal overlap of late-activating GCs and the inferior olive-driven complex spike. This means that plasticity is induced only at synapses made by GCs active at the end of the CS interval, sparing the synapses made by all other active GCs.

This population motif could also mediate learned timing in movements. For instance, if the cerebellum fails to end a movement on time, it would be classified as dysmetric, a hallmark of cerebellar damage. That is, the way the cerebellum monitors elapsed time during a movement could be directly analogous to the temporal basis set hypothesized for DEC. If the GCL encoded a temporal basis set during a reaching movement, for example, then, under conditions of an overshoot, granule cells active near the endpoint would be subject to synaptic weakening as a result of complex spike feedback. This would mediate a subsequent pause in PC activity that would lead to adaptive deceleration to end the movement on target in subsequent trials[84]. This proposal is highly analogous to the mechanisms thought to underlie DEC[17]. One limitation to this hypothesis is that it is unclear whether the motor goal is represented within the temporal basis set, which would be required for the circuit

to slow down at a particular location relative to a target rather than at a particular time relative to the start of the movement. An alternative idea posits that mossy fibres that convey corollary discharge (information about motor commands) and goal information do not converge within the granule cell layer and instead create parallel channels that convey information about eye-velocity and goals (Fig. 5). These channels could then be used to compute a dynamic motor error, via a mechanism in which velocity information is integrated by MLIs to estimate eye displacement and is fed forward onto PCs that receive a static goal signal from a distinct population of granule cells. A recent study has proposed that, in this scenario, PCs will pause when displacement matches the goal, generating a stop signal[87] and showed that mossy fibres have been encode these types of information. Thus, in principle, the required components are present for this computation. This hypothesis predicts that target and sensorimotor estimates are not merged in the granule cell layer, while the former hypothesis predicts that they are. Future studies examining granule cell layer encoding will therefore aid in evaluating these predictions.

The prediction that the GCL encodes a temporal basis set has received empirical support from studies of the electric fish's cerebellar-like electrosensory lobule[148]. Below we review additional empirical tests of the idea that granule cell populations encode elapsed time or spatiotemporal state to support temporally restricted learning in mammals and discuss mechanisms by which a temporal basis set may emerge. These ideas remain in their infancy in terms of high dimensional mapping to lower-level control policies; however, in principle, the idea of temporal bases generated by the granule cell layer is highly consistent with these proposals.

David Marr and James Albus introduced the influential hypothesis that the vast population of granule cells serve a specific computational function for cerebellar learning[38,130]. Through sparsening, i.e. representing a smaller spatiotemporal signal than is present in the mossy fibers, of mossy fibre input activity, it was proposed that they perform input pattern separation, enabling better learning by PCs. This idea was extended into the temporal domain to account for the cerebellum's role in temporal learning through a related hypothesis that suggests that the granule cell layer encodes a temporal basis set, resulting in temporally sparse encoding by granule cells that facilitates temporal learning in PCs [17,22,149]. Subsequent population imaging approaches examining cerebellar granule layer activity during a variety of behaviors have often focused on population sparseness in order to evaluate these hypotheses. In 2017, a bevy of papers reported the outcomes of experiments and theories probing the Marr and Albus prediction of granule cell layer population activity sparseness[150–154]. Disagreement initially arose around the question of whether the predicted granule cell sparseness matched the in vivo data. A number of groups reported relatively dense granule cell population activity during behaviors that included DEC in mice, forelimb lever push-pull tasks in mice, and sensory responses in larval zebrafish[150–152]. Debates continue about these observations and how they may or may not align with the Marr and Albus theory of the granule cell layer. However, the current consensus is that statistical reformatting of mossy fiber information by the granule cell layer, resulting in decorrelation and temporal expansion (i.e. more granular representation of time-evolving activity), can enhance learning, even without extensive sparsening[155–157]. Additionally, when the behavior under investigation is higher dimensional, involving many

body parts simultaneously and faster temporally this appears to increase granule cell population information dimensionality[155], suggesting that evidence of low dimensional, dense activation could reflect task demands, rather than the coding principle at work.

Particularly relevant to the focus of this Review is the debate about whether the granule cell layer encodes elapsed time in its population code. Calcium imaging and serial single-cell recordings from granule cells lend support to the notion that a time-evolving population code exists, particularly during limb movements in mice and following electric-organ discharges in electric fish[148,156]. Recordings from granule cell populations in mice during tasks with varying delay intervals between operant movement completion and reward delivery showed that the granule cell population activity both tiled time, with unique subpopulations active at evolving time intervals, and flexibly matched the temporal intervals of the task[156]. On the other hand, during tasks involving sensory stimuli (such as the tones used in DEC), granule cell activity timing was either very dense and nonstationary over learning[151], or inconsistent across trials[157], neither of which is easily reconciled with the temporal basis set hypothesis. Ongoing discussions about these observations persist. As technological advances enhance the temporal precision and spatial extent of population recordings in the granule cell layer, the numerous predictions and computational models that posit roles for temporal bases in learning will become newly testable. This will include studies identifying how the temporal sharpness of granule cell layer population activity influences the timing of learnable motor control signals and how population statistics such as shared bases across similar movements, influence motor learning generalization. Relationships between features of circuit physiology and behavioral learning would provide mechanistic underpinnings of fundamental psychophysical properties of motor learning [149,158].

Beyond questions of whether the granule cell layer generates a temporal basis set are questions of how it might do so. A handful of mechanistic hypotheses have been put forward. As predicted by Marr, mossy fibres from diverse precerebellar sources converge onto single granule cells, at least in some parts of the cerebellum[157,159,160], although some regions may support less mixing[161–163]. At scale, this mixing of inputs from different sources, coupled with the subthreshold excitatory post synaptic currents (EPSCs) that they generate in granule cells, meaning that coincident inputs are required to produce spiking, will satisfy some of the theoretical requirements for sparsification[36,130,148,160,164–167]. Physiologically, the synapses of mossy fibres derived from different sources also show distinct short-term plasticity profiles: for example, both high release probability but strongly depressing primary sensory afferents and low release probability but facilitating premotor or secondary sensory afferents have been observed to synapse onto granule cells [165]. Modeling shows that this diversity in synaptic dynamics can confer temporal diversity to firing in postsynaptic granule cells, and that the emergent timing diversity follows Weber-law distributions with explanatory power for temporal learning[165,168,169]. The temporal dynamics of mossy fibre firing patterns may impart further temporal diversity onto postsynaptic granule cell populations: when mossy fibres with mixed temporal tuning to ongoing movements converge onto granule cells, the coincidence detection and thresholding rules that govern high-dimensional static representations also occur in the temporal domain, resulting in the emergence of temporal bases[148,166,167]. Granule cell excitability gradients are also observed across the dorsal-ventral axis of the layer: thus if

identical inputs were shared across that axis, timing diversity could emerge simply through excitability differences[43,170].

Unipolar brush cells (UBCs), an excitatory interneuron within the granule cell layer, are thought to amplify and extend brief input signals through feedforward excitation within the granule cell layer [148,171]. By acting as a volley to maintain excitatory drive after mossy fibers stop firing, these neurons can extend the timecourse of stimulus representations. Furthermore, recurrent inhibition through Golgi cells, inhibitory granule layer interneurons, could also establish time-varying spike thresholds in granule cells, achieving similar diversification of threshold and therefore timing heterogeneity in granule cells[22,40]. This plethora of diverse features indicates a rich mechanistic reservoir for generation of temporal bases. However, the details of these population codes will remain obscure until methods for population recording that maintain high temporal fidelity become accessible. Taken together, synaptic input properties, dynamic input patterns and intrinsic circuit features are all thought to contribute to generating granule cell population activity patterns that support pattern separation and temporal learning.

Feedforward control vs feedback control

The preceding sections were intended to critically discuss mechanistic correlates between DEC and other movements to evaluate the idea that the cerebellum's role in making movements fast, smooth and accurate could be explained by its capacity to use reliable contextual cues encoded in the granule cell layer to generate anticipatory control signals through learning in PCs. As described above, there are echoes of cerebellar DEC mechanisms in multiple behaviors. For example, explicit similarities exist between DEC and reaching movements. In both cases, temporally organized bursting activity in the cerebellar nuclei drives short latency motor control that precedes behaviorally salient timepoints (airpuff or endpoint); in both cases there is some evidence of PC spike suppression timed to disinhibit the nuclei at those salient timepoints; and in both cases timing within the spatiotemporal context (the tone or early phase movement) can be learned and adjusted to over trials. Thus, to a rough approximation, it seems that many of the core mechanistic principles of the framework of DEC hold even in a complex behavior such as unconstrained, multi-joint reach-to-grasp movements.

Nevertheless, holes remain in our empirical knowledge of feedforward control mechanisms in all paradigms, including DEC. One of the limitations of this viewpoint is that it is difficult to account for the evidence that the cerebellum is intimately involved in rapid feedback control[172,173]. This feedback control is more easily explained by internal model hypotheses since these models often posit that sensory feedback is explicitly compared to predicted feedback to generate feedback-based corrections [8, 9]. These ideas have enjoyed the greatest traction in human and non-human primate psychophysical studies, in which mechanism is unconstrained and the localization of computations is inferred from motor control deficits observed in people or animals with cerebellar diseases. From these studies it has been argued that the cerebellum is the site where predicted and actual sensory feedback are compared, using a computational variant called a Smith predictor [174]. The behavioral consequences of perturbations of the cerebellum using transcranial magnetic stimulation

appear consistent with interruption of a sensory state estimate that is used to generate accurate future motor commands[175]. Moreover, the idea that sensory prediction and sensory prediction errors are encoded by the cerebellum — either as PC codes[80,122] or as cancelled sensory reafference in the cerebellar nuclei[51,61,89] — is supported mechanistically.

Many other brain regions such as the motor cortex have also been proposed to compute internal models[176] and to compare them to sensory feedback to generate motor corrections [176–182]. The motor cortex is involved in issuing error corrective signals, with a modulatory (but not strictly dependent) role proposed for cerebellum[178,181]. More recent work examining corrective submovements in tongue and limb movements reinforces the idea that areas outside the cerebellum generate sensory predictions and compute sensory prediction error[179,183–185]. Further, the evidence that the cerebellum is required for feedback control in humans is mixed, with findings of intact, but altered, feedback control gain in people with cerebellar degeneration[33,186]. Thus, internal models, error signals, or other salient sensory prediction error cues could originate outside the cerebellum. Through pathways to and from the cerebellum, these signals could be inherited, modified, and learned from, rather than being computed de novo or in isolation by the cerebellum itself[187]. Such segregation of computations could allow the cerebellum to adapt to errors detected and corrected elsewhere in the motor system[181,188]. These studies therefore call into question whether the cerebellum is required for types of internal model calculations often ascribed to it.

