Computational anatomy: the cerebellar microzone computation, a hypothesis

Mike Gilbert[1] *, Anders Rasmussen[2]

[1]School of Psychology, College of Life and Environmental Sciences, University of Birmingham, Birmingham B15 2TT, UK ORCID: 0000-0003-4415-6915

[2]Department of Experimental Medical Science, Lund University, BMC F10, 22184 Lund, Sweden ORCID: 0000-0003-4477-369X

*Correspondence: m.gilbert.1@bham.ac.uk;

Abstract

The cerebellum is a large brain structure. It is poorly understood how the cerebellum codes and processes information. It is generally assumed that the representation of information in the brain is in high resolution and replicable with high fidelity, and that those features make it precise and characterise the inputs to sophisticated computations. Replicability is an important assumption of neural network learning models, including cerebellar models, which depend on replicability because they rely on pattern memory. We challenge those ideas. We argue that the cerebellar microzone computation is the aggregate of simple unit computations and a passive effect of anatomy, unaided and unlearned, which we are able to recreate *in silico* in the form of a network of unit functions. Unit functions are attributed to identified anatomical hardware. The functions themselves are linear, with evidence. In this form, we are able to model large cell networks in high relevant definition, without simplifying or omitting parameters that affect the computation. Conversion of inputs to output of the system is mediated by a chain of rate codes, and not supplanted by learning. Unit populations are organised functionally in layers that span the network. A learning algorithm and replicability are unnecessary to explain the evidence.

Keywords cerebellum; theory; model; computation; anatomy; microzone; network; code

Lay summary

It is poorly understood – i.e., not understood – how the cerebellum represents information and uses it to generate motor outputs, and how outputs are coordinated with each other and with behaviour to control movement. We try to answer those questions.

The paper is in two main parts. In the first, we model the cerebellar network computation, and in particular the microzone computation, the middle layer of the network. We propose the computation is unlearned, simply a passive effect of cell morphologies and network geometry. This is still sufficient to distil a single output rate – motor output of the cerebellum – from 80,000 input signals.

In the second part, we propose that the topography of input and output of cerebellar motor circuits could automate unconscious control of motor sequences. To illustrate how that could work, we model swimming. We propose that the cerebellum connects proprioception-motor loops so that movement itself can drive coordinated motor output with only sparse (go, no go) executive input. The role of the cerebellum and functional topography of connections may remain today a variation of the same principles.

1. Introduction

Most of the mass and volume of the cerebellum is made up of the cerebellar cortex, which is divided into an inner and outer layer, the granular layer and molecular layer. The molecular layer is divided functionally into long, thin strips called microzones [1-3]. Microzones lie at right angles to the main portion of the axons of granule cells, called parallel fibres, which themselves lie parallel to each other and the cerebellar surface. A microzone is populated by a few hundred Purkinje cells. Purkinje cells receive contact in passing from parallel fibres and are the sole outputs of the cerebellar cortex (Fig 1). Firing rates of the output cells of the cerebellum are bidirectionally sensitive to Purkinje cell rates [4]. This provides the main control of output cell rates.

Neuronal firings rates are thought to represent values received as inputs to network computations. We present the simple idea that there is a passive, unlearned computational effect of microzone anatomy. It is contentious partly because it conflicts with traditional learning models, and partly because it depends on linear communication between cells, for which there is significant but incomplete evidence.

Do we need a rethink? We believe probably yes, partly because the old view has not aged well and partly to account for new evidence. In the classic cerebellar learning model [5-7], the chief instrument of the network computation is the learned

modification of parallel fibre synaptic weights. The 'purpose' is that remembered patterns – i.e., received by training-modified synapses – return a learned response. The learned response displaces the naive response to unprocessed input rates. However, experimental confirmation of the predicted role of synaptic modification in motor learning did not materialise [8-11], and naturally-occurring grades of synaptic depression were not found [12-14].

There is a significant body of evidence that rate information is linearly conserved in signalling between cerebellar neurons [15-18], including synaptic signalling in the molecular layer [19-22], and that cells are adapted for faithful transmission [23-29]. Firing rates linearly code behavioural and task metrics 'at all levels of the cerebellar circuit' [30 p.239, citing 30 references]. Neurobiology is highly adapted to control for biophysical and other noise. There are cerebellar examples of exactly this, where cell-to-cell communication is adapted to control for variables that are not code, i.e., to conserve fidelity of linear transmission of rate information [15-18, 23, 24, 27-29, 31-33].

Perhaps surprisingly, it is not obvious how to account for linear signalling as part of a computational mechanism. Also, it is probably still the mainstream view that synaptic weights will be involved somehow. We propose here a mechanism of the microzone computation in which anatomy has an unlearned computational role. This is passive, simply a result of cell morphologies and network geometry. Nonetheless, it is able,

plausibly, to run uninterrupted calculations in parallel, in real time, and doesn't need synaptic weights.

We propose – as artificial neural network models do – that signals traffic flows through computational units organised in layers, but with the difference that:

1) Our units are matched with physiological hardware (and physical counterparts are not synapses);

2) Units form functional populations whose size and spatial dimensions are a consequence of known physiology; and

3) Unit functions are not conjecture, they are linear, with the basis in evidence of linear signalling, cited in context.

Note that we do not propose to be mathematically interesting. The heart of the proposal is that the physiological computation is in fact not sophisticated, against expectation – at least, there is a plausible argument that this is the case, which we present as an alternative to current models. For that reason, physiology has a high profile in the paper. The sophistication of the biological design is found in the fact that physiology is adapted to isolate computationally relevant parameters (and computations themselves) from interference by other variables.

