



HAL
open science

WHAT IF...? RETHINKING THE ROLE OF LEARNING IN NEURAL NETWORKS

Mike Gilbert, Anders Rasmussen

► **To cite this version:**

Mike Gilbert, Anders Rasmussen. WHAT IF...? RETHINKING THE ROLE OF LEARNING IN NEURAL NETWORKS. 2025. <hal-04863982>

HAL Id: hal-04863982

<https://hal.science/hal-04863982v1>

Preprint submitted on 4 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

1

2

3

4 **WHAT IF...? RETHINKING THE ROLE OF LEARNING IN NEURAL NETWORKS**

5 **Mike Gilbert^a (corresponding author), Anders Rasmussen^b**

6 ^aSchool of Psychology, College of Life and Environmental Sciences, University of

7 Birmingham B15 2TT, UK

8 ORCID: 0000-0003-4415-6915

9 ^bDepartment of Experimental Medical Science, Lund University, BMC F10, 22184

10 Lund, Sweden

11 ORCID: 0000-0003-4477-369X

12

13 Correspondence: m.gilbert.1@bham.ac.uk

14

15

16 **Abstract**

17 The cerebellum is a large brain structure which is highly conserved across vertebrate
18 groups and widely connected with other parts of the brain. It is a mainstay of many
19 neural network models, including cerebellar models, that changes of synaptic
20 transmission strength learned by experience or under instruction displace the naive
21 response to information coded in firing rates. This idea has strongly influenced
22 theoretical thinking. However, it requires replicability – meaning, data are coded as
23 bespoke neural activity which is accurately repeatable. This is intuitive because it is
24 typically our experience of language, including computer code: the same language is
25 used to repeat the same information. This may not be a safe assumption. We reasoned
26 that, in the cerebellum, there must be a reliable learning outcome whether patterns
27 are repeated or not, including the scenario that they are never repeated. Indeed, this
28 may be a core cerebellar function – distilling a clean signal from untidy inputs. Here, we
29 propose a mechanism and function. There is no learning algorithm or pattern memory,
30 both of which are unnecessary to explain the evidence. If we are correct, 50 years of
31 assumptions that have traditionally underpinned neural network models may need to
32 be rethought.

33 **Keywords** cerebellum; theory; learning; synapse; model; network; code

34

35

36

37

38

39

40

41

42 **1. Introduction**

43 The cerebellum is modular – made up of repeating cell networks. Each network
44 includes a microzone [1, 2], a long, thin strip of the cerebellar cortex which contains a
45 few hundred Purkinje cells. Purkinje cells carry the sole output of the cerebellar cortex.

46 Artificial neural networks typically depend on learned synaptic modifications to convert
47 input data coded in firing rates into network output. In cerebellar models, the key site
48 of learning is where parallel fibres make synaptic contact on Purkinje cells and in some
49 models also on interneurons that inhibit Purkinje cells. Parallel fibres are the axons of

50 granule cells, small and very numerous cells which receive the principle excitatory
51 input to the cerebellum, from mossy fibres.

52 Synaptic changes following training are reported at both parallel fibre synapses on
53 Purkinje cells [3-5] and on interneurons [6-11] (Fig 1). (We use 'interneuron' for
54 'molecular layer interneuron'.) However, confirmation of the theoretically-expected
55 role of long-term depression in motor learning remains pending [12-16]. There is
56 evidence of linear signalling at both the parallel fibre-Purkinje cell [17, 18] and parallel
57 fibre-interneuron [19, 20] synapses, as well as between other cerebellar cells [20-24].
58 Cerebellar physiology includes adaptations for high fidelity transmission of rate
59 information [25-31]. Firing rates 'at all levels of the cerebellar circuit' [32 p.239, citing
60 30 sources] are linearly related to behavioural metrics. It is difficult to reconcile linear
61 signalling with learning models since it is the function of learned synaptic weight
62 changes is to *displace* a purely rate-coded response.

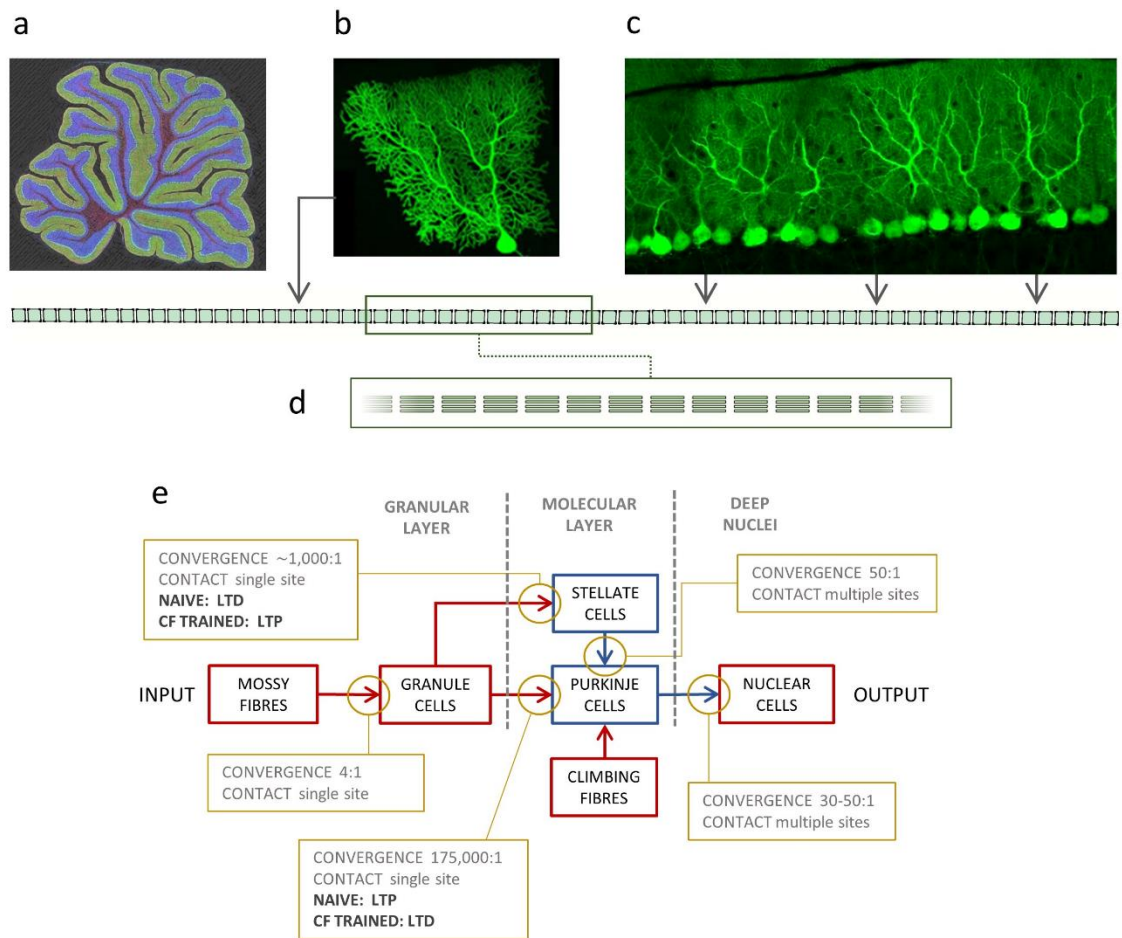
63 Moreover, learning models are predicated on the assumption that neurally-coded data
64 are replicable – a repeat of the same data reliably takes the same form of neural
65 activity. Otherwise, signals would not be received at the correct synapses. Replicability
66 may not be a safe assumption. Possibly, patterns are *never* repeated. Single neurons
67 and synapses do not behave with the reliable consistency often envisaged by models.
68 The performance of single cells can be variable, including cells of the same type, and
69 synaptic transmission involves random variables. ~8,000 mossy fibres drive parallel

70 fibre input to a single Purkinje cell. It is doubtful that even an exact repeat of
71 conditions upstream (if that ever happens) would trigger a computer-like replica of
72 input to the cerebellum. In addition, the input layer of the cerebellar network, the
73 granular layer, where mossy fibres terminate, seems to be *designed* to convert input
74 into pseudo-randomised internal activity [33, 34], with the result that variation of input
75 becomes a wholesale change to the pattern of active and inactive parallel fibres.