One of the satisfying features of the feedforward control model is the clear cascade of representations that it proposes as information flows through the cerebellar cortical adaptive filter to generate control signals through the cerebellar outputs. Most internal model hypotheses do not yet enjoy this level of precise attribution of circuit implementation within the circuit. The neural correlates of sensory prediction and sensory cancellation have been observed[51,89], but aside from a few examples, understanding how these signals emerge from inputs is not clear. The broader question also persists, where is an internal model generated? Where is the comparison between an internal model and feedback made? And how is that comparison used? On balance, it is plausible that PC patterns that resemble internal models could emerge through associative tuning of synaptic weights, as described for the feedforward control hypothesis; however, to achieve the complexity of these models upstream of the controllers in the cerebellar nuclei, additional PC learning rules may need to exist within convergent modules. PC firing rates can correlate with upcoming movements or with movements that just past, arguing for predictive (feedforward) functions and feedback (internal-model like) control functions, respectively [113,114]. How these response features, however, drive nuclear firing, or inform downstream computations more broadly, is unknown. These remain open questions that are exciting and active avenues of current investigation.

Setting aside implementation, a resolution between these disparate views may exist. Computations that use intra-movement feedback to cue learned feedforward control may allow moment-by-moment control to guide goal-directed movements without the need for explicit internal models[189,190]. Could the advantages of internal-model based control, which incorporates useful feedback into outputs, and the computational simplicity of feedforward policies, be unified into a single framework? An elegant model of adaptive intermittent

control points to a potential way to achieve this, by using intermittent feedforward associations rather than continuous feedback control[190]. This association view is akin to a lookup table **[G]**, in which the cerebellum generates learned commands from associative memories at discrete rather than continuous intervals[190]. Such chains of associations-to-actions can be seen in sequences of associative eyeblinks[96] and could, in theory, give rise to chains of anticipatory complex actions by using feedback as the associative cue upon which the next action is evoked. Further, this low-level feedforward control policy view may be more in alignment with emerging artificial neural network implementations of motor control than the dense computational requirements of internal models, which assume high dimensional representations of vast contingencies and complex physics. This model predicts that these control chains are learned, which may present a crucial challenge to the theory, given that it is not immediately clear how this associationist view can account for evidence that dynamic motor error monitoring can be used to rapidly adjust movements without training[191] (although see[89]). Thus, it may be the case that predictive control signals implement computations that resemble the output of internal models, but do so without representational models of physics, interaction torques, sensory feedback and goals. Instead, they may use implicit associations of these high dimensional conditions with learned motor outputs that minimize errors.

Conclusion

The literature suggests a simple model by which the cerebellum can rapidly learn to implement predictive movement control by learning to use cues to generate anticipatory control signals (Fig. 6). Many behaviors could benefit from such flexible anticipatory control. For example, reaching movements can be sped up and anticipatorily ended on target by such mechanisms[27,32,41,53,55,56,59,84]. According to this model, both efference copy signals and sensory feedback can form the drivers for temporally diverse firing within the granule cell layer. This allows mapping of the associations between experience and errors, so that subsequent movements which invoke the same sensorimotor experience, drive learned control signals that mitigate error[42,84]. With the recent appreciation of anatomical feedback loops and PC response diversity, this core computation could lead to elaborate chains of responses that link muscle groups in sequences; abbreviate net disinhibition of the cerebellar nuclei; or implement error detection through comparisons between sensory-privileged channels with those yoked to efference copy pathways. Depending on the actuation channel of each cerebellar module, this basic control system could speed and simplify an individual's control during behaviors as diverse as movements[27], social interactions[68] and food intake[192]. However, this view also leaves open many unresolved questions, including whether or what the optimization function of cerebellar circuits may be; how cerebellar output feeds back to instructive signaling to establish homeostatic learning levels; how time and sleep reorganize learned responses within and outside the cerebellar cortex through consolidation, and how recurrent input from the cortex and other brain areas is processed to flexibly operate in a dynamical environment. Beyond the questions relating to the implementation of cerebellar computations, the root question of how this curious structure has facilitated the massive growth and prowess of vertebrate body plans as well as the expansion of the human prefrontal cortex, will demand further attention. Artificial neural network models

may be used to identify the comparative advantages of such network architectures, which differ so extremely from the deep network models that have proved to be so powerful. Determining whether having such a relatively shallow wide-input layer network operating alongside recurrent neural networks (RNNs) makes computations faster, learning easier or better, or confers other key advantages will be an exciting step forward in the conversation about the differences between real brains and artificial networks.

Acknowledgements

We thank members of the Person lab and our three reviewers for their thoughtful and constructive suggestions. The authors acknowledge the following grant support NS134561 to K.P.N. and NS114430 and NS131839 to A.L.P.

Glossary

Adaptive filter
A neural circuit computation that adapts its outputs to changing inputs, achieving flexible goals.

Associative learning
A learning process that links together two stimuli. This is often used in context of Pavlovian processes in which one stimulus predicts the occurrence of a second.

Cerebellar ataxia
A suite of motor discoordination profiles characteristic of cerebellar dysfunction.

Cerebellum proper
The cerebellum as distinct from cerebellar-like structures such as the dorsal cochlear nucleus or electrosensory lobule of the electric fish. The cerebellum proper is unique in that it possess climbing fiber inputs.

Classical conditioning
A Pavlovian learning paradigm in which a neutral “conditioning” stimulus is repeatedly paired with a reflex-inducing “unconditioned” stimulus. Animals learn to associate the CS with the US through the process of associative learning.

Complex spike
A burst of action potentials generated by Purkinje cells in response to climbing fiber inputs. We use the term here to include dendritic Ca events driven by climbing fiber inputs to PCs, even though these signals are mechanistically distinct from the spikelets emitted by the soma.

Control theory
A field of engineering that formalizes processes for generating a control variable behave in a desired way by using inputs in various ways. These inputs can be feedback or copies of output commands, relevant examples for biological versions of controllers.

Eligibility trace

A hypothesized molecular mechanism that renders a synapse eligible for synaptic plasticity. Eligibility traces are thought to be labile in time and can be specific to individual synapses.

Efference copy
A copy of a motor command being sent to the periphery.

Internal models
A hypothesized framework in which the brain generates world- and body-based models used for functions as diverse as motor control to social cognition.

Long-term depression (LTD)
Use-dependent weakening of synaptic strength.

Long-term potentiation (LTP)
Use-dependent strengthening of synaptic strength.

Lookup table
An engineering term depicting a discrete mapping of an input onto an output, as in a table indexed by rows and columns.

Microzone
A small cerebellar module defined by interconnected olivary climbing fiber projections to a subset of PCs, their convergent targets in the cerebellar nuclei, and their projections back to the same olivary region.

Population coding
The principle that neurons encode information as ensembles with temporally evolving dynamics rather than individually.

Saccades
Brief, nearly-ballistic eye movements from one target to another.

Simple spike
An action potential type in PCs. Simple spikes are the typical type of Na-spike that propagates down the axon. They are in contrast to complex spikes which have distinct characteristics.

Supervised learning
A machine learning perspective in which networks are trained through labeled instructive signals.

Temporal difference learning
A reinforcement learning algorithm that estimates future outcomes based on the difference between predicted and actual outcomes using intermediate cues. In the case of cerebellar learning this refers to a CS becoming instructive through learned representation in the climbing fiber pathway.

Trace eyelid conditioning

A variation on a classical conditioning paradigm in which the end of the CS is separated in time from the US. By contrast, in delay eyelid conditioning CS and US co-occur.

REFERENCES

1. Eccles JC, Ito M & Szentágothai J The Cerebellum as a Neuronal Machine. (Springer Berlin Heidelberg, Berlin, Heidelberg, 1967). doi:10.1007/978-3-662-13147-3.
2. Cisek P Evolution of behavioural control from chordates to primates. Philos Trans R Soc Lond B Biol Sci 377, 20200522 (2022). [PubMed: 34957850] This review discusses neural function through the lens of evolutionary transitions that required novel computations.
3. Shadmehr R & Mussa-Ivaldi S Biological Learning and Control. MIT Press https://mitpress.mit.edu/9780262549554/biological-learning-and-control/.This book provides an invaluable introduction to computational approaches to studying biological motor control.
4. Holmes G The Croonian Lectures On The Clinical Symptoms Of Cerebellar Disease And Their Interpretation. Lecture I. 1922. Cerebellum 6, 142–147; discussion 141 (2007). [PubMed: 17510914] This is a fascinating and thorough description of motor deficits that emerged in World War 1 soldiers who suffered gunshot wounds to the cerebellum.
5. Shadmehr R Population coding in the cerebellum: a machine learning perspective. J Neurophysiol 124, 2022–2051 (2020). [PubMed: 33112717]
6. Spong MW & Fujita M Control in robotics. The Impact of Control Technology: Overview, Success Stories, and Research Challenges (Samad T and Annaswamy A, eds.). IEEE Control Systems Society (2011).
7. Wolpert DM & Miall RC Forward Models for Physiological Motor Control. Neural Netw 9, 1265–1279 (1996). [PubMed: 12662535] This paper beautifully describes how forward models can be used to solve many motor control problems common in biological systems.
8. Wolpert DM, Miall RC & Kawato M Internal models in the cerebellum. Trends in Cognitive Sciences 2, 338–347 (1998). [PubMed: 21227230]
9. Shadmehr R, Smith MA & Krakauer JW Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci 33, 89–108 (2010). [PubMed: 20367317]
10. Cullen KE Internal models of self-motion: neural computations by the vestibular cerebellum. Trends Neurosci 46, 986–1002 (2023). [PubMed: 37739815]
11. Ohyama T, Nores WL, Murphy M & Mauk MD What the cerebellum computes. Trends Neurosci. 26, 222–227 (2003). [PubMed: 12689774] This perspective article proposes feedforward control principles underlying delay eyelid conditioning could be extended to other motor control domains.
12. Raymond JL, Lisberger SG & Mauk MD The Cerebellum: A Neuronal Learning Machine? Science 272, 1126–1131 (1996). [PubMed: 8638157]
13. Bastian AJ Learning to predict the future: the cerebellum adapts feedforward movement control. Curr Opin Neurobiol 16, 645–649 (2006). [PubMed: 17071073]
14. McCormick DA & Thompson RF Cerebellum: essential involvement in the classically conditioned eyelid response. Science 223, 296–299 (1984). [PubMed: 6701513]
15. Moyer JR, Deyo RA & Disterhoft JF Hippocampectomy disrupts trace eye-blink conditioning in rabbits. Behav Neurosci 129, 523–532 (2015). [PubMed: 26214217]
16. Kalmbach BE, Ohyama T, Kreider JC, Riusech F & Mauk MD Interactions between prefrontal cortex and cerebellum revealed by trace eyelid conditioning. Learn Mem 16, 86–95 (2009). [PubMed: 19144967]
17. Medina JF, Nores WL, Ohyama T & Mauk MD Mechanisms of cerebellar learning suggested by eyelid conditioning. Curr Opin Neurobiol 10, 717–724 (2000). [PubMed: 11240280]
18. Jirenhed D-A, Bengtsson F & Hesslow G Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. J. Neurosci 27, 2493–2502 (2007). [PubMed: 17344387]
19. Jirenhed D-A & Hesslow G Are Purkinje Cell Pauses Drivers of Classically Conditioned Blink Responses? Cerebellum 15, 526–534 (2016). [PubMed: 26400585]