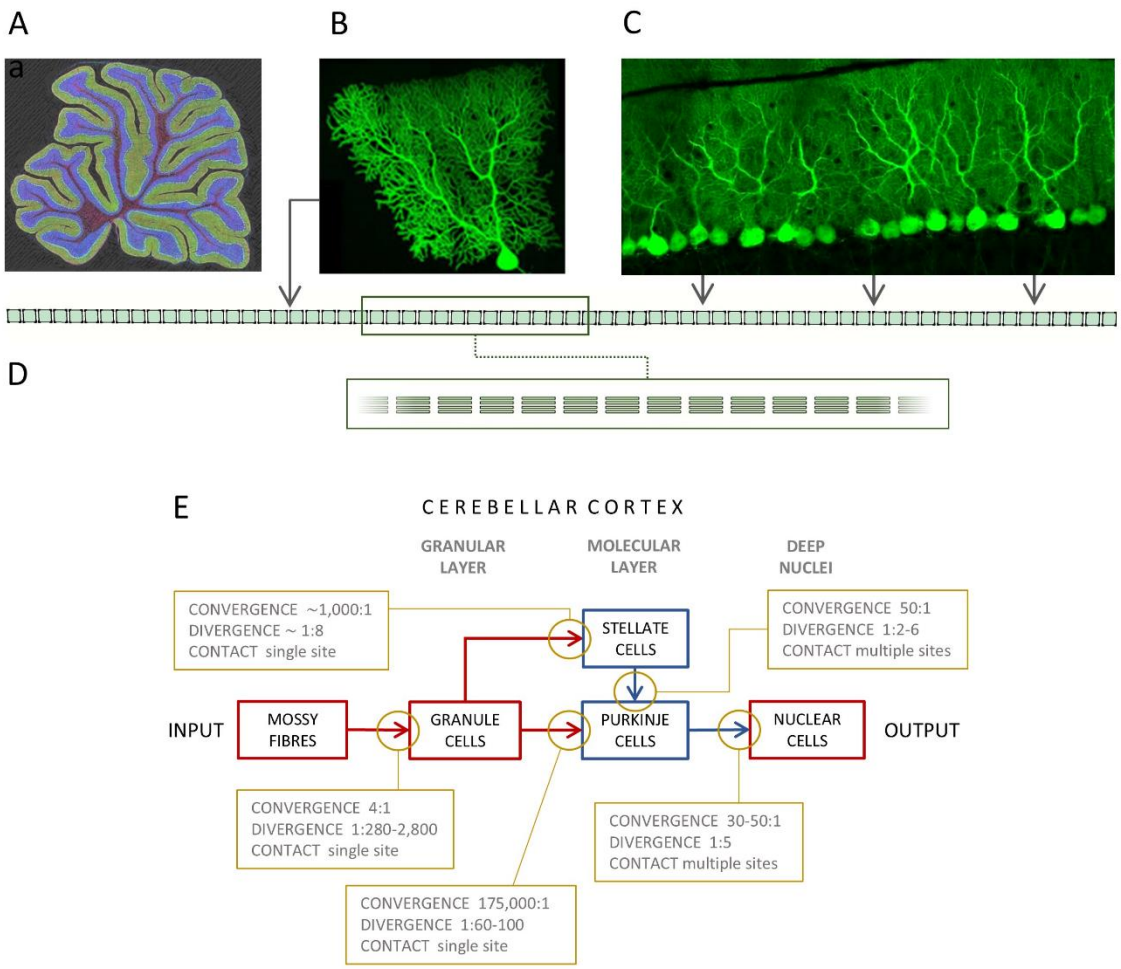


Fig 1. Schematic of data flow in the cerebellum.

A Sagittal section of the rat cerebellum (obtained by multi-photon microscopy, courtesy Tom Deerinck). The outer layer, the molecular layer, of the cerebellar cortex is stained green and the inner layer, the granular layer, is blue. **B** and **C** Purkinje cells in a sagittal slice (two-photon laser scanner micrograph, courtesy Mike Häusser). The Purkinje cell arbour is large from this view, filling the molecular layer vertically. It is flattened in the plane orthogonal to parallel fibres. **D** Top. Schematic of a parallel fibre view (i.e., the

same view as a–c) of a row of 80 Purkinje cell territories, but straightened out. The molecular layer is divided functionally into long, thin strips called microzones. At a mid-range estimate, a microzone is 75–100 Purkinje cells long and 4 wide (measuring 1,500–2,000 x 150 µm). Inset. A short section viewed from the cerebellar surface. **E** Simplified circuit wiring, highly preserved across vertebrates. Red: glutamate; blue: GABA. Brown boxes contain information about network connections. Microzones lie at right angles to parallel fibres. Tens of millions of parallel fibres pass through a microzone, making contact in passing (if at all) on Purkinje cells and interneurons. END CAPTION

2. Methods

2.1 The principles of the computation

The cerebellar network (as distinct from organisation into microcircuits that include a cell group in the inferior olive) is formed of a microzone (the middle layer), the region of the granular layer which supplies parallel fibre input to it (the input layer), and a cell group in the deep cerebellar nuclei which includes the output cells of the network (which are also the output cells of the cerebellum, the outer layer) (Fig 2). For non-cerebellar physiologists, a layer of a network has a different meaning to the same word used in granular layer and molecular layer. For example, a microzone, the middle layer

of a network, is a division of the molecular layer, a stratum of the cerebellar cortex. Each successive network layer – input, middle, output – contains a smaller number of cells than the last. The input layer contains millions. The number of output cells may be around 50.

The cerebellum, we argue, exploits statistical effects of taking random samples. Say a population of values (with any frequency distribution) is randomly sampled. Samples are independent, sometimes called sampling with replacement – an illustration sometimes used is that coloured balls taken from a bag are replaced in the bag between samplings. There are *n* samples, sample size *m*. If we take the mean (or sum) of each sample, and *n* is large, the sample means have a normal or near-normal distribution, by the central limit theorem. (This works even if *m* is a range.) The new distribution is centred on the same mean as the sampled distribution, but narrower.