76 We reasoned that learning must return a reliable outcome whether patterns are
77 repeated or not. On the face of it, this idea has the weakness that it is not obvious how
78 information would be captured in a pattern that is never the same, or how learning
79 would ever have a useful outcome. We propose how the cerebellum may have solved
80 those problems. We will argue that training teaches synaptic weights to polarise.

81 Memory is stored in the ratio of binary synaptic weights, not the pattern. The
82 microzone is the minimum functional unit of learning and memory, so that there is a
83 coherent learning outcome at that level. The function is to calibrate the relationship of
84 inputs and output of a microzone, not to 'compute' or fundamentally change it. The
85 conversion itself is unlearned – control by firing rates is not displaced by learning.

86 Calibration is indiscriminating – the same effect is applied to all input to the system, in
87 all patterns of active and inactive cells. The learning outcome does not depend on
88 pattern replication. We make a case in evidence, with computational support.



89

90 **Fig 1. Schematic of data flow in the cerebellum.**

91 a Sagittal section of the rat cerebellum (obtained by multi-photon microscopy, courtesy Tom
 92 Deerinck). The outer layer, the molecular layer, of the cerebellar cortex is stained green and the
 93 inner layer, the granular layer, is blue. **b** and **c** Purkinje cells in a sagittal slice (two-photon laser
 94 scanner micrograph, courtesy Mike Häusser). The Purkinje cell arbour is large from this view,
 95 filling the molecular layer vertically. It is flattened in the plane orthogonal to parallel fibres. **d**
 96 Top. Schematic of a parallel fibre view (i.e., the same view as a–c) of a row of 80 Purkinje cell
 97 territories, but straightened out. The molecular layer is divided functionally into long, thin

98 strips called microzones. At a mid-range estimate, a microzone is 75–100 Purkinje cells long and
99 4 wide (measuring 1,500–2,000 x 150 μm). Inset. A short section viewed from the cerebellar
100 surface. e Simplified circuit wiring, highly preserved across vertebrates. Red: glutamate; blue:
101 GABA. Brown boxes contain information about network connections. Microzones lie at right
102 angles to parallel fibres. Tens of millions of parallel fibres pass through a microzone, making
103 contact in passing (if at all) on Purkinje cells and interneurons.

104

105 **2. What is the naturally-occurring synaptic learning outcome?**

106 Purkinje cells receive strong innervation from climbing fibres. Long-term depression
107 (LTD) of the parallel fibre-Purkinje cell synapse can be induced by repeated pairing of
108 synaptic activation which commences shortly before a climbing fibre signal. Plasticity at
109 this synapse is much studied. Early reports of parallel fibre LTD were of a group 1
110 metabotropic glutamate receptor (mGluR1)-dependent mechanism expressed
111 postsynaptically, as a reduction in the number of functional AMPA receptors [3, 5].
112 mGluR1s are expressed postsynaptically in a perisynaptic zone around AMPA receptors.
113 Presynaptic LTD has also been reported, which is NMDA (but not mGluR1) dependent
114 [4], along with presynaptic [35, 36] and postsynaptic [37, 38] long-term potentiation,
115 so that LTD is reversible and in theory can be 'reset' [39].

116 Naturally occurring long-term depression of the parallel fibre-Purkinje cell synapse
117 blocks a somatic response of electrical stimulation of parallel fibres in slices [16]. Most

118 parallel fibre-Purkinje cell synapses (estimated at 80–85% [16] or more [15]) are
119 naturally long-term depressed, and while they are not all fully stripped of AMPA
120 receptors – although some are [40] – they are reported to be ‘silent’ [15].
121 Plasticity can be trained in both directions at synapses of parallel fibres on MLIs. MLIs
122 express both AMPA and NMDA receptors extrasynaptically [41] but only AMPA
123 receptors synaptically [42]. Studies of the parallel fibre-MLI synapse are fewer than at
124 the parallel fibre-Purkinje cell synapse, and the postsynaptic induction mechanism
125 incompletely understood, but it appears to occur in the same conditions as the parallel
126 fibre-Purkinje cell synapse, but with the sign reversed (*in vitro* [6-8], *in vivo* [9-11]).
127 Naturally occurring LTD seems to be uniformly severe [15] – there is no response even
128 to stimulation at hundreds of Hz.

129

130 **3. What is a ‘pattern’ of parallel fibre activity?**

131 The granule cell axon rises from the granular layer into the outer layer of the cerebellar
132 cortex, the molecular layer, where it divides in two to form parallel fibres. The two
133 branches travel in opposite directions parallel to the cerebellar surface and to other
134 parallel fibres, in the direction of the long axis of folia. Parallel fibres lie at right angles
135 to microzones and to the plane of the severely flattened Purkinje cell dendritic arbour.
136 The Purkinje cell arbour is large, filling the molecular layer vertically. As estimated

137 350,000 parallel fibres pass through a single Purkinje cell territory [43] and many
138 millions pass through a microzone.

139 It is thought that co-active parallel fibres are a quasi-random sample of the general
140 population. Activity is said to be decorrelated [34, 44], meaning that the subset of
141 granule cells which are active at any moment, and therefore the pattern of active and
142 inactive parallel fibres, has no evident correlation with the pattern in which mossy
143 fibres are active upstream.

144 Granule cells typically fire in short, high-frequency bursts. The reported lengths of
145 bursts varies: 10-20 ms has been reported in adult cats [45] and 8-40 ms in rabbits [46].
146 For part of a burst, in most cases, the whole cell is active. That is, transmission time of
147 a spike is shorter than the sum of interspike intervals, even with transmission times
148 that are 'probably the slowest...in the whole brain' [47 p.41]. However, the active part
149 of a signal seems to be the initial segment. At the parallel fibre-Purkinje cell synapse,
150 high frequency bursts cause short-lived facilitation of release probability within the first
151 few spikes rapidly followed by a reduction of neurotransmitter release [46]. Granule
152 cell bursts are only a few spikes long anyway, so this may be a noise reduction strategy,
153 i.e., a physiological method of eliminating a variable effect of a randomly variable
154 number of spikes in a burst.

155 Patterns, as opposed to signals, however, are not discrete. During behaviour, in
156 locomotor circuits, there is a constant turnover of active cells. As a result, Purkinje cells

157 receive a rapid turnover of signals which is constantly refreshed. That is, at this
158 collective level, data are streamed, notwithstanding that individual signals are short
159 and discrete. The relative timing of signals is randomised (there is no correlation
160 between cells of onset and offset times). A full turnover of active cells that pass
161 through an orthogonal plane (and therefore both a Purkinje cell and a microzone)
162 occurs in approximately the mean duration of a burst.

163 To receive modification under climbing fibre tuition, onset of parallel fibre synaptic
164 activation must occur in a time window ending before the teaching signal [48]. Timing
165 varies from cell to cell and perhaps regionally [49]. Duration of the window is
166 comfortably long enough to contain more than a single turnover of parallel fibre
167 signals, and in at least some cases long enough to contain several (an optimal lead time
168 of 50–200 ms has been reported[48]). The number of synapses on a Purkinje cell which
169 are activated in a learning window is therefore larger – and may be several times larger
170 – than a snapshot of activations (and may be further increased because it is bookended
171 by sub-‘optimal’ learning).

172 A ‘pattern’ of unpaired synaptic activations is made up of all activations that fall outside
173 a climbing fibre learning window. Climbing fibres fire at a low rate (never more than a
174 few Hz). Accordingly, in locomotor circuits during behaviour, the unpaired learning
175 window has what is presumably an unpredictably variable duration and also contains
176 multiple turnovers.

177 **4. Parameters of the synaptic learning function**

178 Transmission strength of a parallel fibre synapse depends on its training history, and
179 more exactly its more recent history, measured in learning episodes, because change is
180 reversible in both directions. We hypothesised that the parameters of the synaptic
181 learning function are confined to selected variables, and that biophysics are
182 subordinated to this result. We discuss function parameters in turn.