20. Halverson HE, Khilkevich A & Mauk MD Relating cerebellar purkinje cell activity to the timing and amplitude of conditioned eyelid responses. J. Neurosci 35, 7813–7832 (2015). [PubMed: 25995469]
21. Ten Brinke MM et al. Dynamic modulation of activity in cerebellar nuclei neurons during pavlovian eyeblink conditioning in mice. Elife 6, (2017).
22. Mauk MD & Buonomano DV The neural basis of temporal processing. Annu Rev Neurosci 27, 307–340 (2004). [PubMed: 15217335]
23. Rasmussen A, Jirenhed D-A, Wetmore DZ & Hesslow G Changes in complex spike activity during classical conditioning. Front Neural Circuits 8, 90 (2014). [PubMed: 25140129]
24. Ohmae S & Medina JF Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. Nat Neurosci 18, 1798–1803 (2015). [PubMed: 26551541]
25. ten Brinke MM et al. Evolving Models of Pavlovian Conditioning: Cerebellar Cortical Dynamics in Awake Behaving Mice. Cell Rep 13, 1977–1988 (2015). [PubMed: 26655909]
26. Boele H-J et al. Impact of parallel fiber to Purkinje cell long-term depression is unmasked in absence of inhibitory input. Sci Adv 4, eaas9426 (2018). [PubMed: 30306129]
27. Vilis T & Hore J A comparison of disorders in saccades and in fast and accurate elbow flexions during cerebellar dysfunction. in Progress in Brain Research (eds. Freund H-J, Büttner U, Cohen B & Noth J.) vol. 64 207–215 (Elsevier, 1986). [PubMed: 3088673]
28. Flament D, Hore J & Vilis T Braking of fast and accurate elbow flexions in the monkey. J. Physiol. (Lond.) 349, 195–202 (1984). [PubMed: 6737291]
29. Vilis T & Hore J Effects of changes in mechanical state of limb on cerebellar intention tremor. J Neurophysiol 40, 1214–1224 (1977). [PubMed: 409809]
30. Pasalar S, Roitman AV, Durfee WK & Ebner TJ Force field effects on cerebellar Purkinje cell discharge with implications for internal models. Nat. Neurosci 9, 1404–1411 (2006). [PubMed: 17028585]
31. Robinson DA The use of control systems analysis in the neurophysiology of eye movements. Annu. Rev. Neurosci 4, 463–503 (1981). [PubMed: 7013640]
32. Robinson FR & Fuchs AF The role of the cerebellum in voluntary eye movements. Annu. Rev. Neurosci 24, 981–1004 (2001). [PubMed: 11520925]
33. Smith MA, Brandt J & Shadmehr R Motor disorder in Huntington's disease begins as a dysfunction in error feedback control. Nature 403, 544–549 (2000). [PubMed: 10676962]
34. Kawato M & Gomi H A computational model of four regions of the cerebellum based on feedback-error learning. Biol Cybern 68, 95–103 (1992). [PubMed: 1486143]
35. Flash T & Sejnowski TJ Computational approaches to motor control. Current Opinion in Neurobiology 11, 655–662 (2001). [PubMed: 11741014]
36. Dean P & Porrill J Adaptive-filter models of the cerebellum: computational analysis. Cerebellum 7, 567–571 (2008). [PubMed: 18972182]
37. Heiney SA, Wojaczynski GJ & Medina JF Action-based organization of a cerebellar module specialized for predictive control of multiple body parts. Neuron 109, 2981–2994.e5 (2021). [PubMed: 34534455]
38. Albus JS A theory of cerebellar function. Mathematical Biosciences 10, 25–61 (1971).A beautiful work presenting a sweeping hypothesis of cerebellar control of movements the proves a useful framework over 50 years later.
39. Safo P & Regehr WG Timing dependence of the induction of cerebellar LTD. Neuropharmacology 54, 213–218 (2008). [PubMed: 17669443]
40. Medina JF & Mauk MD Computer simulation of cerebellar information processing. Nat Neurosci 3 Suppl, 1205–1211 (2000). [PubMed: 11127839]
41. Becker MI & Person AL Cerebellar Control of Reach Kinematics for Endpoint Precision. Neuron (2019) doi:10.1016/j.neuron.2019.05.007.
42. Guo J-Z et al. Disrupting cortico-cerebellar communication impairs dexterity. Elife 10, e65906 (2021). [PubMed: 34324417]
43. Gao Z et al. Excitatory Cerebellar Nucleocortical Circuit Provides Internal Amplification during Associative Conditioning. Neuron 89, 645–657 (2016). [PubMed: 26844836]

44. Wang X, Yu S-Y, Ren Z, De Zeeuw CI & Gao Z A FN-MdV pathway and its role in cerebellar multimodular control of sensorimotor behavior. Nat Commun 11, 6050 (2020). [PubMed: 33247191]
45. Noda H et al. Saccadic eye movements evoked by microstimulation of the fastigial nucleus of macaque monkeys. J. Neurophysiol 60, 1036–1052 (1988). [PubMed: 3171655]
46. Sun Z, Junker M, Dicke PW & Thier P Individual neurons in the caudal fastigial oculomotor region convey information on both macro- and microsaccades. Eur J Neurosci 44, 2531–2542 (2016). [PubMed: 27255776]
47. Özcan OO et al. Differential Coding Strategies in Glutamatergic and GABAergic Neurons in the Medial Cerebellar Nucleus. J Neurosci 40, 159–170 (2020). [PubMed: 31694963]
48. Burton JE & Onoda N Interpositus neuron discharge in relation to a voluntary movement. Brain Res. 121, 167–172 (1977). [PubMed: 832153]
49. Fortier PA, Kalaska JF & Smith AM Cerebellar neuronal activity related to whole-arm reaching movements in the monkey. Journal of Neurophysiology 62, 198–211 (1989). [PubMed: 2754472]
50. Zhai P et al. Whisker kinematics in the cerebellum. J Physiol 602, 153–181 (2024). [PubMed: 37987552]
51. Brooks JX, Carriot J & Cullen KE Learning to expect the unexpected: rapid updating in primate cerebellum during voluntary self-motion. Nat. Neurosci 18, 1310–1317 (2015). [PubMed: 26237366] This study provides a pristine example of learning and reafference cancellation within the primate cerebellum.
52. Fuchs AF, Robinson FR & Straube A Role of the caudal fastigial nucleus in saccade generation. I. Neuronal discharge pattern. J. Neurophysiol 70, 1723–1740 (1993). [PubMed: 8294949]
53. Low AYT et al. Precision of Discrete and Rhythmic Forelimb Movements Requires a Distinct Neuronal Subpopulation in the Interposed Anterior Nucleus. Cell Rep 22, 2322–2333 (2018). [PubMed: 29490269]
54. Thanawalla AR, Chen AI & Azim E The Cerebellar Nuclei and Dexterous Limb Movements. Neuroscience 450, 168–183 (2020). [PubMed: 32652173]
55. Dacre J et al. A cerebellar-thalamocortical pathway drives behavioral context-dependent movement initiation. Neuron 109, 2326–2338.e8 (2021). [PubMed: 34146469]
56. Nashef A, Cohen O, Harel R, Israel Z & Prut Y Reversible Block of Cerebellar Outflow Reveals Cortical Circuitry for Motor Coordination. Cell Rep 27, 2608–2619.e4 (2019). [PubMed: 31141686]
57. Rispal-Padel L, Cicirata F & Pons C Cerebellar nuclear topography of simple and synergistic movements in the alert baboon (Papio papio). Exp Brain Res 47, 365–380 (1982). [PubMed: 6889975]
58. Rispal-Padel L, Cicirata F & Pons C Contribution of the dentato-thalamo-cortical system to control of motor synergy. Neurosci. Lett 22, 137–144 (1981). [PubMed: 7231805]
59. Thach WT Cerebellar output: Properties, synthesis and uses. Brain Research 40, 89–97 (1972). [PubMed: 4624494]
60. Apps R & Garwicz M Anatomical and physiological foundations of cerebellar information processing. Nat. Rev. Neurosci 6, 297–311 (2005). [PubMed: 15803161]
61. Brown ST & Raman IM Sensorimotor Integration and Amplification of Reflexive Whisking by Well-Timed Spiking in the Cerebellar Corticonuclear Circuit. Neuron 99, 564–575.e2 (2018). [PubMed: 30017394] This study discovered conditions under which cerebellar output is coupled or decoupled from cerebellar cortical signaling, providing precedence for considering how sensory and motor feedback are handled differently by the cerebellum.
62. Proville RD et al. Cerebellum involvement in cortical sensorimotor circuits for the control of voluntary movements. Nat Neurosci 17, 1233–1239 (2014). [PubMed: 25064850]
63. Chen S, Augustine GJ & Chadderton P Serial processing of kinematic signals by cerebellar circuitry during voluntary whisking. Nat Commun 8, 232 (2017). [PubMed: 28794450]
64. Gaffield MA & Christie JM Movement Rate Is Encoded and Influenced by Widespread, Coherent Activity of Cerebellar Molecular Layer Interneurons. J Neurosci 37, 4751–4765 (2017). [PubMed: 28389475]

65. Gao Z et al. A cortico-cerebellar loop for motor planning. Nature 563, 113–116 (2018). [PubMed: 30333626]

66. Chabrol FP, Blot A & Mrsic-Flogel TD Cerebellar Contribution to Preparatory Activity in Motor Neocortex. Neuron 103, 506–519.e4 (2019). [PubMed: 31201123]

67. Jung SJ et al. Novel Cerebello-Amygdala Connections Provide Missing Link Between Cerebellum and Limbic System. Front Syst Neurosci 16, 879634 (2022). [PubMed: 35645738]

68. Stoodley CJ & Tsai PT Adaptive Prediction for Social Contexts: The Cerebellar Contribution to Typical and Atypical Social Behaviors. Annu Rev Neurosci 44, 475–493 (2021). [PubMed: 34236892]

69. Fallahnezhad M et al. Cerebellar control of a unitary head direction sense. Proc Natl Acad Sci U S A 120, e2214539120 (2023). [PubMed: 36812198]

70. Chen CH et al. A Purkinje cell to parabrachial nucleus pathway enables broad cerebellar influence over the forebrain. Nat Neurosci 26, 1929–1941 (2023). [PubMed: 37919612]