Large values of *n* and *m* give a precisely-shaped new distribution. Modest values give an untidier (but still normal-tending) shape. However, if the new distribution is then itself independently randomly sampled, and the process is repeated, the results nonetheless become more focused, and the relationship of the means of successive distributions is well preserved. We will argue that the cerebellar microzone provides (part of) the physical hardware for the cerebellar network to distil output from input in this way, noting that at each step the relationship of the means includes a constant.

2.2 The relationship of physiology and model parameters

In this contention, the neurophysiology of a microzone is the physical form of data organised in layers. A data layer spans a network, so that it extends from end to end of a microzone. Data are computed by units. In its computational form, a unit is a linear function. In its physical form it is the component (a cell or part of a cell) that executes the unit function and returns a value. The biophysical form of a value depends on the layer. A unit randomly samples the afferent data layer. During behaviour, units receive a constant turnover of inputs and compute ceaselessly, so that the value coded by a unit is an analogue signal.

This is a third use of the word layer, again with a different meaning. Here, a layer is a layer of computations. Units in their physical form themselves form several layers in a single microzone, but that is immaterial to the computation. A microzone is the computational equivalent of a plane at right angles to the cerebellar surface and parallel fibres, where all cells sample the same convergent parallel fibre signals.

We argue that the microzone computation is the passive effect of anatomy which we are able to recreate *in silico* in the form of a network of unit functions. Functions are attributed to identified anatomical hardware and the functions themselves are evidenced. The evidence is discussed in context, in the main text. We simulate a moment in time, and more exactly the computation performed on a very short period

173 of inputs. In this form, we are able to model large cell networks in high relevant
174 definition, without simplifying assumptions. Put another way, if we are correct about
175 the physical form of units, and we are also correct that unit functions are linear and
176 about the physical dimensions of data layers, the model includes all computationally
177 relevant parameters. The sophistication of the system is biological, not computational.
178 It lies in the fact (as noted) that neurophysiology is adapted to exclude an effect of
179 other variables, including interference by (a redundant effect of) biophysical variables
180 which execute transmission of signals.

181 The simulation provides a demonstration in principle that processing by anatomy
182 provides a mechanism that:

183 1) Autonomously strictly regulates the fraction of active parallel fibres that converge at
184 any time on a microzone; and

185 2) Converts mossy fibre firing rates received by the input layer of a network into well
186 predicted coherent firing of middle-layer, microzone-grouped Purkinje cells.

187 Units are connected by known anatomy. Information about anatomy is detailed and
188 not contentious. There is some variation between species, and indeed between
189 estimates, but network architecture is strongly conserved, suggesting that variation
190 does not affect the form of the computation. The number of active inputs to a unit is in
191 some cases randomly variable with time. However, we can simulate that by, for each

192 unit, selecting the number at random from the reported range (with a distributed
193 probability if we know what that is).

194 It is unnecessary to know which individual cells contact each other because
195 (biologically and in the model) contact is at random inside microzone sectors defined
196 by a combination of topography and scaffold-like architecture of the cerebellar cortex.
197 Biologically, a sector is a volume with spatial dimensions. *In silico*, we can represent
198 that as a population of unit functions. We know from anatomy (or can derive) the size
199 (and shape) of unit populations, population ratios, and convergence and divergence
200 ratios.

201

202 2.3 Assumptions

203 2.3.1 THE GRANULE CELL CODE

204 We have previously simulated the granular layer mechanism which converts mossy
205 fibre signals into granule cell and therefore parallel fibre signals [34]. The granular
206 layer methodology and a review of the evidence appears in Supplemental Materials.
207 We use the granular layer mechanism to generate parallel fibre input to the microzone
208 computation. Previously, we modelled a sagittal strip of the granular layer. In this
209 paper, we extend that to the larger region (around 40 strips) that supplies parallel fibre

input to the microzone that equally bisects the region viewed from the cerebellar surface – i.e., the input and middle layers of a network, respectively. The outputs of the input layer mechanism are the homeostatic regulation of the fraction of active input layer granule cells, and therefore the fraction of active parallel fibres which converge on the central microzone, and the parallel fibre code.

In this paper, the workings of the granular layer mechanism are effectively an assumption. The parallel fibre code has invariant features, regardless of the information it contains. Information is coded at collective level, in the joint activity of cells which are active at the same time. More exactly, it is coded in the frequency distribution of the rates at which those cells fire, and more exactly still this reflects the status of cells in a moment of time where they intersect an orthogonal plane. Therefore, it is coded in the frequency distribution of active inputs to the central microzone.

Firing rates are randomly distributed among active parallel fibres and active parallel fibres are a random subset of the general population. Accordingly, any random sample of active inputs is (at any moment) a random sample of the same distribution, and because all sectors of a microzone receive a large number of active inputs at any time, all sectors receive the same code, as one. In this paper, we treat the form of the parallel fibre code as an assumption. The reader may, if they wish, read the granular layer paper, but if not, they are not at a disadvantage.