183 **4.1 The sequence and ratio of synaptic activations**

184 Two snapshots even of sparse randomly-patterned parallel fibre activity are statistically
185 assured to overlap to some extent at the numbers received by a Purkinje cell. A pattern
186 of co-activated synapses is denser than the regulated density of co-active parallel
187 fibres, so there is heavier overlap. A single synapse may participate in multiple patterns
188 (as long envisaged [44]), both paired with a climbing fibre signal and not paired. The
189 number and frequency of occurrences of a pattern, and sequence and ratios in which
190 patterns occur, are outside cerebellar control. Since participation of a single synapse in
191 a pattern of activations is unrelated to its participation in any other pattern, each
192 synapse has a unique history – sequence and ratio – of climbing-fibre-paired and
193 unpaired activations. This is unrelated to its participation in any particular pattern, and
194 randomly related to the sequence and ratio at any other synapse. Moreover, that
195 relationship is dynamic.

196 **4.2 Learning step size**

197 It is likely that the amount of synaptic change per activation depends partly on
198 variables that are affected by prior learning. For example, if synaptic changes affect the
199 availability of postsynaptic calcium, on which learning is dependent, it follows there is a
200 learned effect on further training. Learning may also receive an effect of a prior
201 reduction or increase of the number of postsynaptic AMPA receptors. An important
202 mechanism of learned postsynaptic change is removal or insertion of AMPA receptors
203 from or into the postsynaptic membrane [39, 50]. Removal of a fixed fraction per
204 learning episode would cause a diminishing per-episode number to be internalised, so
205 that step size decreases towards the limits of the range. Conversely, at high numbers,
206 the number of receptors inserted per episode may diminish towards saturation.

207 **4.3 Nucleo-olivary feedback**

208 Cerebellar circuits send an inhibitory projection to the contralateral olivary complex
209 [51, 52]. Training teaches inhibitory feedback. Feedback is precisely targeted and timed
210 to intercept the lesson that trained it at source [53]. The function of feedback is not
211 agreed. It has been argued that climbing fibre signals teach a graded lesson that
212 depends on the number of spikes they contain [54, 55], and counter-argued that the
213 same evidence in fact shows they don't [56]. In the counter-proposal, signals have an
214 all-or-nothing signature, as originally reported [57, 58]. There is recent evidence

215 against graded signals [59]. If lessons are binary (on any occasion, either there is a
216 lesson or there is no lesson), it follows that feedback is also binary – not the firing rate
217 of the feedback signal, but because it can only have two outcomes, a full block or no
218 effect. If it is strong enough to counteract excitatory drive to the inferior olive that
219 would otherwise evoke a climbing fibre signal, the outcome is to block the lesson on
220 that occasion. If not, there is no effect. The effect of a block is to reverse the direction
221 of learning in that training episode. So, again, learning has an effect on further
222 learning, but this time not a graduated effect but acting through a threshold.

223

224 **5. An iterative model predicts that synaptic weights polarise**

225 These parameters have previously been combined in an iterative model [60] where
226 training was with a randomly ordered sequence of patterns which share synapses,
227 some paired with climbing fibre instruction, some not (summarised in Box 1). The
228 result is that synaptic weights *within* a pattern polarise, so that, following training,
229 some synapses are severely depressed and other synapses (the rest) are robustly
230 potentiated. This is because the strength of feedback is a compound effect of synaptic
231 changes, and when it is strong enough to block instruction in paired trials, there is a
232 learned effect on the ratio of paired and unpaired trials. As a result, at more weakly

233 trained synapses thereafter the sum of the size of steps in unpaired trials is larger than
234 positive. Synapses where learning is more advanced proceed towards saturation.
235 Weights polarise because the limits of the learning curve are more stable than the
236 middle. As weights always polarise, the only (variable) learning output is the ratio of
237 silent to working synapses. The ratio depends on the feedback threshold, i.e., the
238 strength of feedback needed to block an instruction signal, which depends on the
239 strength of excitatory drive to the inferior olive (and more exactly the relative strength
240 of drive that excites instruction signals and drive that excites parallel fibre signals that
241 are paired with it). All synaptic groups learn the same ratio of trained to untrained
242 (silent to working) synapses, notwithstanding that all synapses have a unique history. It
243 makes no difference to this outcome how many patterns parallel fibres cycle through,
244 or in what order, or how often (if ever) they are repeated.

245

246

BOX 1

247 **An iterative model of parallel fibre synaptic learning**

248 This is a summary. A full description has appeared previously [60]. The core of the model
249 is:

250

$$L_n = L_{n-1} + k(T_n - L_{n-1})$$

251 L represents synaptic strength. It has the range $0 < L < 1$. L_n is the outcome of learning
 252 episode n . Let us say an increasing value of L represents progressive weakening of
 253 synaptic strength. T_n is either 1 or 0, representing respectively a climbing fibre lesson, or
 254 no lesson. k is a constant which adjusts Δ /trial (step size). $L_{\lim \rightarrow 1}$ is where an effect of
 255 further training is exhausted. This does not necessarily – but may – mean AMPA
 256 receptors are entirely absent from the postsynaptic membrane.

257 As learning can be in both directions, and the size of learning steps diminishes towards
 258 the limits (also in both directions), the equation is extended to

$$259 \quad L_n = L_{n-1} + k(T_n - L_{n-1}) * \begin{cases} f_A(L_{n-1}), & T_n = 1 \\ f_A(1 - L_{n-1}), & T_n = 0 \end{cases}$$

260 This modifies learning steps in the lower range of L . $f_A(x)$ is a function which adjusts the
 261 size of the effect, treated as a constant equal to 1.

262 Training teaches inhibitory feedback which is precisely targeted and timed to block the
 263 signals that trained it [53]. If climbing fibre instruction signals are functionally binary, as
 264 we think, then in any trial, feedback either blocks the lesson in that trial, or has no effect.

265 In a trial where a lesson is blocked, the direction of learning is reversed. Using the
 266 Heaviside function, $H(x)$, we might express that as:

$$267 \quad T_n = H(f\langle r \rangle - \langle rL_{n-1} \rangle)$$

268 where rL_{n-1} is a measure of the joint effect at a single synapse of afferent signal

269 frequency, represented by r , and prior learning, which we take as the product. The effect
 270 of training on feedback is a compound effect of the learning outcome to date at all
 271 synapses, represented by $\langle rL_{n-1} \rangle$, the mean of the products. This is an abstraction which
 272 we use to represent the strength of inhibitory feedback. It is related to learning because
 273 synaptic depression disinhibits the feedback pathway. $f\langle r \rangle$ is a function of the mean of
 274 all rates, the threshold where a rising value of $\langle rL_{n-1} \rangle$ blocks olivary discharge, so that
 275 there is no climbing fibre lesson in that instance. So:

$$276 \quad L_n = L_{n-1} + k([H(f\langle r \rangle) - \langle rL_{n-1} \rangle] - L_{n-1}) * \begin{cases} L_{n-1}, & T_n = 1 \\ 1 - L_{n-1}, & T_n = 0 \end{cases}$$

277 The threshold represents excitatory drive to the inferior olive, which we take as a
 278 function of parallel fibre rates – that is, we take them to be related. The function that
 279 relates them determines the ultimate ratio of silent to working synapses: the learning
 280 outcome. END BOX.

281

282 **6. What does that do?**

283 Speaking generally, neurons fire at a rate that varies with the number of co-active
 284 inputs, with a larger number having a stronger effect, other things being equal.

285 Decorrelation has the consequences that depressed and working parallel fibre-Purkinje
 286 cell synapses are randomly intermingled, and that active inputs to a Purkinje cell are

287 received at the same ratio of weights as the ratio in the general synaptic population.
288 This result is the same regardless of the pattern in which parallel fibres are active,
289 whether patterns are replicated, sometimes replicated or never replicated. Therefore,
290 the ratio of weights controls the number of concurrent activations of working synapses,
291 assuming that active parallel fibres are a constant fraction of the general population
292 (this condition is satisfied [61]).

293 Thus, learning applies a pattern independent adjustment to the balance between
294 excitation and inhibition of Purkinje cells. Purkinje cell and granule cell morphologies,
295 and cortical architecture, bring about this result. The size of the adjustment is
296 proportional to the ratio of synaptic weights.

297

298 **7. Microzones learn as a single unit**

299 Microzones are defined by their climbing fibre input [1, 62], which is from a
300 functionally discrete group of olivary cells. A Purkinje cell receives contact from a single
301 climbing fibre and a single climbing fibre innervates 1–10 Purkinje cells. As instruction
302 signals are functionally invariant, and the number of climbing fibres that contact a
303 Purkinje cell is always one, the lesson is invariant. As far as we know, there is no
304 evidence of internal organisation of the olivary group-microzone pathway.