71. Frontera JL et al. Bidirectional control of fear memories by cerebellar neurons projecting to the ventrolateral periaqueductal grey. Nat Commun 11, 5207 (2020). [PubMed: 33060630]

72. Schmahmann JD Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. J Neuropsychiatry Clin Neurosci 16, 367–378 (2004). [PubMed: 15377747]

73. Uusisaari M, Obata K & Knöpfel T Morphological and electrophysiological properties of GABAergic and non-GABAergic cells in the deep cerebellar nuclei. J. Neurophysiol 97, 901–911 (2007). [PubMed: 17093116]

74. Heiney SA, Kim J, Augustine GJ & Medina JF Precise control of movement kinematics by optogenetic inhibition of Purkinje cell activity. J. Neurosci 34, 2321–2330 (2014). [PubMed: 24501371]

75. Garcia KS & Mauk MD Pharmacological analysis of cerebellar contributions to the timing and expression of conditioned eyelid responses. Neuropharmacology 37, 471–480 (1998). [PubMed: 9704988]

76. Attwell PJE, Ivarsson M, Millar L & Yeo CH Cerebellar mechanisms in eyeblink conditioning. Ann N Y Acad Sci 978, 79–92 (2002). [PubMed: 12582043]

77. Ebner TJ A role for the cerebellum in the control of limb movement velocity. Curr Opin Neurobiol 8, 762–769 (1998). [PubMed: 9914240]

78. Fu QG, Suarez JI & Ebner TJ Neuronal specification of direction and distance during reaching movements in the superior precentral premotor area and primary motor cortex of monkeys. J. Neurophysiol 70, 2097–2116 (1993). [PubMed: 8294972]

79. Hewitt AL, Popa LS, Pasalar S, Hendrix CM & Ebner TJ Representation of limb kinematics in Purkinje cell simple spike discharge is conserved across multiple tasks. Journal of Neurophysiology 106, 2232–2247 (2011). [PubMed: 21795616]

80. Herzfeld DJ, Kojima Y, Soetedjo R & Shadmehr R Encoding of action by the Purkinje cells of the cerebellum. Nature 526, 439–442 (2015). [PubMed: 26469054] This study invigorated the field of cerebellar cortical coding by identifying a population coding principle in which complex-spike tuning defined populations of Purkinje neurons that, despite heterogeneous individual tuning, as a group were exquisitely tuned to behavior.

81. Herzfeld DJ, Kojima Y, Soetedjo R & Shadmehr R Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. Nat. Neurosci 21, 736–743 (2018). [PubMed: 29662213]

82. Zobeiri OA & Cullen KE Cerebellar Purkinje cells in male macaques combine sensory and motor information to predict the sensory consequences of active self-motion. Nat Commun 15, 4003 (2024). [PubMed: 38734715]

83. Raghavan RT & Lisberger SG Responses of Purkinje cells in the oculomotor vermis of monkeys during smooth pursuit eye movements and saccades: comparison with floccular complex. J Neurophysiol 118, 986–1001 (2017). [PubMed: 28515286]

84. Calame DJ, Becker MI & Person AL Cerebellar associative learning underlies skilled reach adaptation. Nat Neurosci 26, 1068–1079 (2023). [PubMed: 37248339]

85. Thier P, Dicke PW, Haas R & Barash S Encoding of movement time by populations of cerebellar Purkinje cells. Nature 405, 72–76 (2000). [PubMed: 10811220]
86. Sedaghat-Nejad E, Pi JS, Hage P, Fakharian MA & Shadmehr R Synchronous spiking of cerebellar Purkinje cells during control of movements. Proc Natl Acad Sci U S A 119, e2118954119 (2022). [PubMed: 35349338]
87. Fakharian MA, Shoup AM, Hage P, Elseweifi HY & Shadmehr R A vector calculus for neural computation in the cerebellum. bioRxiv 2024.11.14.623565 (2024) doi:10.1101/2024.11.14.623565.This study discovered an organizational logic within cerebellar cortical population coding that implements output null and output potent signaling, suggesting that not all cerebellar activity exerts control.
88. Hage P et al. Control of tongue movements by the Purkinje cells of the cerebellum. bioRxiv 2024.07.25.604757 (2025) doi:10.1101/2024.07.25.604757.
89. Sawtell NB Neural Mechanisms for Predicting the Sensory Consequences of Behavior: Insights from Electrosensory Systems. Annual Review of Physiology 79, 381–399 (2017).
90. Wang X, Novello M, Gao Z, Ruigrok TJH & De Zeeuw CI Input and output organization of the mesodiencephalic junction for cerebro-cerebellar communication. J Neurosci Res 100, 620–637 (2022). [PubMed: 34850425]
91. Person AL & Raman IM Purkinje neuron synchrony elicits time-locked spiking in the cerebellar nuclei. Nature 481, 502–505 (2011). [PubMed: 22198670]
92. Witter L, Canto CB, Hoogland TM, de Gruijl JR & De Zeeuw CI Strength and timing of motor responses mediated by rebound firing in the cerebellar nuclei after Purkinje cell activation. Front Neural Circuits 7, 133 (2013). [PubMed: 23970855]
93. De Zeeuw CI Bidirectional learning in upbound and downbound microzones of the cerebellum. Nat Rev Neurosci 22, 92–110 (2021). [PubMed: 33203932]
94. Blot FGC et al. Purkinje cell microzones mediate distinct kinematics of a single movement. Nat Commun 14, 4358 (2023). [PubMed: 37468512]
95. Ohmae S, Ohmae K, Heiney S, Subramanian D & Medina J A recurrent circuit links antagonistic cerebellar modules during associative motor learning. 2021.11.16.468438 Preprint at 10.1101/2021.11.16.468438 (2021).
96. Khilkevich A, Zambrano J, Richards M-M & Mauk MD Cerebellar implementation of movement sequences through feedback. Elife 7, (2018).
97. Geminiani A et al. Mesoscale simulations predict the role of synergistic cerebellar plasticity during classical eyeblink conditioning. PLoS Comput Biol 20, e1011277 (2024). [PubMed: 38574161]
98. De Zeeuw CI, Koppen J, Bregman GG, Runge M & Narain D Heterogeneous encoding of temporal stimuli in the cerebellar cortex. Nat Commun 14, 7581 (2023). [PubMed: 37989740]
99. Raymond JL & Medina JF Computational Principles of Supervised Learning in the Cerebellum. Annu Rev Neurosci 41, 233–253 (2018). [PubMed: 29986160]
100. Kozareva V et al. A transcriptomic atlas of mouse cerebellar cortex comprehensively defines cell types. Nature 598, 214–219 (2021). [PubMed: 34616064]
101. Tang T et al. Heterogeneity of Purkinje cell simple spike–complex spike interactions: zebrin- and non-zebrin-related variations. J Physiol 595, 5341–5357 (2017). [PubMed: 28516455]
102. Mauk MD, Steinmetz JE & Thompson RF Classical conditioning using stimulation of the inferior olive as the unconditioned stimulus. Proc Natl Acad Sci U S A 83, 5349–5353 (1986). [PubMed: 3460097]
103. Steinmetz JE, Lavond DG & Thompson RF Classical conditioning in rabbits using pontine nucleus stimulation as a conditioned stimulus and inferior olive stimulation as an unconditioned stimulus. Synapse 3, 225–233 (1989). [PubMed: 2718098]
104. Silva NT, Ramírez-Buriticá J, Pritchett DL & Carey MR Climbing fibers provide essential instructive signals for associative learning. Nat Neurosci 27, 940–951 (2024). [PubMed: 38565684]
105. Ito M Error detection and representation in the olivo-cerebellar system. Front Neural Circuits 7, 1 (2013). [PubMed: 23440175]
106. Yang Y & Lisberger SG Purkinje-cell plasticity and cerebellar motor learning are graded by complex-spike duration. Nature 510, 529–532 (2014). [PubMed: 24814344]

107. Medina JF & Lisberger SG Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. Nat. Neurosci 11, 1185–1192 (2008). [PubMed: 18806784]
108. Muller SZ et al. Complex spikes perturb movements and reveal the sensorimotor map of Purkinje cells. Curr Biol S0960-9822(23)01311-8 (2023) doi:10.1016/j.cub.2023.09.062.
109. Pi JS et al. The olivary input to the cerebellum dissociates sensory events from movement plans. Proc Natl Acad Sci U S A 121, e2318849121 (2024). [PubMed: 38630714]
110. Popa LS, Streng ML, Hewitt AL & Ebner TJ The Errors of Our Ways: Understanding Error Representations in Cerebellar-Dependent Motor Learning. Cerebellum 15, 93–103 (2016). [PubMed: 26112422]
111. Hull C Prediction signals in the cerebellum: beyond supervised motor learning. Elife 9, e54073 (2020). [PubMed: 32223891]
112. Kitazawa S, Kimura T & Yin PB Cerebellar complex spikes encode both destinations and errors in arm movements. Nature 392, 494–497 (1998). [PubMed: 9548253]
113. Streng ML, Popa LS & Ebner TJ Modulation of sensory prediction error in Purkinje cells during visual feedback manipulations. Nat Commun 9, 1099 (2018). [PubMed: 29545572]
114. Popa LS, Streng ML & Ebner TJ Purkinje Cell Representations of Behavior: Diary of a Busy Neuron. Neuroscientist 25, 241–257 (2019). [PubMed: 29985093] This is a fantastic review of the complex field of Purkinje neuron tuning.
115. Bouvier G et al. Cerebellar learning using perturbations. Elife 7, (2018).
116. Heffley W et al. Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. Nat. Neurosci 21, 1431–1441 (2018). [PubMed: 30224805]
117. Bina L, Romano V, Hoogland TM, Bosman LWJ & De Zeeuw CI Purkinje cells translate subjective salience into readiness to act and choice performance. Cell Rep 37, 110116 (2021). [PubMed: 34910904]
118. Wagner MJ et al. A neural circuit state change underlying skilled movements. Cell 184, 3731–3747.e21 (2021). [PubMed: 34214470]
119. Kostadinov D, Beau M, Pozo MB & Häusser M Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. Nat. Neurosci 22, 950–962 (2019). [PubMed: 31036947]
120. Imamizu H & Kawato M Neural correlates of predictive and postdictive switching mechanisms for internal models. J. Neurosci 28, 10751–10765 (2008). [PubMed: 18923050]
121. Streng ML et al. Mesoscale Ca2+ imaging reveals networks of Purkinje cell dendritic and somatic modulation, with divergent roles of activity versus correlation during behavior. 2023.10.05.561090 Preprint at 10.1101/2023.10.05.561090 (2023).
122. Nguyen V & Stell BM Rapid Motor Adaptation via Population-level Modulation of Cerebellar Error Signals. 2024.01.03.574031 Preprint at 10.1101/2024.01.03.574031 (2024).
123. Ojakangas CL & Ebner TJ Purkinje cell complex spike activity during voluntary motor learning: relationship to kinematics. J Neurophysiol 72, 2617–2630 (1994). [PubMed: 7897479]
124. Mann-Metzer P & Yarom Y Electrotonic coupling interacts with intrinsic properties to generate synchronized activity in cerebellar networks of inhibitory interneurons. J Neurosci 19, 3298–3306 (1999). [PubMed: 10212289]
125. Lefler Y, Yarom Y & Uusisaari MY Cerebellar inhibitory input to the inferior olive decreases electrical coupling and blocks subthreshold oscillations. Neuron 81, 1389–1400 (2014). [PubMed: 24656256]
126. Kim OA, Ohmae S & Medina JF A cerebello-olivary signal for negative prediction error is sufficient to cause extinction of associative motor learning. Nat Neurosci 23, 1550–1554 (2020). [PubMed: 33169031]
127. Hoang H et al. Electrical coupling controls dimensionality and chaotic firing of inferior olive neurons. PLoS Comput Biol 16, e1008075 (2020). [PubMed: 32730255]
128. Hoang H et al. Dynamic organization of cerebellar climbing fiber response and synchrony in multiple functional components reduces dimensions for reinforcement learning. Elife 12, e86340 (2023). [PubMed: 37712651]