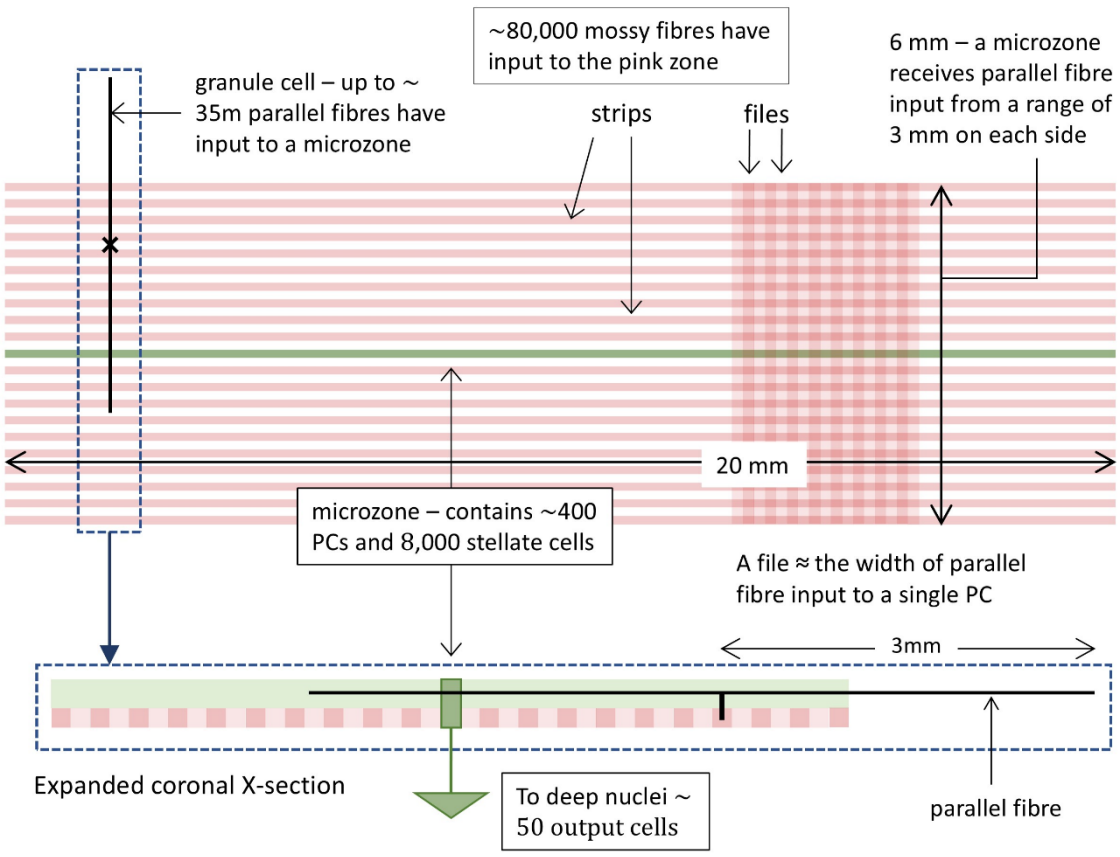


Fig 2. Schematic of a network.

The number of cells in each network layer diminishes from the input layer to the middle layer to the output layer. We use functional and nominal divisions of a network to quantify the proposals. A 'strip' is a functional division of the granular layer lying parallel to microzones (white stripes are also strips). A 'file' is a nominal division of the granular layer lying at right angles to strips, with approximately the width of a Purkinje cell. A 'field' is the intersection of a strip and a file; it is also a nominal division, merely a device used for modelling. END CAPTION

239 It is a final but important feature that code in this form mimics sampling with
240 replacement. Purkinje cells and molecular layer interneurons are both severely
241 flattened in the plane orthogonal to parallel fibres. Both receive contact from a
242 random sample of passing parallel fibres, and therefore receive contact from a random
243 sample that are active. They each receive, therefore, individually, at any moment, a
244 random sample of what is effectively the same distribution of parallel fibre rates,
245 which is independent of the sample received by any other cell, notwithstanding that in
246 most cases they receive contact from a sample of a different population of cells.

247 **2.3.2 PARALLEL FIBRE SYNAPTIC MEMORY**

248 The traditional learning model, which centrally features synaptic memory, has not
249 received confirmation by experimental work. However, there is still a mainstream
250 feeling that the computation will turn out to involve synaptic weights. In this view,
251 core assumptions are retained, and the aim of further work is to understand how
252 plasticity at multiple sites is related, and related to function. Historically, there has
253 been strong emphasis on researching learning [35, 36]. It is regarded as naive to have a
254 different opinion.

255 We think it is unnecessary to interpret experimentally induced learning as evidence of
256 functionally graded synaptic weights. There is now a range of evidence – both in the
257 granular layer, reviewed in Supplemental Materials, and in the molecular layer, cited
258 and discussed in context, in the main text – that firing rates are linearly related, and

259 that synaptic transmission is adapted for linear transmission, and of linear
260 relationships between firing rates and behavioural metrics (suggesting linearity is not
261 merely an artefact).

262 One of the present authors has previously proposed a learning outcome which is
263 consistent with linear transmission. In that proposal, the result of learned parallel fibre
264 synaptic modification is to polarise synaptic weights such that all synapses either
265 transmit robustly (working synapses) or do not transmit at all (silent synapses). This
266 was first proposed, with a synaptic mechanism, here [37]. The learning outcome is the
267 ratio of working to silent synapses. Working synapses are a random subset of the
268 population. There is no pattern memory. The effect on the response of the
269 postsynaptic cell is independent of the pattern in which parallel fibres are active and
270 inactive. Following training, all permutations are received at the same ratio of working
271 and silent synapses regardless of whether or not any particular pattern was involved in
272 training, even if it has never been received before. The effect of learning is to make a
273 gain adjustment to the network computation.

274 We make it a further assumption in this paper that the learning outcome has these
275 features. Therefore, it is unnecessary to include learning here.

276 The learned ratio is the same on all Purkinje cells in the same microzone. Microzones
277 are defined topographically by their climbing fibre input [1, 2] which is received from a
278 discrete cell group in the contralateral inferior olive. In response to stimulation,

climbing fibres in a group fire together, so that a microzone that receives the output of a group receives climbing fibre tuition as a volley. Discharge synchrony of a climbing fibre group is enhanced by dendritic electrotonic coupling of olivary cells [38] and is seen in anaesthetised mice in response to sensory stimuli [3]. Coherence of complex spikes between functionally-grouped Purkinje cells (Purkinje cells reliably respond to climbing fibre discharge with a complex spike) is low when it includes spontaneous discharge but high in response to sensory stimuli [39]. This provides the substrate for microzone-grouped Purkinje cells to receive the same training, so that they learn as a unit.