305 In response to stimulation, climbing fibres in a group fire together, so that a microzone
306 that receives the output of a group receives climbing fibre tuition as a volley. Discharge
307 synchrony of a climbing fibre group is enhanced by dendritic electrotonic coupling of
308 olivary cells [63] and is seen in anaesthetised mice in response to sensory stimuli [64].
309 Coherence of complex spikes between functionally-grouped Purkinje cells (Purkinje
310 cells reliably respond to climbing fibre discharge with a complex spike) is low when it
311 includes spontaneous discharge but high in response to sensory stimuli [65]. This
312 provides the substrate for microzone-grouped Purkinje cells to receive the same
313 training at the same times, so that they learn as a unit.

314 Climbing fibre communication with stellate cells is exclusively by glutamate spillover
315 (Szapiro and Barbour 2007). Interneurons express both AMPA and NMDA receptors
316 extrasynaptically. The standard shape and dimensions of microzones suggests that the
317 peak amplitude and time course of spillover is the same at all locations, so again the
318 whole population receive the same lesson as a unit.

319

320 **8. Does learning at the parallel fibre-stellate cell synapse affect** 321 **calibration of the balance?**

322 As interneurons learn in the same conditions as Purkinje cells, we might expect the
323 ratio of silent to working synapses to be the same but reversed, provided there is no

324 effect of cell morphology on the ratio. By our calculation, there is no effect of
325 morphology (Box 2). As (we have proposed) synaptic weights are binary, a possible
326 effect of learning is limited to a possible effect of (severe, as noted) synaptic
327 depression. We can determine if there is an effect by comparing the distributed
328 probabilities of the number of active inputs to a stellate cell and the number received
329 at working synapses.

330 Many fewer parallel fibres contact an interneuron than a Purkinje cell. The stellate cell
331 territory is flattened in the same plane as Purkinje cells. Stellate cells form planar
332 networks [66] that occupy the space (about 40 μm) between Purkinje cells [67]. The
333 Purkinje cell dendritic territory is $\sim 300 \times 200 \mu\text{m}$. The stellate cell dendritic territory is
334 around $80 \times 80 \mu\text{m}$ for superficially located ('outer-level') cells and around 120×120
335 μm for deeper lying ('inner-level') cells [67 pp.217-221]. Around 350,000 parallel fibres
336 pass through the Purkinje cell arbour [43]. Therefore, $350,000 / ((200 \times 300) / 120^2)$ or
337 $\sim 84,000$ parallel fibres therefore pass through an inner-level stellate cell dendritic
338 territory, assuming parallel fibres are uniformly distributed. Of those, around 420 are
339 active if 0.5% are active in the general population.

340 It is thought a stellate cell receives 'several hundred excitatory inputs from granule
341 cells' [67, 68, p.1310], perhaps as many as 1,000. 1,000 would mean around 1 in every
342 84 parallel fibres which pass through a stellate cell territory makes contact (i.e., 1,000
343 out of 84,000). Accordingly, there is a probability of $1/84 = \sim 0.0119$ that an active cell

344 makes contact. The probability of contact by k (out of 420) active cells is therefore
 345 given by

346
$$\frac{420!}{k!(420 - k)!} * 0.0119^k * (1 - 0.0119)^{420-k}$$

347 This gives the Table 1 probabilities for the range $k = 0$ to 12, where k is the number of
 348 active parallel fibres that contact a stellate cell and p is the probability of that number.
 349 There is a high probability ($p = 0.8934$) that the number lies between 1 and 8. This is in
 350 good agreement with the low reported number suggested by 'two to eight substantial
 351 EPSPs [excitatory postsynaptic potentials]' [10 p.9628].

352 **Table 1 The probability, p , that a stellate cell receives contact from k active parallel**
 353 **fibres**

k	0	1	2	3	4	5	6	7	8	9	10	11	12
p	.0066	.0331	.0836	.1403	.1762	.1765	.147	.1047	.0651	.0359	.0178	.008	.0033

354

355

356

BOX 2

357 **Cell morphology does not affect the fraction of climbing fibre-trained parallel fibre**
 358 **synapses**

359 The total number of climbing fibre-trained synapses given n stored patterns is

360
$$x_n = x_{n-1} + a((m - x_{n-1})/m)$$

361 such that the percentage left untrained is $y_n = 100(1 - (x_n/m))$ where m is the total
362 number of parallel fibres which make synaptic contact on the target cell (175,000 on a
363 Purkinje cell and perhaps a thousand on a stellate cell) and a is the average number that
364 are active per pattern (a weighted average for stellate cells, derived from a probability
365 distribution). The value of m is an estimate derived from observation and $x_1 = ar$,
366 where $a = df$ and $r = m/w$. Here, d is the density of parallel fibre activity, f is dendritic
367 territory size, and w is the total number of parallel fibres intersecting a territory. Territory
368 'size' means area in the sagittal plane and the density of active parallel fibres is the
369 number per unit area. w and f are from observation, and d is an estimate. As expected,
370 we get the same result for all values of f and r – that is, the result is independent of cell
371 morphology, including depth-dependent morphology of interneurons [67, 69, 70].

372 There is no variable in the calculation for the number of times a single parallel fibre
373 makes contact on a stellate cell. Contact is assumed to be at a single synapse. This is
374 probable because the stellate cell dendritic territory is very highly flattened, barely wider
375 than dendritic diameter [71], and parallel fibres pass through at right angles. END BOX.

376

377 A random pattern of active inputs to a stellate cell may be to all working synapses, all
378 silent, or a mixture (with all silent the least likely). The probability that n active inputs

379 are to working synapses is the probability of n active inputs to any synapse (working or
 380 not, per Table 1) reduced by the probability that at least one input is to a silent synapse
 381 (so: the product of P_n and $1 - P(\text{none})$ which becomes $P_n(1 - 0.85^n)$) but increased by
 382 the sum of the probabilities that a higher total number of active inputs, a , is reduced
 383 by a number, b , to silent synapses, such that $a - b = n$. For example, the odds that
 384 there are two active inputs to working synapses is increased by the product of the
 385 probabilities that there are three active inputs in total, and any one is to a silent
 386 synapse. It is also increased by the product of the probabilities that there are four
 387 active inputs, and any two are to silent synapses, and so on. The probability of n active
 388 inputs at working synapses is accordingly given by:

$$389 \quad P_n - P_n(1 - 0.85^n) + \sum_{y=1}^z P_{n+y} \left[\frac{(n+y)!}{y!n!} * 0.15^y * (1 - 0.15)^n \right]$$

390 where P_n is the probability of n inputs (to any synapse, derived in Table 1) and $n + z$ is
 391 the maximum number of active inputs with more than insignificant odds (so around 12
 392 because higher numbers have a very low probability). This gives the Table 2
 393 probabilities.

394 **Table 2 The probability, p , that n concurrently active inputs are received at working**
 395 **synapses**

n	0	1	2	3	4	5	6	7	8	9	10	11	12
p	.0086	.0435	.1067	.1715	.2034	.1906	.1472	.0966	.0549	.0276	.0124	.005	.0019

396

397 The range and distribution of probabilities is very close to Table 1. That is, in any given
398 instance, the effect of learning on the number of active inputs to working synapses is
399 confined to a weak probability bias and marginal compared to the effect of uncertainty
400 at small numbers. Or: there is no effect of learning. Synaptic depression, though
401 profound, does not materially change the number of active inputs to working synapses.
402 Therefore, the learned adjustment of the balance between excitation and inhibition of
403 Purkinje cells is *exclusively controlled through the parallel fibre-Purkinje cell synapse*.

404 Note: At outer level, in mammals, there is evidence that stellate cells have good
405 memory for cutaneous receptive fields, but this seems to be limited to that level (Box
406 3).