129. Van Der Giessen RS et al. Role of olivary electrical coupling in cerebellar motor learning. Neuron 58, 599–612 (2008). [PubMed: 18498740]
130. Marr D A theory of cerebellar cortex. The Journal of Physiology 202, 437–470 (1969). [PubMed: 5784296]
131. Ito M Cerebellar control of the vestibulo-ocular reflex--around the flocculus hypothesis. Annu Rev Neurosci 5, 275–296 (1982). [PubMed: 6803651]
132. Nguyen-Vu TDB et al. Cerebellar Purkinje cell activity drives motor learning. Nat. Neurosci 16, 1734–1736 (2013). [PubMed: 24162651]
133. Kimpo RR, Rinaldi JM, Kim CK, Payne HL & Raymond JL Gating of neural error signals during motor learning. Elife 3, e02076 (2014). [PubMed: 24755290]
134. Rowan MJM et al. Graded Control of Climbing-Fiber-Mediated Plasticity and Learning by Inhibition in the Cerebellum. Neuron 99, 999–1015.e6 (2018). [PubMed: 30122378]
135. Jörntell H & Hansel C Synaptic memories upside down: bidirectional plasticity at cerebellar parallel fiber-Purkinje cell synapses. Neuron 52, 227–238 (2006). [PubMed: 17046686]
136. Sejnowski TJ Storing covariance with nonlinearly interacting neurons. J Math Biol 4, 303–321 (1977). [PubMed: 925522]
137. Suvrathan A, Payne HL & Raymond JL Timing Rules for Synaptic Plasticity Matched to Behavioral Function. Neuron 92, 959–967 (2016). [PubMed: 27839999]
138. Jayabal S et al. Experience Alters the Timing Rules Governing Synaptic Plasticity and Learning. 2022.11.28.518128 Preprint at 10.1101/2022.11.28.518128 (2024).
139. Lang EJ, Sugihara I, Welsh JP & Llinás R Patterns of Spontaneous Purkinje Cell Complex Spike Activity in the Awake Rat. J. Neurosci 19, 2728–2739 (1999). [PubMed: 10087085]
140. Medina JF, Nores WL & Mauk MD Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. Nature 416, 330–333 (2002). [PubMed: 11907580]
141. Mauk MD & Donegan NH A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. Learn Mem 4, 130–158 (1997). [PubMed: 10456059]
142. Zhang K et al. Molecular layer disinhibition unlocks climbing-fiber-instructed motor learning in the cerebellum. bioRxiv 2023.08.04.552059 (2023) doi:10.1101/2023.08.04.552059.
143. Bonnan A, Zhang K, Gaffield MA & Christie JM Expression of a Form of Cerebellar Motor Memory Requires Learned Alterations to the Activity of Inhibitory Molecular Layer Interneurons. J Neurosci 43, 601–612 (2023). [PubMed: 36639897]
144. Gaffield MA, Rowan MJM, Amat SB, Hirai H & Christie JM Inhibition gates supralinear Ca2+ signaling in Purkinje cell dendrites during practiced movements. Elife 7, e36246 (2018). [PubMed: 30117806]
145. Lackey EP et al. Specialized connectivity of molecular layer interneuron subtypes leads to disinhibition and synchronous inhibition of cerebellar Purkinje cells. Neuron 112, 2333–2348.e6 (2024). [PubMed: 38692278]
146. Blot A & Barbour B Ultra-rapid axon-axon ephaptic inhibition of cerebellar Purkinje cells by the pinceau. Nat Neurosci 17, 289–295 (2014). [PubMed: 24413696]
147. Mittmann W, Koch U & Häusser M Feed-forward inhibition shapes the spike output of cerebellar Purkinje cells. J Physiol 563, 369–378 (2005). [PubMed: 15613376]
148. Kennedy A et al. A temporal basis for predicting the sensory consequences of motor commands in an electric fish. Nat Neurosci 17, 416–422 (2014). [PubMed: 24531306]
149. Narain D, Remington ED, Zeeuw CID & Jazayeri M A cerebellar mechanism for learning prior distributions of time intervals. Nat Commun 9, 469 (2018). [PubMed: 29391392]
150. Wagner MJ, Kim TH, Savall J, Schnitzer MJ & Luo L Cerebellar granule cells encode the expectation of reward. Nature 544, 96–100 (2017). [PubMed: 28321129]
151. Giovannucci A et al. Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. Nat Neurosci 20, 727–734 (2017). [PubMed: 28319608]
152. Knogler LD, Markov DA, Dragomir EI, Štih V & Portugues R Sensorimotor Representations in Cerebellar Granule Cells in Larval Zebrafish Are Dense, Spatially Organized, and Non-temporally Patterned. Curr Biol 27, 1288–1302 (2017). [PubMed: 28434864]

153. Cayco-Gajic NA, Clopath C & Silver RA Sparse synaptic connectivity is required for decorrelation and pattern separation in feedforward networks. Nat Commun 8, 1116 (2017). [PubMed: 29061964]

154. Litwin-Kumar A, Harris KD, Axel R, Sompolinsky H & Abbott LF Optimal Degrees of Synaptic Connectivity. Neuron 93, 1153–1164.e7 (2017). [PubMed: 28215558]

155. Lanore F, Cayco-Gajic NA, Gurnani H, Coyle D & Silver RA Cerebellar granule cell axons support high-dimensional representations. Nat Neurosci 24, 1142–1150 (2021). [PubMed: 34168340]

156. Garcia-Garcia MG et al. A cerebellar granule cell-climbing fiber computation to learn to track long time intervals. Neuron 112, 2749–2764.e7 (2024). [PubMed: 38870929]

157. Fleming EA, Field GD, Tadross MR & Hull C Local synaptic inhibition mediates cerebellar granule cell pattern separation and enables learned sensorimotor associations. Nat Neurosci 27, 689–701 (2024). [PubMed: 38321293]

158. Thoroughman KA & Shadmehr R Learning of action through adaptive combination of motor primitives. Nature 407, 742–747 (2000). [PubMed: 11048720]

159. Huang C-C et al. Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. Elife 2, e00400 (2013). [PubMed: 23467508]

160. Nguyen TM et al. Structured cerebellar connectivity supports resilient pattern separation. Nature 613, 543–549 (2023). [PubMed: 36418404]

161. Bengtsson F & Jörntell H Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. Proc. Natl. Acad. Sci. U.S.A 106, 2389–2394 (2009). [PubMed: 19164536]

162. Jörntell H & Ekerot C-F Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. J Neurosci 26, 11786–11797 (2006). [PubMed: 17093099]

163. Rancz EA et al. High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. Nature 450, 1245–1248 (2007). [PubMed: 18097412]

164. Billings G, Piasini E, Lőrincz A, Nusser Z & Silver RA Network structure within the cerebellar input layer enables lossless sparse encoding. Neuron 83, 960–974 (2014). [PubMed: 25123311]

165. Chabrol FP, Arenz A, Wiechert MT, Margrie TW & DiGregorio DA Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. Nat Neurosci 18, 718–727 (2015). [PubMed: 25821914]

166. Fujita M Adaptive filter model of the cerebellum. Biol Cybern 45, 195–206 (1982). [PubMed: 7171642]

167. Gilmer JI et al. An emergent temporal basis set robustly supports cerebellar time-series learning. J Neurophysiol 129, 159–176 (2023). [PubMed: 36416445]

168. Buonomano DV & Maass W State-dependent computations: spatiotemporal processing in cortical networks. Nat Rev Neurosci 10, 113–125 (2009). [PubMed: 19145235]

169. Barri A, Wiechert MT, Jazayeri M & DiGregorio DA Synaptic basis of a sub-second representation of time in a neural circuit model. Nat Commun 13, 7902 (2022). [PubMed: 36550115]

170. Straub I et al. Gradients in the mammalian cerebellar cortex enable Fourier-like transformation and improve storing capacity. Elife 9, e51771 (2020). [PubMed: 32022688]

171. Huson V & Regehr WG Realistic mossy fiber input patterns to unipolar brush cells evoke a continuum of temporal responses comprised of components mediated by different glutamate receptors. Elife 13, RP102618 (2025). [PubMed: 39819796]

172. Herzfeld DJ et al. Contributions of the cerebellum and the motor cortex to acquisition and retention of motor memories. NeuroImage 98, 147–158 (2014). [PubMed: 24816533]

173. Xu-Wilson M, Chen-Harris H, Zee DS & Shadmehr R Cerebellar Contributions to Adaptive Control of Saccades in Humans. J. Neurosci 29, 12930–12939 (2009). [PubMed: 19828807]

174. Miall RC, Weir DJ, Wolpert DM & Stein JF Is the cerebellum a smith predictor? J Mot Behav 25, 203–216 (1993). [PubMed: 12581990]

175. Miall RC, Christensen LOD, Cain O & Stanley J Disruption of state estimation in the human lateral cerebellum. PLoS Biol 5, e316 (2007). [PubMed: 18044990]