So, all Purkinje cells receive, at working synapses at any moment, a large random sample of parallel fibre firing rates (and therefore the same code) and the same number of signals (with no effect of the pattern).

2.4 Regulation of the fraction of active parallel fibres

2.4.1 A CASE FOR SPARSE CODE (SMALL FRACTION)

To simulate the computational effect of random sampling, we need to know or derive sample sizes (the number of inputs to each type of unit), and therefore in turn what fraction of parallel fibres is active at any time. There is a long-standing idea that the

297 fraction of parallel fibres that are active at any time is regulated [6], fixed [5], and low
298 (termed sparse code). We disagree with the theoretical reasons, but agree the code is
299 sparse, on these grounds.

300 1) Low density is stable (Fig 3). Randomly fluctuating density would be noise. Stability
301 may be enhanced by feedback via granule cell ascending axon contact on Golgi cell
302 basal dendrites [25] and parallel fibre contact on Golgi cell apical dendrites, acting
303 partly through gap junctions [40].

304 2) Low density is energy efficient. Granule cells are very numerous, and they are not
305 myelinated (so spike propagation is costly).

306 3) The reported number of coincidentally active inputs to a stellate cell is in single
307 figures [41], consistent with low density of co-active parallel fibres (see Table 1).

308 4) It is thought that the pattern of active and inactive parallel fibres, and the pattern in
309 which firing rates are distributed among active cells, are decorrelated – effectively,
310 quasi-randomised. Low density enhances practical decorrelation, that is, if
311 decorrelation is not perfectly random – as it may not be [42] – low density increases
312 the randomness of the pattern of synaptic activations received by targets.

313 For these reasons, we calibrated the granular layer model to convert inputs to a
314 network into a low fraction of active granule cells. We chose 0.5% active because it has
315 support from (as far as we know) the only estimate of the number of co-active inputs

316 to a stellate cell (per section 2.5). We suspect it may be higher – perhaps 1% – but a
317 very low fraction is a harder test that an anatomical computation is feasible (because
318 sample size is smaller).

319 **2.4.2 HOW STRICT IS REGULATION OF THE FRACTION OF ACTIVE PARALLEL FIBRES?**

320 It would cause interference with the computation if the fraction of active parallel fibres
321 varied randomly between sectors of a microzone, or with time. There is no
322 experimental evidence that the fraction is stable or that it is not stable. The granular
323 layer computation simulates the interaction of excitatory input to the cerebellum that
324 is received by granule cells and by granular layer interneurons called Golgi cells, and
325 concurrent inhibition of granule cells by Golgi cells (Supplementary Materials). We
326 adapted the granular layer model to generate parallel fibre input to a microzone. One
327 of the outputs of the model is the number of active granule cells per unit volume of
328 the granular layer.

329 The cerebellar cortex is divided only functionally into microzones and networks.
330 Anatomically, it appears as a seamless carpet of cells. To calculate the number of
331 active granule cells we divided it nominally into sagittal strips (the direction of
332 microzones) and bands which cross strips at right angles (the direction of parallel
333 fibres). Parallel fibre input to a microzone is supplied by 20 strips either side. These are
334 crossed by 100 bands, making a total of 4,100 200 µm x 150 µm 'fields' where a band

335 crosses a strip (the extra 100 is the strip of the granular layer under the middle
336 microzone). So, there are 100 fields in a strip and 41 in a band.

337 Together, strips and bands represent a region of the granular layer with dimensions
338 based on a 3 mm range of parallel fibres in each direction from the granule cell soma
339 [43, 44] and a microzone dimensions of 150 µm x 2,000 µm viewed from the cerebellar
340 surface [2, 3, 45].

341 The results are shown in Figure 3. The number of active granule cells in each field gives
342 the number in each file. The number in a file is equal to the number of active parallel
343 fibres at the intersection of that file with the middle microzone. Local regulation of the
344 number of active granule cells in each field in a strip (Fig 3 columns 2 and 5) becomes
345 still tighter regulation of the number of active parallel fibres in a file (Fig 3 columns 3
346 and 6), a statistical effect of cerebellar cortical architecture. Density is independent of
347 the frequency distribution of mossy fibre rates (Fig 3 columns 1 and 4). At low density
348 of active granule cells, the regulated density of active parallel fibres is hyper strict.

349 Physiological regulation may be still stricter, because the number of active granule
350 cells in each field may not vary as much as it does in the simulation. Local feedback via
351 the granule cell ascending axon [25, 46] may condense the physiological range of field-
352 to-field variation. Feedback via parallel fibres [5, 40] would (additionally) reduce the
353 range of file-to-file variation. These are not represented in the simulation.

So, by our calculation, data are not corrupted by noisy variation of the fraction of parallel fibres that are active, because the fraction is strictly regulated by the granular layer computation.