407

408

BOX 3

409 **Stellate cells at outer level have good memory but only for a topographically narrow**
410 **source of signals.**

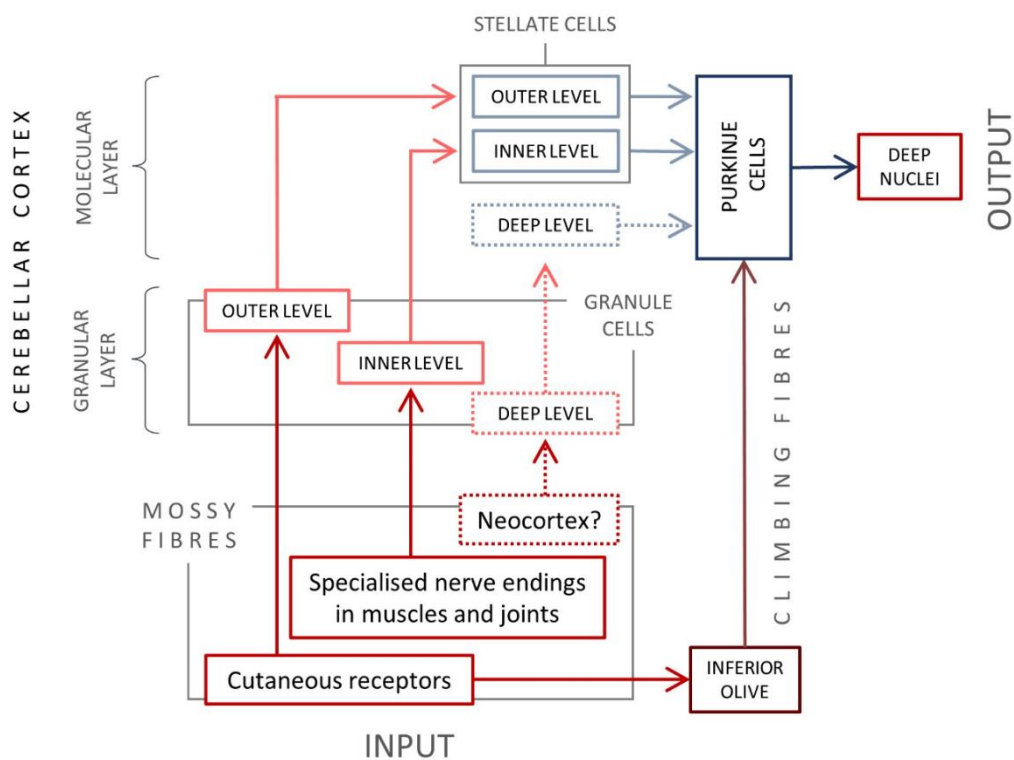
411 Both the granular layer and molecular layer of the cerebellar cortex contain vertical
412 organisation. Mossy fibre input to the granular layer is topographically stratified [45,
413 table 4]. There is 'a specific depth distribution of granule cells depending on the type of
414 input they received' [45, p.11795, 72] (C3 region, anaesthetised adult cats). Input to the

415 outer level of the granular layer is from cutaneous receptors, and at deeper level (the
416 largest stratum) is driven by passive forelimb movement and joint flexion, suggesting
417 that during behaviour it is driven by movement itself, by signals originating in muscles
418 and joints: proprioception. Topography is preserved in the molecular layer, with 'granule
419 cells in the inner granule cell layer giving rise to PFs [parallel fibres] in the inner molecular
420 layer and granule cells in the outer granule cell layer giving rise to PFs in the outer
421 molecular layer' [73, p.122]. Boundaries are faded, but this is the 'prevalent rule' [67,
422 p.66]. So, parallel fibre signals activity is modally stratified.

423 This has significance that relates to evidence that some stellate cells have good memory
424 for receptive fields. In a series of experiments, C3 stellate cells were reported to respond
425 (by firing) exclusively to stimulation of an associated discrete region of the body surface,
426 or receptive field, [9-11, 74], and not to stimulation of other fields. This is thought to
427 reflect synaptic plasticity at the parallel fibre-stellate cell synapse trained by cutaneous
428 stimulation, which evokes both mossy fibre and climbing fibre signals that travel by
429 different pathways to converge on the same strip of the cerebellar cortex [2, 62].
430 Climbing fibres terminate on Purkinje cells while mossy fibres terminate on granule cells
431 in the subjacent granular layer.

432 The interneurons used in these studies were activated by cutaneous stimulation,
433 effectively making it a selection requirement that they were outer-level cells, although
434 they were not expressly selected for their depth. So, stellate cells at outer level have

435 good memory but only for signals from a narrow topographical source, and only
 436 discriminate between signals from that source and from everywhere else. The special
 437 relationship of climbing fibre signals and mossy fibre signals is absent below outer level
 438 because mossy fibre signals are not evoked by the same stimulus. Therefore, receptive
 439 field memory is confined to outer level.



440

441 **Parallel fibre signals are modally stratified.** Reds: glutamate; blues: GABA.

442 END BOX.

443

444 9. What is learned calibration of the excitation-inhibition

445 balance for?

446 It is unknown why Purkinje cells receive simultaneous excitation and inhibition, both
447 coding granule cell rates. There are linear relationships of granule cell to Purkinje cell
448 firing (in slices with inhibition blocked) [17, 18], and granule cell to interneuron firing
449 (behaving animal) [19]. The granule cell → interneuron → Purkinje cell pathway drives
450 a fall in the firing rate of Purkinje cells as a linear function of granule cell firing rates
451 [20]. (Excitation was applied to mossy fibres, driving inhibition of laterally positioned
452 Purkinje cells which did not receive excitation.) There are linear relationships of the
453 balance between excitation and inhibition of Purkinje cells and depolarising or
454 hyperpolarising dendritic membrane voltage, and of membrane voltage and
455 bidirectional modulation of Purkinje cell simple spike output (behaving animal) [19].

456 To state that briefly, there is linear inhibition of Purkinje cells without excitation, and
457 linear excitation without inhibition. We moot that the net effect is modulation of
458 Purkinje cell firing which is a linear function of the sum of excitation and inhibition,
459 such that there is an inverse linear relationship of granule cell rates and Purkinje cell
460 firing which is calibrated by learning. Purkinje cells powerfully and linearly modulate
461 firing of excitatory projection neurons in the deep cerebellar nuclei, the output cells of
462 the cerebellum [75, 76].

463 However, the relationship of granule cell rates and Purkinje cell rates cannot be a
464 simple inversion. Purkinje cells fire spontaneously at robust rates [66, 77]
465 (approximately 60 Hz or 90 Hz [78], depending on whether they are in a zebrin-positive
466 or zebrin-negative zone), but well short of the higher end of their range during
467 behaviour. In walking rats, rates may exceed 200 Hz [79]. That is, at times, there is net
468 excitation of Purkinje cells, and at other times net inhibition, which *both code granule*
469 *cell rates*.

470 On the face of it, this is difficult to explain. But a possible explanation emerges if the
471 influence of disynaptic inhibition is weaker than direct excitation at low parallel fibre
472 rates, but there is a rate-dependent balance of influence such that, as parallel fibre
473 rates increase, inhibition becomes stronger at a faster rate than excitation. This has
474 support. Stellate cells form tonically self-inhibiting networks [67, 70]. Tonic inhibition
475 attenuates the effect of weak parallel fibre excitation [66] assisted by
476 endocannabinoid-mediated suppression of neurotransmitter release [80]. Weakening
477 competes for influence with enhancement of parallel fibre synaptic transmission to
478 stellate cells as a result of GABA spillover from stellate cell synapses [81, 82], acting
479 through GABA_A receptors which parallel fibres express extrasynaptically [83].
480 Enhancement is progressively strengthened by increasing parallel fibre rates, assisted
481 by positive feedback [84]. For more detail see Box 4.

482 Both pathways have a linear effect on Purkinje cell firing but with a rate-dependent
483 balance of influence, we suggest. In the lower part of the parallel fibre range, excitation
484 of stellate cells (and therefore feed-forward inhibition of Purkinje cells) is weakened
485 with the result that Purkinje cells receive net excitation. In the higher part of the
486 parallel fibre range, excitation of stellate cells (and therefore feed-forward inhibition of
487 Purkinje cells) is strengthened with the result that Purkinje cells receive weak net
488 excitation and ultimately net inhibition. The result is a larger Purkinje cell functional
489 range but a linearly inverted response to parallel fibre rates. The ratio of silent to
490 working synapses calibrates the balance.