176. Pisella L et al. An ‘automatic pilot’ for the hand in human posterior parietal cortex: toward reinterpreting optic ataxia. Nat Neurosci 3, 729–736 (2000). [PubMed: 10862707]
177. Desmurget M et al. Role of the posterior parietal cortex in updating reaching movements to a visual target. Nat Neurosci 2, 563–567 (1999). [PubMed: 10448222]
178. Vilis T, Hore J, Meyer-Lohmann J & Brooks VB Dual nature of the precentral responses to limb perturbations revealed by cerebellar cooling. Brain Res 117, 336–340 (1976). [PubMed: 825192]
179. Bollu T et al. Cortex-dependent corrections as the tongue reaches for and misses targets. Nature 594, 82–87 (2021). [PubMed: 34012117]
180. Bollu T et al. Motor cortical inactivation impairs corrective submovements in mice performing a hold-still center-out reach task. J Neurophysiol 132, 829–848 (2024). [PubMed: 39081209]
181. Pruszynski JA et al. Primary motor cortex underlies multi-joint integration for fast feedback control. Nature 478, 387–390 (2011). [PubMed: 21964335]
182. Desmurget M & Grafton S Forward modeling allows feedback control for fast reaching movements. Trends Cogn Sci 4, 423–431 (2000). [PubMed: 11058820]
183. Schneider DM, Sundararajan J & Mooney R A cortical filter that learns to suppress the acoustic consequences of movement. Nature 561, 391–395 (2018). [PubMed: 30209396]
184. Holey BE & Schneider DM Sensation and expectation are embedded in mouse motor cortical activity. Cell Rep 43, 114396 (2024). [PubMed: 38923464]
185. Mathis MW, Mathis A & Uchida N Somatosensory Cortex Plays an Essential Role in Forelimb Motor Adaptation in Mice. Neuron 93, 1493–1503.e6 (2017). [PubMed: 28334611]
186. Zimmet AM, Cao D, Bastian AJ & Cowan NJ Cerebellar patients have intact feedback control that can be leveraged to improve reaching. Elife 9, e53246 (2020). [PubMed: 33025903]
187. Wagner MJ et al. Shared Cortex-Cerebellum Dynamics in the Execution and Learning of a Motor Task. Cell 177, 669–682.e24 (2019). [PubMed: 30929904]
188. Diedrichsen J, Hashambhoy Y, Rane T & Shadmehr R Neural Correlates of Reach Errors. J. Neurosci 25, 9919–9931 (2005). [PubMed: 16251440]
189. Todorov E & Jordan MI Optimal feedback control as a theory of motor coordination. Nat Neurosci 5, 1226–1235 (2002). [PubMed: 12404008]
190. Sakaguchi Y, Tanaka M & Inoue Y Adaptive intermittent control: A computational model explaining motor intermittency observed in human behavior. Neural Networks 67, 92–109 (2015). [PubMed: 25897510]
191. Soetedjo R & Horwitz GD Closed-Loop Optogenetic Perturbation of Macaque Oculomotor Cerebellum: Evidence for an Internal Saccade Model. J Neurosci 44, e1317232023 (2024). [PubMed: 38182420]
192. Low AYT et al. Reverse-translational identification of a cerebellar satiation network. Nature 600, 269–273 (2021). [PubMed: 34789878]
193. Mars Exploration Rovers: Spirit and Opportunity - NASA Science. https://science.nasa.gov/mission/mars-exploration-rovers-spirit-and-opportunity/ (2024).
194. Azim E & Alstermark B Skilled forelimb movements and internal copy motor circuits. Current Opinion in Neurobiology 33, 16–24 (2015). [PubMed: 25588912]

BOX 1:

Sensory delays in controlled systems

An extreme, if whimsical, example of sensory feedback delays in engineered systems is seen in the Mars rover missions of Spirit and Opportunity. With one-way communication delays of 7-20 minutes and limited autonomous driving capabilities, the rovers had to be driven extremely slowly by Earth-based engineers to avoid obstacles, and after a decade of operation had traversed only[193]..

Although orders of magnitude faster, the brain faces a conceptually similar problem: sensory feedback from the periphery indicating that targets have been reached or obstacles are present occurs after motor commands are sent. This sensory feedback can face delays of between 50 to 120 ms across species as diverse as mice and monkeys, depending on whether the feedback comes from fast proprioceptors, somatosensation or vision[194]; similarly, any corrective response is subject to return delays of processing. Like all control systems that rely on sensory feedback for control, the brain's performance will suffer from these delays: even if sensory feedback is fast, the information received will be out-of-date relative to output control signals, thus any response to feedback interpreted as an error will only take place after the error has occurred. Evidence that the cerebellum is involved in the brain's solution to this control systems problem comes from the observation that movement endpoints become pseudo-oscillatory in the face of cerebellar damage[4,5,11].

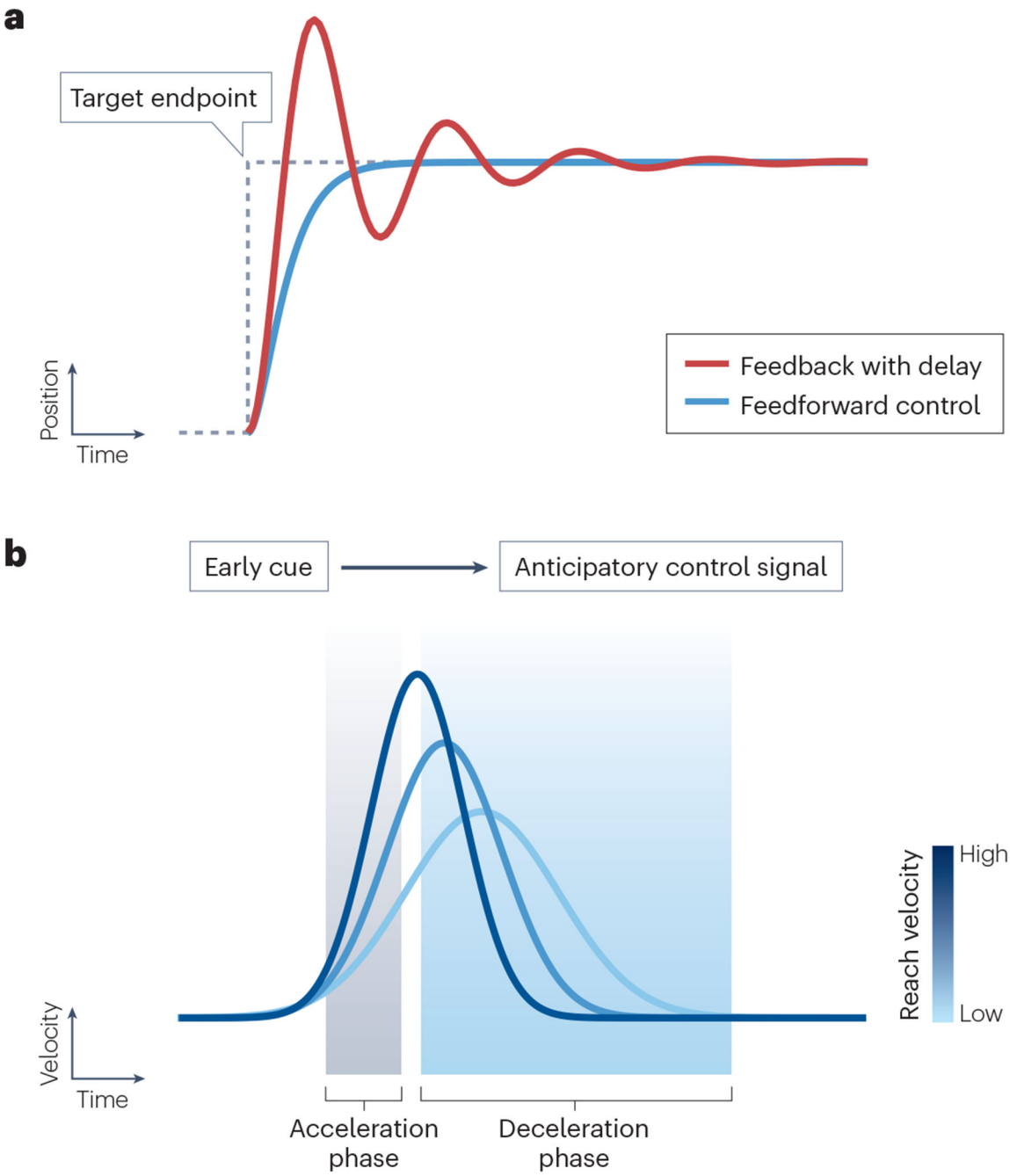

Fig 1: Effects and use of sensory delays in motor control systems.

a, Oscillations around a target endpoint emerge in feedback driven control systems with long sensory delays[4,11]. Such poorly engineered feedback control systems update their output based on feedback that is out-of-date with respect to the current state, resulting in overshooting and overcorrecting and creating oscillatory endpoints. Control systems that incorporate learned feedforward control avoid these endpoint oscillations, approaching but not overshooting the target[9]. **b**, A learned feedforward policy can use contextual information early in the movement to generate an anticipatory control signal that guides the effector

predictively to the target. The example illustration shows how the kinematics during the early phase of a movement can be used as cues to shape the late phase kinematics. Here the velocity profile of a reach remains bell-shaped regardless of the reach velocity (represented by the shade of the line) because the acceleration phase is used as a cue to structure the deceleration phase.