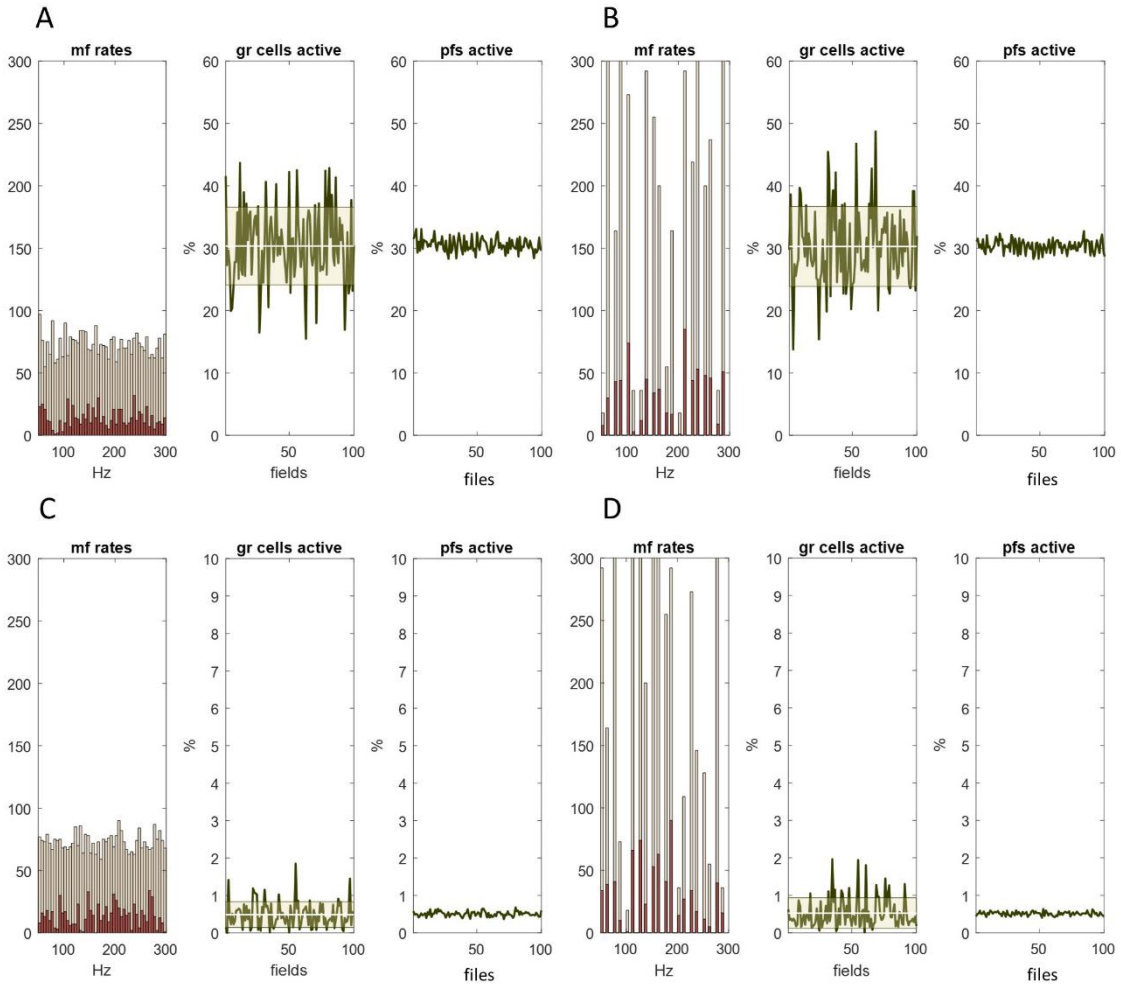


Fig 3. Recoding in the granular layer regulates the fraction of active parallel fibres.

Columns 1 and 4. Firing rates received as input to the granular layer region that supplies parallel fibre input to a microzone. Data represent inputs to a sagittal strip that traverses the region, selected at random. Inputs to each strip were independently randomly

generated with a uniform and discontinuous distribution, respectively. Data represent a moment of time. **Columns 2 and 5.** The percentage of granule cells that fire in each field of a single strip, again selected at random. **Columns 3 and 6.** The percentage of granule cell that fire in each band, therefore also the percentage of active parallel fibres in each band. **Top row.** ∼30% of granule cells are active. **Bottom row**. Golgi cell inhibition of granule cells is adjusted by a constant so that a lower fraction of granule cells fire, around 0.5%. y axes are scaled to the data (so the contraction of the data range in the bottom row is even greater than it appears). The shape of the sampled distribution (columns 1 and 4) does not affect the results, suggesting that regulation of analogue inputs at changing rates has the same result. END CAPTION

2.5 Functional anatomy

Cell anatomies have a function in the computation. This is not incidental but a main driver of their physical form and the scaffold-like geometry of the contact they make on each other. Cell morphologies, cell ratios, orientation to each other, spatial organisation, the shape and size of dendritic and axonal territories, convergence and divergence ratios, and the probabilities that they make and receive contact, all contribute to function. We describe them here, in the methods section, because we derive parameters from anatomy.

The number of Purkinje cells in a microzone probably varies. We use 400. Around 350,000 parallel fibres pass through the dendritic field of a single Purkinje cell, of which one in two make contact, normally at a single synapse (sometimes two, average 1.24) [47]. Purkinje cells are severely flattened in the plane orthogonal to parallel fibres and to the cerebellar surface, filling the molecular layer vertically. As noted, microzones lie across the path of parallel fibres at right angles, so that parallel fibres pass through by the shortest distance.

Purkinje cells receive inhibition from interneurons, stellate cells and in some vertebrate groups, basket cells. We do not consider basket cells. The stellate cell territory is sagittally flattened in the plane orthogonal to parallel fibres. Stellate cells form planar networks [48] that occupy the space (about 40 µm) between Purkinje cells [49]. The stellate cell dendritic territory varies with the depth of the cell body in the molecular layer. For superficially located ('outer-level') cells it is around 80 x 80 µm. At deeper level ('inner level') it is larger, around 120 x 120 µm [49 pp.217-221]. The number of parallel fibres which pass through an inner-level stellate cell territory is given by $a/((b \text{ x } c)/120^2)$, assuming parallel fibres are uniformly distributed, where a is the number that pass through a Purkinje cell territory (350,000 [50]), and b and c are the dimensions of the Purkinje cell territory in the same plane (200 x 300 µm), giving ~84,000. If 0.5% are active in the general population, around 420 of those are active.