491

492 **BOX 4**

493 **Weak excitation of interneurons is further weakened**

494 Stellate cells form tonically self-inhibiting networks [67, 70] (which include basket cells),
495 making contact on each other at the soma. Each cell receives contact from a low number
496 of neighbours, perhaps just one or two [66]. Stellate cells are very readily excited with
497 inhibition blocked [68], but in physiological conditions somatic hyperpolarisation causes
498 strong attenuation of passive current flow, reducing charge transfer from synapse to
499 soma [66]. Although a stellate cell receives contact from (many) more parallel fibres than
500 from other stellate cells, the number of concurrently active excitatory inputs is low

501 (single figures, Table 2) and excitatory contact is predominantly dendritic [85] so that
502 inhibition 'stands between' excitatory input and a somatic effect. Tonic inhibition
503 reduces the stellate cell dendritic membrane excitatory time constant. Summation is
504 therefore likely to be absent or 'severely affected' [i.e., much weakened: 66, p.675].
505 Elevation of postsynaptic intracellular calcium triggered by synaptic activation sums
506 linearly or supralinearly [80]. Calcium ion current density is too low to contribute to
507 depolarisation, but causes a branch-specific, endocannabinoid-mediated, short-term
508 suppression of neurotransmitter release [80].

509 **Strong excitation of interneurons is further strengthened**

510 GABA spillover from stellate cell synapses enhances parallel fibre synaptic transmission
511 to stellate cells [81, 82]. 'It is well established that parallel fibers express functional
512 GABA_ARs, that their activation depolarizes the axon, and that they increase release
513 probability at parallel fiber synapses' [83 p.16924]. Strong parallel fibre signals cause
514 more spillover, so that the effect of strong parallel fibre signals is strengthened. The
515 relationship is reciprocal, creating a positive feedback loop. 'GABA release from
516 molecular layer interneurons activates parallel fiber GABA_A receptors, and this, in turn,
517 increases release probability at synapses between parallel fibers and molecular layer
518 interneurons' [84 p.565]. END BOX.

519

520 **10. Discussion**

521 Cerebellar modelling has widely influenced network modelling. It was and remains an
522 assumption that information is represented by a replicable neural correlate, so that,
523 following training, signals are received at the correct synapses. In cerebellum models,
524 the key site of learning is the parallel fibre synapse, and training teaches a memory of
525 the permutation of active and inactive parallel fibres. However, Experimental evidence
526 has since contradicted the predictions of cerebellar models, failing to confirm the role
527 of LTD [12-14] predicted by cerebellar learning models, or the functional fine-grading
528 of synaptic weights [15, 16]. Also, for reasons stated at the beginning of the paper, it
529 may not be safe to assume that patterns are dependably replicated, *particularly* in the
530 cerebellum, which converts input to quasi-randomised internal activity [34, 44], so that
531 changes to input are amplified.

532 There is a growing body of evidence that communication in the cerebellum is linear,
533 i.e., firing rates are related by linear functions, including signalling at both the parallel
534 fibre-Purkinje cell [17, 18] and parallel fibre-interneuron [19, 20] synapses, and
535 generally that cerebellar synaptic transmission is highly adapted for high fidelity
536 transmission of rate information [25-31], calling into question the traditionally
537 envisaged function of synaptic memory.

538 One-time code (where the representation of information becomes independent of the
539 form of neural activity) has traditionally snagged on the conceptual problem that it is
540 not obvious how neural activity would represent information, or how learning could
541 have a useful outcome. In the present proposal, the learning outcome is not dependent
542 on replication. In fact, it may not be necessary or meaningful to think in terms of
543 patterns at all. Training teaches the same outcome even if the permutation of active
544 parallel fibres is never repeated. Training polarises synaptic weights, removing
545 interference by synaptic noise with linear communication. It is a key function of the
546 cerebellar (and any) network to eliminate or negate redundant variables (i.e., any
547 effect that codes nothing). In a linear system, that includes variable synaptic
548 transmission strength, because it displaces faithful transmission of rate information. In
549 this view, it is not merely plausible but necessary that pattern discrimination is absent
550 from the learned response.

551 The cerebellum is thought to be modular, made up of networks with the same
552 repeating basic wiring, which is well preserved across species [86, 87]. Learned
553 calibration means the modular cerebellar network can work blind, without regard to
554 the source or nature of inputs – i.e., without specialising – precisely because it does
555 *not* discriminate. As a result, the same standard unit has been able to support the
556 diversification of vertebrate body plans.

557

558 **Grants**

559 Funding: This work was supported by the Leverhulme Trust [grant number ECF-2022-
560 079] (Mike Gilbert); the Swedish Research Council [grant number 2020-01468]; the Per-
561 Eric och Ulla Schybergs Foundation [grant number 42630]; and the Crafoord
562 Foundation [grant numbers 20180704, 20200729, 20220776, & 20230655] (Anders
563 Rasmussen).

564 **Author contributions**

565 M.G. Researched, conceived and developed the ideas, performed the analysis, wrote
566 the code, prepared the figures and wrote the manuscript. A.R. provided expertise and
567 feedback, and reviewed and provided editorial comments on the manuscript.

568 **Declaration of interests**

569 The authors declare no competing interests.

570

571 **REFERENCES**

572

- 573 1. Oscarsson, O., *Functional units of the cerebellum - sagittal zones and microzones.*
 574 Trends in Neurosciences, 1979. **2**: p. 143-145.
- 575 2. Garwicz, M., C.F. Ekerot, and H. Jörntell, *Organizational principles of cerebellar*
 576 *neuronal circuitry.* News in Physiological Sciences, 1998. **13**: p. 28-34.
- 577 3. Hansel, C., D.J. Linden, and E. D'Angelo, *Beyond parallel fiber LTD: the diversity of*
 578 *synaptic and non-synaptic plasticity in the cerebellum.* Nat Neurosci, 2001. **4**(5): p.
 579 467-75.
- 580 4. Qiu, D.L. and T. Knopfel, *Presynaptically expressed long-term depression at*
 581 *cerebellar parallel fiber synapses.* Pflugers Arch, 2009. **457**(4): p. 865-75.
- 582 5. Ito, M. and M. Kano, *Long-lasting depression of parallel fiber-Purkinje cell*
 583 *transmission induced by conjunctive stimulation of parallel fibers and climbing*
 584 *fibers in the cerebellar cortex.* Neurosci Lett, 1982. **33**(3): p. 253-8.
- 585 6. Smith, S.L. and T.S. Otis, *Pattern-dependent, simultaneous plasticity differentially*
 586 *transforms the input-output relationship of a feedforward circuit.* Proc Natl Acad
 587 Sci U S A, 2005. **102**(41): p. 14901-6.
- 588 7. Soler-Llavina, G.J. and B.L. Sabatini, *Synapse-specific plasticity and*
 589 *compartmentalized signaling in cerebellar stellate cells.* Nat Neurosci, 2006. **9**(6):
 590 p. 798-806.
- 591 8. Rancillac, A. and F. Crépel, *Synapses between parallel fibres and stellate cells*
 592 *express long-term changes in synaptic efficacy in rat cerebellum.* J Physiol, 2004.
 593 **554**(Pt 3): p. 707-20.
- 594 9. Jörntell, H. and C.F. Ekerot, *Reciprocal bidirectional plasticity of parallel fiber*
 595 *receptive fields in cerebellar Purkinje cells and their afferent interneurons.* Neuron,
 596 2002. **34**(5): p. 797-806.
- 597 10. Jörntell, H. and C.F. Ekerot, *Receptive field plasticity profoundly alters the*
 598 *cutaneous parallel fiber synaptic input to cerebellar interneurons in vivo.* J
 599 Neurosci, 2003. **23**(29): p. 9620-31.
- 600 11. Ekerot, C.F. and H. Jörntell, *Parallel fiber receptive fields: a key to understanding*
 601 *cerebellar operation and learning.* Cerebellum, 2003. **2**(2): p. 101-9.
- 602 12. Welsh, J.P., et al., *Normal motor learning during pharmacological prevention of*
 603 *Purkinje cell long-term depression.* Proc Natl Acad Sci U S A, 2005. **102**(47): p.
 604 17166-71.
- 605 13. Schonewille, M., et al., *Reevaluating the role of LTD in cerebellar motor learning.*
 606 Neuron, 2011. **70**(1): p. 43-50.
- 607 14. ten Brinke, M.M., et al., *Evolving Models of Pavlovian Conditioning: Cerebellar*
 608 *Cortical Dynamics in Awake Behaving Mice.* Cell Rep, 2015. **13**(9): p. 1977-88.
- 609 15. Ekerot, C.F. and H. Jörntell, *Parallel fibre receptive fields of Purkinje cells and*
 610 *interneurons are climbing fibre-specific.* Eur J Neurosci, 2001. **13**(7): p. 1303-10.
- 611 16. Isope, P. and B. Barbour, *Properties of unitary granule cell-->Purkinje cell synapses*
 612 *in adult rat cerebellar slices.* J Neurosci, 2002. **22**(22): p. 9668-78.
- 613 17. Walter, J.T. and K. Khodakhah, *The linear computational algorithm of cerebellar*
 614 *Purkinje cells.* J Neurosci, 2006. **26**(50): p. 12861-72.