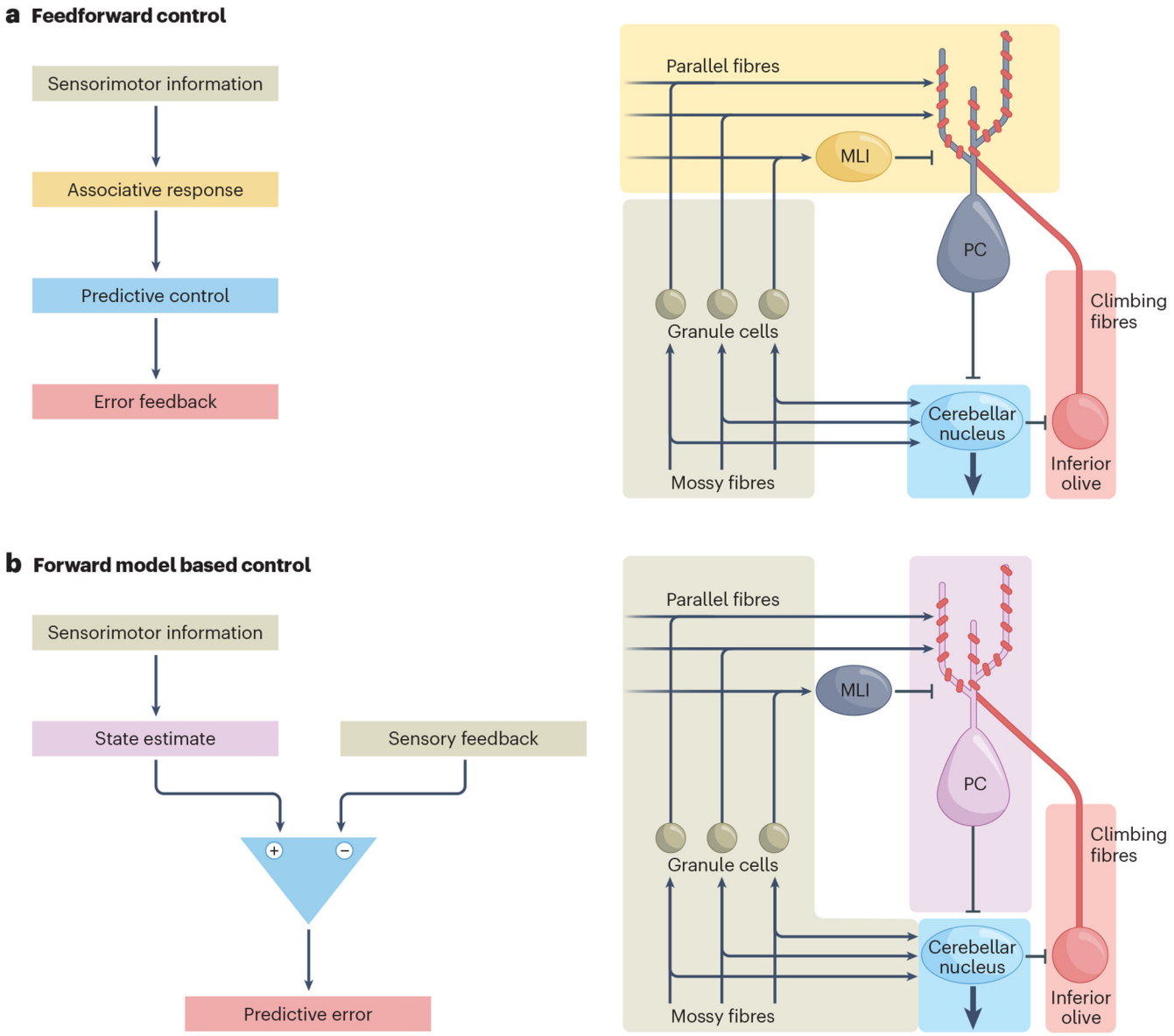

Fig 2: Comparison between alternative control system models.

a, Feedforward controllers use sensorimotor information as a cue to generate an associative response, implementing predictive control. Feedback of motor errors are sent via climbing fibers originating in the inferior olive (IO). As described in the main text, sensorimotor information is conveyed to the cerebellum through mossy fibers and reformatted in the granule cell layer. Granule cell parallel fiber-to-PC synapses are subject to plasticity based on associative learning rules in conjunction with climbing fiber instructive signals, often active following errors. This associative learning leads to PC responses that implement predictive control by the cerebellar nuclei (CbN), enhancing motor precision and accuracy. **b**, Forward-model based controllers use sensorimotor information (conveyed by mossy fibers) to generate a state estimate (learned and represented by PCs) that is used to interpret sensory feedback. Mismatches between predicted and actual feedback generate a sensory

prediction error (within the CbN or IO), which can be used to update a motor plan, identify errors used in update rules to tune the network and differentiate self from other, among other functions[7,177].

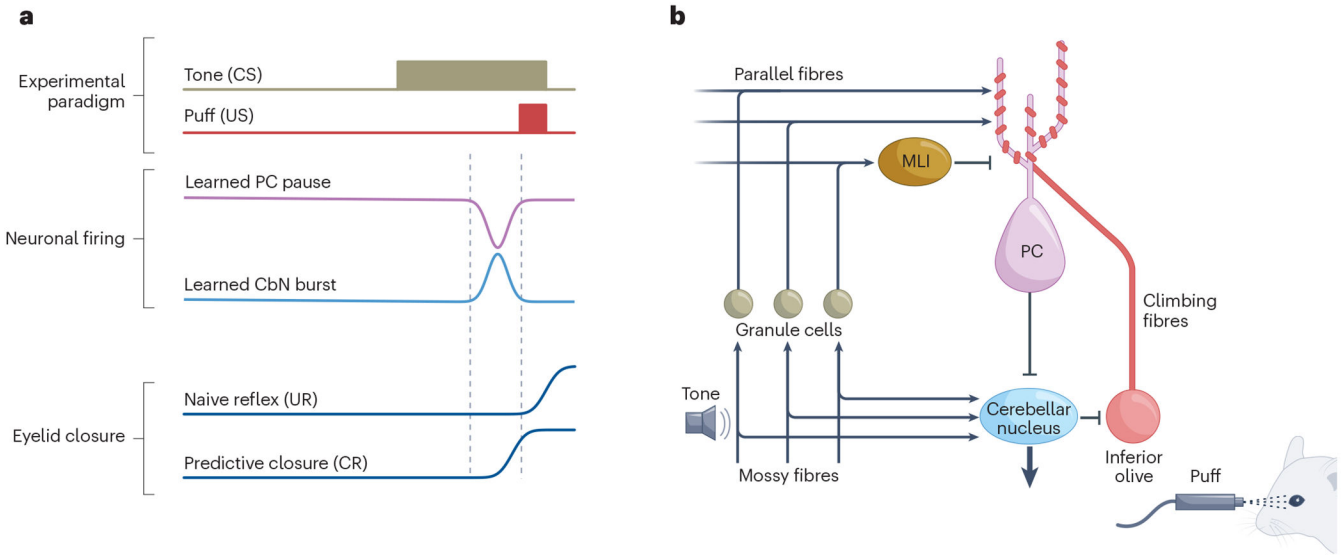

Fig 3: Delay eyelid conditioning as a canonical cerebellar feedforward control paradigm.

The left panel shows a schematic illustration of a widely-used cerebellar learned timing task, delay eyelid conditioning (DEC). The right panel shows the cerebellar circuit proposed to implement learning via feedforward control. In this behavioral paradigm, a tone (the conditioned stimulus (CS)) is paired (over repeated trials) with an airpuff (the unconditioned stimulus (US)) directed at the eye. In naive conditions, the CS elicits no behavioral response but the US elicits a reflexive eyeblink, an unconditioned response (UR). Over training, mossy fibers convey the CS to the cerebellum, where they impinge on granule cells. These in turn project to Purkinje cells (PCs) and molecular layer interneurons (MLI) via parallel fibres. Airpuffs elicit activity in the IO which is conveyed to PCs via climbing fibers. Conjunctive activation of parallel fibers and climbing fibers weakens parallel fiber synaptic strength through LTD. This process occurs following repeated pairing of the CS and US, resulting in animals generating a predictive eyelid closure (the conditioned response, CR), with the timing of this closure determined by a learned pause in PC firing that disinhibits output neurons in the cerebellar nucleus (CbN) burst).

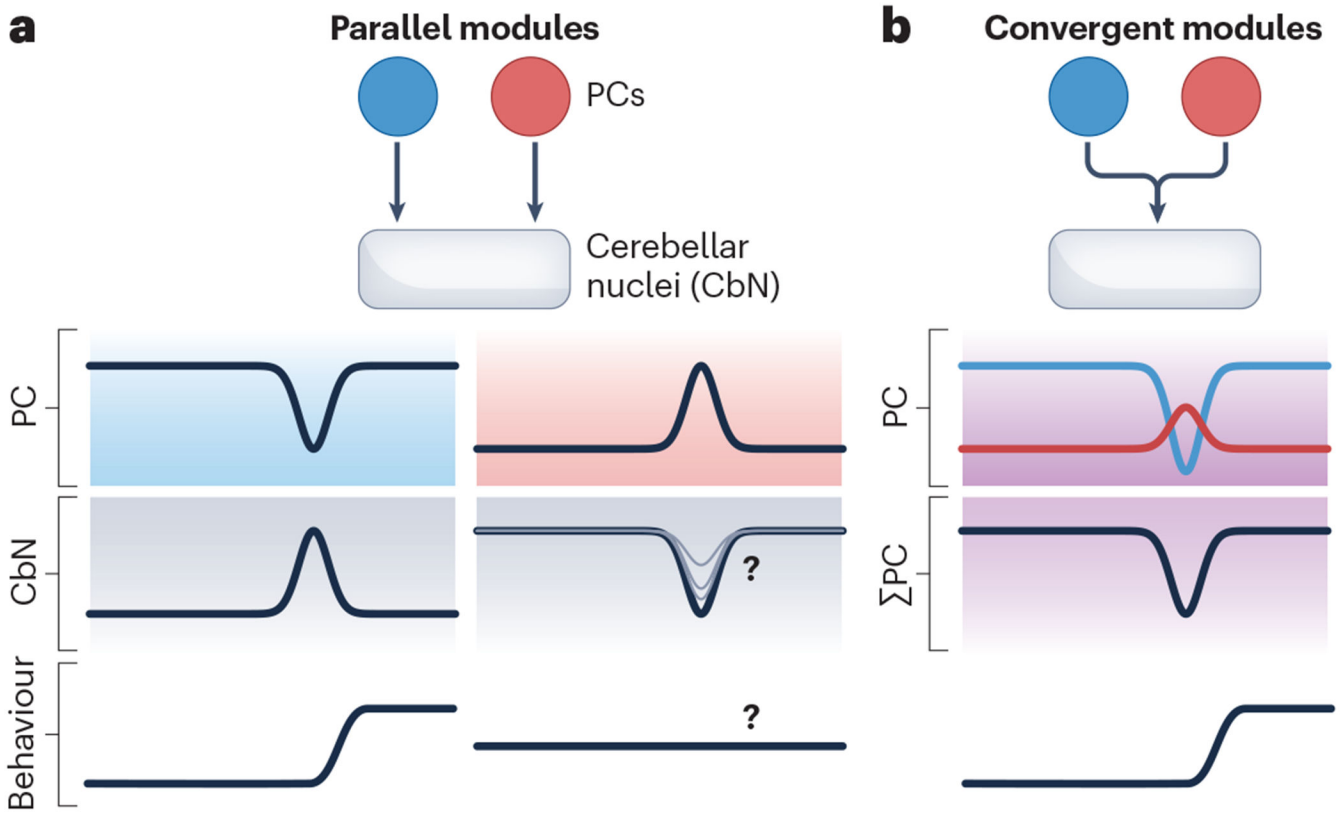

Fig 4: Resolving Purkinje cell coding diversity.

Some cerebellar Purkinje cells (PCs) show net suppression of simple spiking (blue shading) in response to a learned cue such as a tone, or during movements such as saccades, whereas others show facilitation of simple spiking (red shading)[61,80,84]. How these distinct populations collaborate to encode motor-related information is unknown, but two possible scenarios are summarized here. **a**, If PCs that exhibit simple spike suppression during a behavior remain segregated from those that exhibit facilitation, these signals would disinhibit the cerebellar nuclei to actuate control of behaviour. The simple spike facilitation that is commonly observed in some PCs is often tuned to kinematics or to the sensory environment[114]. However, the effect of the facilitation on cerebellar nuclear output or behaviour is not always clear, owing to the unknown levels of concurrent excitation that enter the cerebellar nuclei via mossy fibers(indicated by the question marks). **b**, An alternative possibility is that heterogeneous PC populations (i.e. suppressed and facilitated PCs) converge in the cerebellar nuclei. Such convergence could implement some of the forms of cancellation that have been hypothesized to be mediated by internal models and would resolve some of the ambiguity regarding the relationship between PCs and cerebellar nuclear activity. When considered as a population, the heterogeneous PC activity (∑PC) may show net simple spike suppression at specific timepoints, with a tighter temporal relationship to behavior than suppression or facilitation alone. This hypothesis may also account for inverse nuclear firing rates, where net suppression of PCs corresponds to net increase in CbN activity via disinhibition [84,86–88].