It has been estimated that a stellate cell receives contact from several hundred parallel fibres [49, 51], which we take as 1,000. Working with these estimates, 1 in 84 make contact (i.e., 1,000 out of 84,000), so the probability that an active cell makes contact is 1/84 = ~0.0119. Accordingly, contact by x (out of 420) active cells has a probability given by

$$\frac{420!}{x!\,(420-x)!} * 0.0119^{x} * (1-0.0119)^{420-x}$$

Probabilities are shown in Table 1 for the range x = 0 to 12, where x is the number of active parallel fibres that make contact and p is the probability of that number. The distribution of probabilities is in good agreement with the modest number – single figures – suggested by 'two to eight substantial EPSPs' [41 p.9628].

Table 1 The probability, p, that a stellate cell receives contact from x active parallel fibres

x	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

Caption. Table 1 shows the distributed probabilities of the number of active inputs to inner-level stellate cell.

3. Results

A discussion of related evidence is included in the results section. We believe it gains by context.

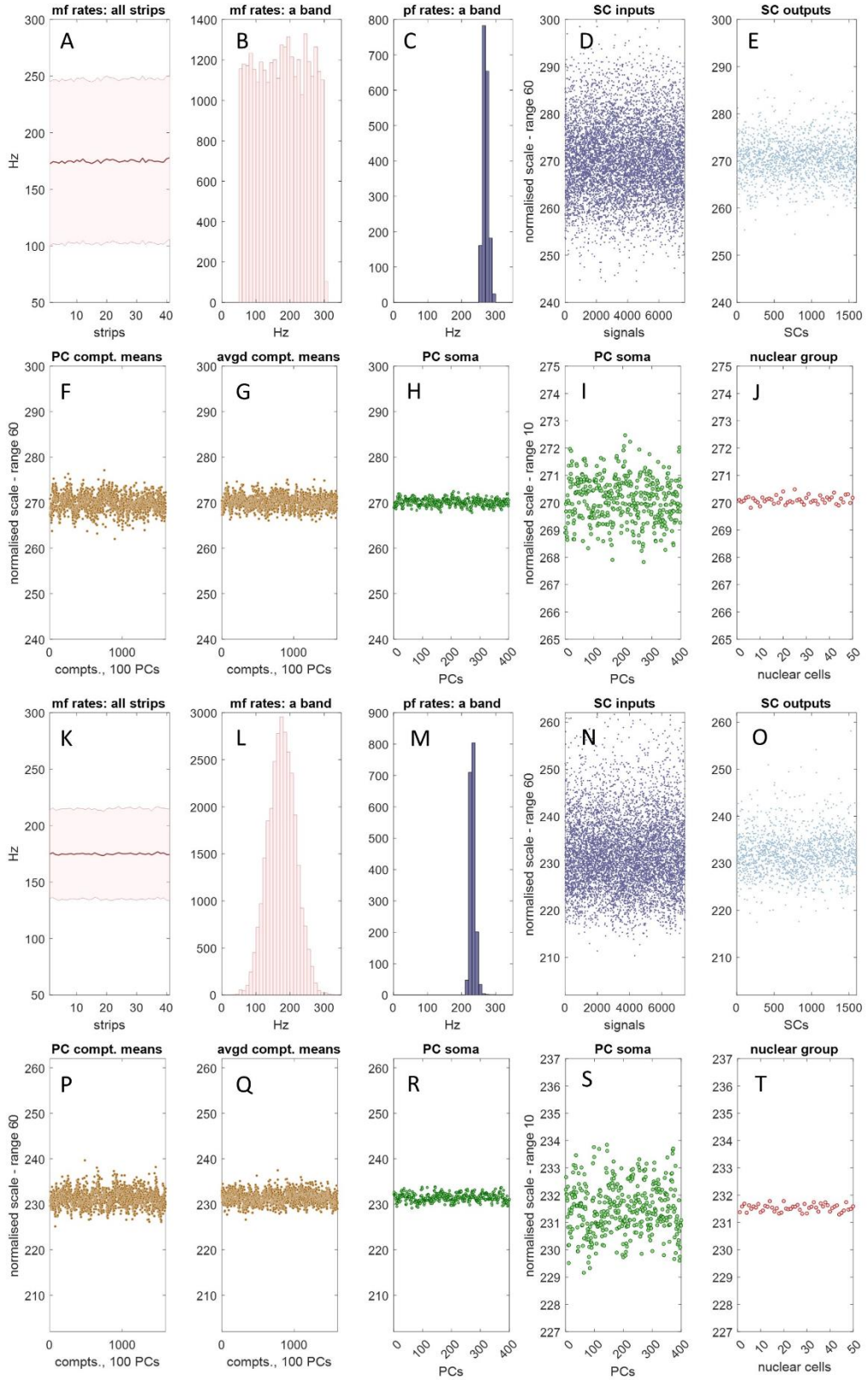
3.1 Stellate cells

We divided cells in a microzone into nominal ‘sectors’. A sector is the size of the intersection of a band and a microzone, so a sector and the field below it together form a column of the cerebellar cortex. Cells in a sector randomly sample parallel fibre rates in that band. Like fields, sectors *are simply a device*. They do not represent a separation of data streams or that there is any difference or compartmentalisation of the way data are processed.

The granular layer model returns the firing rates of granule cells in each field, and therefore the frequency distribution of parallel fibre rates in each file, and so rates received by each sector of the middle microzone, as noted. We next used that to simulate random sampling of parallel fibre firing rates by stellate cells in the middle microzone. There are an estimated 16 stellate cells for each Purkinje cell [52], meaning 80 stellate cells per sector (there are four Purkinje cells per sector so five layers of stellate cells are needed to inhibit them from both sides). Sample size was randomly generated with the probabilities in Table 1. This gave the number of active inputs, and input rates, to each stellate cell in each sector.