- 615 18. Walter, J.T. and K. Khodakhah, *The advantages of linear information processing for*
616 *cerebellar computation*. Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4471-6.
- 617 19. Jelitai, M., et al., *Dendritic excitation-inhibition balance shapes cerebellar output*
618 *during motor behaviour*. Nat Commun, 2016. **7**: p. 13722.
- 619 20. Park, S.M., E. Tara, and K. Khodakhah, *Efficient generation of reciprocal signals by*
620 *inhibition*. J Neurophysiol, 2012. **107**(9): p. 2453-62.
- 621 21. Rancz, E.A., et al., *High-fidelity transmission of sensory information by single*
622 *cerebellar mossy fibre boutons*. Nature, 2007. **450**(7173): p. 1245-8.
- 623 22. Ritzau-Jost, A., et al., *Ultrafast action potentials mediate kilohertz signaling at a*
624 *central synapse*. Neuron, 2014. **84**(1): p. 152-163.
- 625 23. Duguid, I., et al., *Control of cerebellar granule cell output by sensory-evoked Golgi*
626 *cell inhibition*. Proc Natl Acad Sci U S A, 2015. **112**(42): p. 13099-104.
- 627 24. Duguid, I., et al., *Tonic inhibition enhances fidelity of sensory information*
628 *transmission in the cerebellar cortex*. J Neurosci, 2012. **32**(32): p. 11132-43.
- 629 25. Delvendahl, I. and S. Hallermann, *The Cerebellar Mossy Fiber Synapse as a Model*
630 *for High-Frequency Transmission in the Mammalian CNS*. Trends Neurosci, 2016.
631 **39**(11): p. 722-737.
- 632 26. Saviane, C. and R.A. Silver, *Fast vesicle reloading and a large pool sustain high*
633 *bandwidth transmission at a central synapse*. Nature, 2006. **439**(7079): p. 983-7.
- 634 27. Cesana, E., et al. *Novel granule cell-Golgi cell excitatory input in the cerebellar*
635 *granular layer*. in FENS Abstr. 2010.
- 636 28. Moore, J.W. and D.E.J. Blazis, *11 - Simulation of a Classically Conditioned*
637 *Response: A Cerebellar Neural Network Implementation of the Sutton–Barto–*
638 *Desmond Model*, in *Neural Models of Plasticity*, J.H. Byrne and W.O. Berry, Editors.
639 1989, Academic Press. p. 187-207.
- 640 29. Sargent, P.B., et al., *Rapid vesicular release, quantal variability, and spillover*
641 *contribute to the precision and reliability of transmission at a glomerular synapse*. J
642 Neurosci, 2005. **25**(36): p. 8173-87.
- 643 30. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptic Specializations Support*
644 *Frequency-Independent Purkinje Cell Output from the Cerebellar Cortex*. Cell
645 Reports, 2016. **17**(12): p. 12.
- 646 31. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptotagmin 7 confers frequency*
647 *invariance onto specialized depressing synapses*. Nature, 2017. **551**: p. 503.
- 648 32. Raymond, J.L. and J.F. Medina, *Computational Principles of Supervised Learning in*
649 *the Cerebellum*. Annu Rev Neurosci, 2018. **41**: p. 233-253.
- 650 33. Cayco-Gajic, N.A. and R.A. Silver, *Re-evaluating Circuit Mechanisms Underlying*
651 *Pattern Separation*. Neuron, 2019. **101**(4): p. 584-602.
- 652 34. Cayco-Gajic, N.A., C. Clopath, and R.A. Silver, *Sparse synaptic connectivity is*
653 *required for decorrelation and pattern separation in feedforward networks*. Nat
654 Commun, 2017. **8**(1): p. 1116.
- 655 35. Salin, P.A., R.C. Malenka, and R.A. Nicoll, *Cyclic AMP mediates a presynaptic form*
656 *of LTP at cerebellar parallel fiber synapses*. Neuron, 1996. **16**(4): p. 797-803.
- 657 36. Chen, C. and W.G. Regehr, *The mechanism of cAMP-mediated enhancement at a*
658 *cerebellar synapse*. J Neurosci, 1997. **17**(22): p. 8687-94.

- 659 37. Lev-Ram, V., et al., *A new form of cerebellar long-term potentiation is postsynaptic*
660 *and depends on nitric oxide but not cAMP*. Proc Natl Acad Sci U S A, 2002. **99**(12):
661 p. 8389-93.
- 662 38. Coesmans, M., et al., *Bidirectional parallel fiber plasticity in the cerebellum under*
663 *climbing fiber control*. Neuron, 2004. **44**(4): p. 691-700.
- 664 39. Jörntell, H. and C. Hansel, *Synaptic memories upside down: bidirectional plasticity*
665 *at cerebellar parallel fiber-Purkinje cell synapses*. Neuron, 2006. **52**(2): p. 227-38.
- 666 40. Masugi-Tokita, M., et al., *Number and density of AMPA receptors in individual*
667 *synapses in the rat cerebellum as revealed by SDS-digested freeze-fracture replica*
668 *labeling*. J Neurosci, 2007. **27**(8): p. 2135-44.
- 669 41. Szapiro, G. and B. Barbour, *Multiple climbing fibers signal to molecular layer*
670 *interneurons exclusively via glutamate spillover*. Nat Neurosci, 2007. **10**(6): p. 735-
671 42.
- 672 42. Clark, B.A. and S.G. Cull-Candy, *Activity-dependent recruitment of extrasynaptic*
673 *NMDA receptor activation at an AMPA receptor-only synapse*. J Neurosci, 2002.
674 **22**(11): p. 4428-36.
- 675 43. Harvey, R.J. and R.M. Napper, *Quantitative study of granule and Purkinje cells in*
676 *the cerebellar cortex of the rat*. J Comp Neurol, 1988. **274**(2): p. 151-7.
- 677 44. Albus, J.S., *A Theory of Cerebellar Function*. Mathematical Biosciences, 1971.
678 **10**(1-2): p. 25-61.
- 679 45. Jörntell, H. and C.F. Ekerot, *Properties of somatosensory synaptic integration in*
680 *cerebellar granule cells in vivo*. J Neurosci, 2006. **26**(45): p. 11786-97.
- 681 46. van Beugen, B.J., et al., *High Frequency Burst Firing of Granule Cells Ensures*
682 *Transmission at the Parallel Fiber to Purkinje Cell Synapse at the Cost of Temporal*
683 *Coding*. Frontiers in Neural Circuits, 2013. **7**: p. 95.
- 684 47. Sultan, F., *Exploring a critical parameter of timing in the mouse cerebellar*
685 *microcircuitry: the parallel fiber diameter*. Neurosci Lett, 2000. **280**(1): p. 41-4.
- 686 48. Wang, S.S., W. Denk, and M. Hausser, *Coincidence detection in single dendritic*
687 *spines mediated by calcium release*. Nat Neurosci, 2000. **3**(12): p. 1266-73.
- 688 49. Suvrathan, A., H.L. Payne, and J.L. Raymond, *Timing Rules for Synaptic Plasticity*
689 *Matched to Behavioral Function*. Neuron, 2016. **92**(5): p. 959-967.
- 690 50. Malinow, R. and R.C. Malenka, *AMPA receptor trafficking and synaptic plasticity*.
691 Annu Rev Neurosci, 2002. **25**: p. 103-26.
- 692 51. Hesslow, G. and M. Ivarsson, *Inhibition of the inferior olive during conditioned*
693 *responses in the decerebrate ferret*. Exp Brain Res, 1996. **110**(1): p. 36-46.
- 694 52. Bengtsson, F., P. Svensson, and G. Hesslow, *Feedback control of Purkinje cell*
695 *activity by the cerebello-olivary pathway*. Eur J Neurosci, 2004. **20**(11): p. 2999-
696 3005.
- 697 53. Bengtsson, F. and G. Hesslow, *Cerebellar control of the inferior olive*. Cerebellum,
698 2006. **5**(1): p. 7-14.
- 699 54. Mathy, A., et al., *Encoding of oscillations by axonal bursts in inferior olive neurons*.
700 Neuron, 2009. **62**(3): p. 388-99.
- 701 55. Bazzigaluppi, P., et al., *Olivary subthreshold oscillations and burst activity*
702 *revisited*. Front Neural Circuits, 2012. **6**: p. 91.