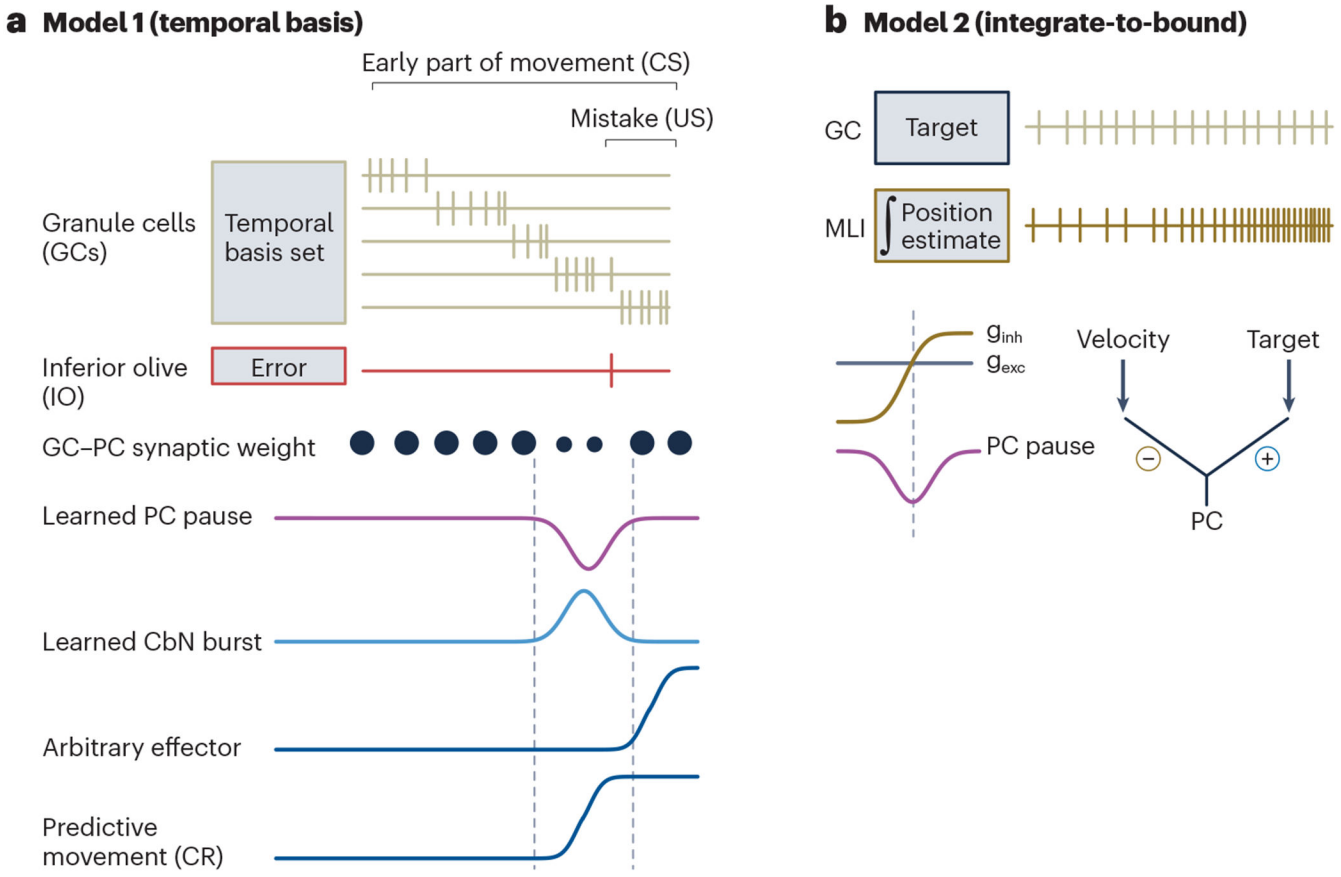

Fig 5: Two models of the cerebellar control of movement.

The cascade of events that take place during delayed eyelid conditioning (Fig. 3) can be viewed as a general framework for feedforward control by the cerebellum. By analogy, stopping movements on target could employ similar computations. Two distinct hypotheses for how this computation may form are illustrated. **a**, In the temporal basis model, granule cell populations (GCs) represent elapsed time or spatiotemporal state within a reach, such that subpopulations of GCs fire at different elapsed times during the movement.. Errors communicated by climbing fibers from the inferior olive induce heterosynaptic long term depression (LTD), at GC to Purkinje cell (PC) synapses, reducing the synaptic efficacy of the subpopulation of GCs that fired just prior to the error. On subsequent reaches, the excitatory drive from this subpopulation of GCs remains depressed, leading to suppressed PC activity (PC pause). Thus, the early phase of motor commands or kinematics (the conditioned stimulus (CS) in the schematic), if paired with a mistake (the unconditioned stimulus (US) in the schematic), could become predictors used to generate pauses, disinhibit the cerebellar nuclei (CbN) and drive predictive stopping or starting (akin to a conditioned response (CR) of an arbitrary effector (i.e. eye, limb) to avoid errors in the future. Theoretically, spatial goals could be represented within the GCL relaying target-encoding mossy fiber information [41,84,167]. **b**, In the integrate-to-bound model, target location is relayed by the mossy fibres to GCs, which convey this information directly to PCs via a topographic code, while the displacement of the effector from the target is

computed from velocity information in the GCs integrated in inhibitory molecular layer interneurons (MLIs). As the target is approached, the level of inhibition grows (summed inhibitory conductance (g_{inh})) to match and then exceed excitation (summed excitatory conductance (g_{exc}), leading to suppression of PC activity and disinhibition of the cerebellar nuclear actuation, stopping the movement [87]. This novel hypothesis for cerebellar cortical computations is conceptually related to classic models of dynamic motor error control of saccades proposed for brainstem circuits[31,191].

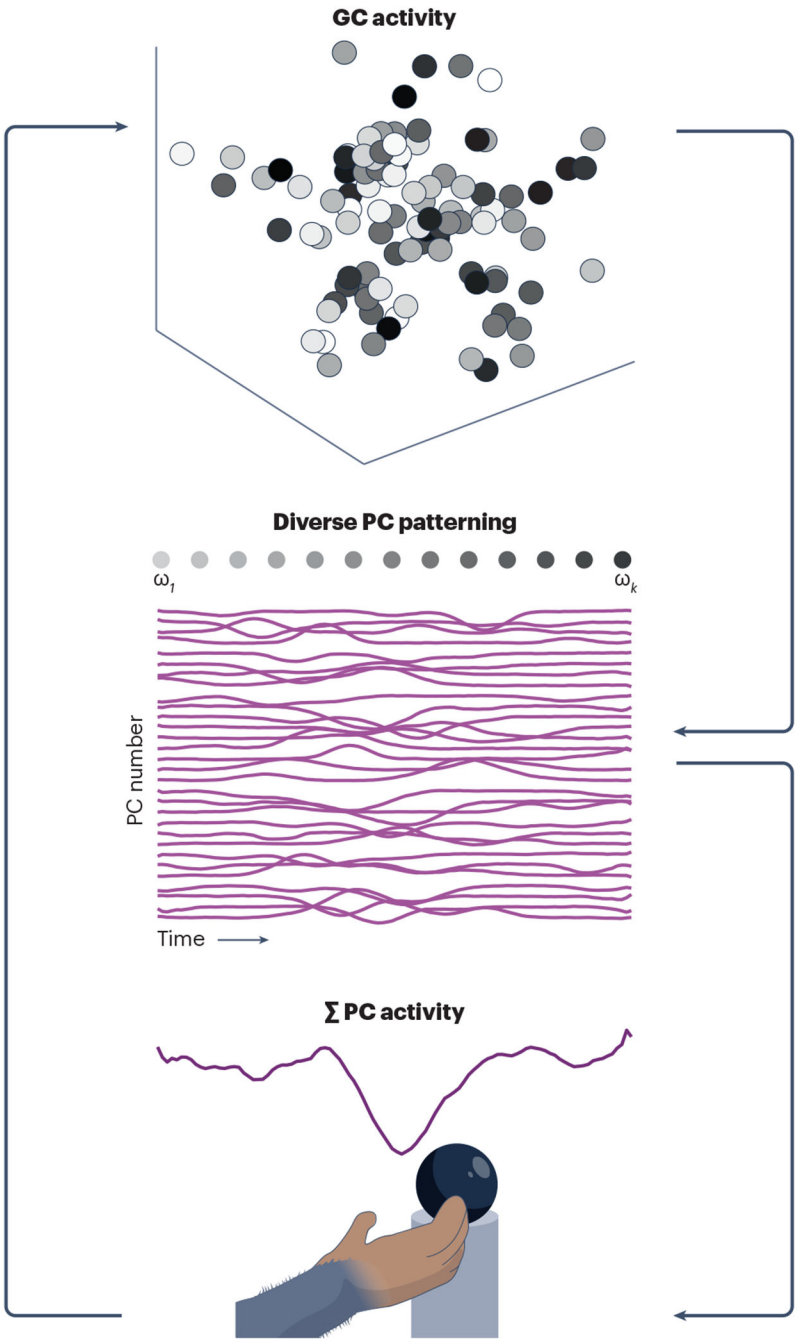

Fig 6: High dimensional granule cell population codes give rise to diverse signals.

High dimensional GC activity (left), schematized in 3 dimensions (axes). The dimensionality of GC representations is very high[155], indicating narrow spatiotemporal tuning. The diverse onset timing of GC activity is indicated in grayscale, representing an embedded temporal basis within the GC population code. When coupled with the diverse instructive feedback from the inferior olive received by PCs, adjusting synaptic weights, this network structure can give rise to diverse PC codes (right, wavy lines representing learned, time-varying PC firing rates during a movement). Some PCs may exhibit firing patterns that

resemble forward models, others feedforward control policies. The function that emerges from their population activity will depend on the tuning of the instructive signals that drive plasticity at the GC-PC synapses. For example, according to the LTD model, if the climbing fiber is tuned to limb target overshoot, PCs will learn to pause near endpoint[84,122], while if the climbing fiber is tuned to reward delivery, PCs will learn to pause in anticipation of reward[116,119,165]. Different learning rules will further diversify PC patterning.