- 703 56. Gilbert, M. and R.C. Miall, *Gating by Functionally Indivisible Cerebellar Circuits: a*
704 *Hypothesis*. *The Cerebellum*, 2021b. **20**(4): p. 518-532.
- 705 57. Crill, W.E., *Unitary multiple-spiked responses in cat inferior olive nucleus*. *J*
706 *Neurophysiol*, 1970. **33**(2): p. 199-209.
- 707 58. Eccles, J.C., R. Llinas, and K. Sasaki, *The excitatory synaptic action of climbing*
708 *fibres on the Purkinje cells of the cerebellum*. *J Physiol*, 1966. **182**(2): p. 268-96.
- 709 59. Öhman, J., et al., *The Effect of Nucleo-Olivary Stimulation on Climbing Fiber EPSPs*
710 *in Purkinje Cells*. *Cerebellum*, 2024.
- 711 60. Gilbert, M., *Gating by Memory: a Theory of Learning in the Cerebellum*. *The*
712 *Cerebellum*, 2021. DOI: **10.1007/s12311-021-01325-9**.
- 713 61. Gilbert, M. and A. Rasmussen, *The cerebellum converts input data into hyper low-*
714 *resolution 3-D pixels*. *Who knew? bioRxiv*, 2024: p. DOI:
715 10.1101/2023.07.14.548987 version 8,
716 <https://www.biorxiv.org/content/10.1101/2023.07.14.548987v8>.
- 717 62. Garwicz, M., H. Jörntell, and C.F. Ekerot, *Cutaneous receptive fields and*
718 *topography of mossy fibres and climbing fibres projecting to cat cerebellar C3*
719 *zone*. *J Physiol*, 1998. **512 (Pt 1)**: p. 277-93.
- 720 63. De Gruijl, J.R., et al., *Inferior Olive: All Ins and Outs*, in *Handbook of the*
721 *Cerebellum and Cerebellar Disorders*, S.J.D. Manto M., Rossi F., Gruol D.L.,
722 Koibuchi N., Editor. 2012, Springer. p. 1013-1058.
- 723 64. Ozden, I., et al., *Reliable coding emerges from coactivation of climbing fibers in*
724 *microbands of cerebellar Purkinje neurons*. *J Neurosci*, 2009. **29**(34): p. 10463-73.
- 725 65. Ju, C., et al., *Neurons of the inferior olive respond to broad classes of sensory input*
726 *while subject to homeostatic control*. *J Physiol*, 2019. **597**(9): p. 2483-2514.
- 727 66. Häusser, M. and B.A. Clark, *Tonic synaptic inhibition modulates neuronal output*
728 *pattern and spatiotemporal synaptic integration*. *Neuron*, 1997. **19**(3): p. 665-78.
- 729 67. Palay, S.L. and V. Chan-Palay, *Cerebellar cortex : cytology and organization*. 1974,
730 Berlin: Springer.
- 731 68. Carter, A.G. and W.G. Regehr, *Quantal events shape cerebellar interneuron firing*.
732 *Nat Neurosci*, 2002. **5**(12): p. 1309-18.
- 733 69. Paula-Barbosa, M.M., et al., *The distribution of stellate cell descending axons in*
734 *the rat cerebellum: a Golgi and a combined Golgi-electron microscopical study*. *J*
735 *Anat*, 1983. **137 (Pt 4)**: p. 757-64.
- 736 70. Sultan, F. and J.M. Bower, *Quantitative Golgi study of the rat cerebellar molecular*
737 *layer interneurons using principal component analysis*. *J Comp Neurol*, 1998.
738 **393**(3): p. 353-73.
- 739 71. Rizza, M.F., et al., *Stellate cell computational modeling predicts signal filtering in*
740 *the molecular layer circuit of cerebellum*. *Scientific Reports*, 2021. **11**(1): p. 3873.
- 741 72. Quy, P.N., et al., *Projection patterns of single mossy fiber axons originating from the*
742 *dorsal column nuclei mapped on the aldolase C compartments in the rat*
743 *cerebellar cortex*. *J Comp Neurol*, 2011. **519**(5): p. 874-99.
- 744 73. Zhang, W. and D.J. Linden, *Calcium influx measured at single presynaptic boutons*
745 *of cerebellar granule cell ascending axons and parallel fibers*. *Cerebellum*, 2012.
746 **11**(1): p. 121-31.

- 747 74. Jörntell, H. and C.F. Ekerot, *Receptive Field Remodeling Induced by Skin*
748 *Stimulation in Cerebellar Neurons in vivo*. Front Neural Circuits, 2011. **5**: p. 3.
- 749 75. Pedroarena, C.M. and C. Schwarz, *Efficacy and short-term plasticity at GABAergic*
750 *synapses between Purkinje and cerebellar nuclei neurons*. J Neurophysiol, 2003.
751 **89**(2): p. 704-15.
- 752 76. Abbasi, S., et al., *Robust transmission of rate coding in the inhibitory Purkinje cell*
753 *to cerebellar nuclei pathway in awake mice*. PLoS computational biology, 2017.
754 **13**(6): p. e1005578-e1005578.
- 755 77. Raman, I.M. and B.P. Bean, *Ionic currents underlying spontaneous action*
756 *potentials in isolated cerebellar Purkinje neurons*. J Neurosci, 1999. **19**(5): p. 1663-
757 74.
- 758 78. Zhou, H., et al., *Cerebellar modules operate at different frequencies*. Elife, 2014. **3**:
759 p. e02536.
- 760 79. Sauerbrei, B.A., E.V. Lubenov, and A.G. Siapas, *Structured Variability in Purkinje*
761 *Cell Activity during Locomotion*. Neuron, 2015. **87**(4): p. 840-52.
- 762 80. Tran-Van-Minh, A., et al., *Differential Dendritic Integration of Synaptic Potentials*
763 *and Calcium in Cerebellar Interneurons*. Neuron, 2016. **91**(4): p. 837-850.
- 764 81. Stell, B.M., et al., *Activation of presynaptic GABA(A) receptors induces glutamate*
765 *release from parallel fiber synapses*. J Neurosci, 2007. **27**(34): p. 9022-31.
- 766 82. Dellal, S.S., R. Luo, and T.S. Otis, *GABAA receptors increase excitability and*
767 *conduction velocity of cerebellar parallel fiber axons*. J Neurophysiol, 2012.
768 **107**(11): p. 2958-70.
- 769 83. Pugh, J.R. and C.E. Jahr, *Activation of Axonal Receptors by GABA Spillover*
770 *Increases Somatic Firing*. The Journal of Neuroscience, 2013. **33**(43): p. 16924-
771 16929.
- 772 84. Pugh, J.R. and C.E. Jahr, *Axonal GABAA receptors increase cerebellar granule cell*
773 *excitability and synaptic activity*. J Neurosci, 2011. **31**(2): p. 565-74.
- 774 85. Lemkey-Johnston, N. and L.M. Larramendi, *Types and distribution of synapses*
775 *upon basket and stellate cells of the mouse cerebellum: an electron microscopic*
776 *study*. J Comp Neurol, 1968. **134**(1): p. 73-112.
- 777 86. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from*
778 *Myxinooids through Birds*. 1967: The University of Minnesota Press.
- 779 87. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from*
780 *Monotremes through Apes*. 1970: The University of Minnesota Press, Minneapolis.

781