Molecular layer interneurons, which include stellate cells, reflect ‘granule cell input with linear changes in firing rate’ [21 p.6]. We assume that stellate cell firing is a linear function of the mean of input rates, which we take to be the mean itself (the sum works as well). This is not to suggest that stellate cell firing rates are equal to the mean of input rates, but that they are some linear function of the mean. As the physiological function is not known, and our purpose is only to demonstrate a statistical relationship of data layers, and not to calculate actual firing rates, we use a normalised scale. The results are shown in Figure 4.

To put that in the language we used earlier, each stellate cell is the physical form of a notional unit, and the afferent data set is the frequency distribution of parallel fibre firing rates. Granule cells fire in short high-frequency bursts [21, 53, 54], and only few are active at any time. But the high turnover of active cells means the distribution is constantly changing. Sample size fluctuates; it varies with time and, at any time, the number of active inputs to a unit has a distributed probability with a result that is independent of other units. The unit computation is a linear function. In physiological conditions, inputs are computed ceaselessly and in parallel by all units – 80 per sector, all sectors, all (at any moment) effectively sampling the same distribution of parallel fibre rates, notwithstanding that they all receive contact from different parallel fibres. The outputs – each an analogue signal – become the data set sampled by the next layer of units.


mf rates: all strips
A
Hz
strips
mf rates: a band
B
Hz
pf rates: a band
C
Hz
SC inputs
D
normalised scale - range 60
signals
SC outputs
E
SCs
PC compt. means
F
normalised scale - range 60
compts., 100 PCs
avgd compt. means
G
compts., 100 PCs
PC soma
H
PCs
PC soma
I
normalised scale - range 10
PCs
nuclear group
J
nuclear cells
mf rates: all strips
K
Hz
strips
mf rates: a band
L
Hz
pf rates: a band
M
Hz
SC inputs
N
normalised scale - range 60
signals
SC outputs
O
SCs
PC compt. means
P
normalised scale - range 60
compts., 100 PCs
avgd compt. means
Q
compts., 100 PCs
PC soma
R
PCs
PC soma
S
normalised scale - range 10
PCs
nuclear group
T
nuclear cells

456

457 **Fig 4. The microzone computation.**

458 SC: stellate cell; PC: Purkinje cell; compt: PC dendritic compartment; mf: mossy fibre; pf:
459 parallel fibre. **A** Mean and SD of mossy fibre input to 41 nominal strips of the granular
460 layer. Strips are 150 microns wide and together make up the region supplying parallel
461 fibre input to the microzone that equally bisects it, indicated by a dashed grey line. The
462 morphology of mossy fibres means that all locations of a strip receive a large random
463 sample of input rates to the whole strip. **B** Frequency distribution of mossy fibre rates
464 received by a single nominal band, 200 microns wide. Bands cross strips at right angles,
465 the direction of parallel fibres. Granule cell firing rates in a band are accordingly parallel
466 fibre rates received by the microzone where they intersect. As bands cross all strips,
467 they all receive the same distribution of mossy fibre rates. There are 100 bands. **C**
468 Parallel fibre rates in a randomly selected band. The frequency distribution of parallel
469 fibres rates is the same in all bands. This is always true, and does not depend on strips
470 receiving the same mossy fibre distribution. **D** Parallel fibre rates received as input to
471 stellate cells in the central microzone, obtained by randomly sampling parallel fibre rates
472 (80 times per file). Sample size was randomly generated with Table 1 probabilities. The
473 y axis range is 60 normalised Hz, down from a mossy fibre range of 250 Hz. **E** Outputs of
474 the stellate cell function. Data represent 16 cells per sector. A sector is the intersection
475 of the central microzone and a band (hence 1,600 data points). There is a further
476 contraction of the data range. **F** The Purkinje cell dendritic arbour was nominally divided
477 into 16 compartments, each randomly sampling local outputs of the stellate cell

478 function. Data are the sample means, shown for a single Purkinje cell per sector in the
479 central microzone (so there are 1,600 values, 16 compartments x 1 Purkinje cell x 100
480 sectors). **G** A Purkinje cell receives inhibition from two planar networks of stellate cells,
481 one each side. Step (F) was repeated, and the data averaged for each compartment. **H**
482 Mean of the compartment means calculated for each Purkinje cell in the central
483 microzone (so 400 data points, 4 Purkinje cells per field x 100 fields), representing
484 somatic integration of dendritic compartment data. **I** The same data but with the y axis
485 range further shortened. A uniform distribution of mossy fibre rates with a range of 50–
486 300 Hz in (A) becomes a very narrow Purkinje cell range with a normal distribution. **J** The
487 output of a microzone is channelled down onto a smaller group of cells which include
488 the output cells of the network. Each output cell receives input from a random sample
489 of 30–50 Purkinje cells [55]. Data are the sample means for 50 output cells. **K–T** We
490 repeated (A)–(J) but with normally distributed mossy fibre rates (SD 40), with the same
491 result. We also obtained the same results (not shown) with a discontinuous mossy fibre
492 distribution (range 50–300 Hz). END CAPTION

493

494 Figure 4 data represent a moment of time. We used values to represent mossy fibre
495 signals that drive parallel fibre input to a microzone, as before. Input rates to each strip
496 were independently randomly generated with a uniform (Fig 4A) and normal (Fig 4K)
497 distribution. Data are displayed as the mean and SD of input to each strip. Firing rates

498 received by each field were derived from input to strips (Supplementary Materials),
499 and input to bands – a row of fields crossing strips at right angles – were derived from
500 input to fields. Input to a randomly selected band is displayed as a histogram (Fig 4B
501 and L). Some very high mossy fibre rates have been recorded but the typical
502 physiological range may be lower [53, 56]. In this figure, strips all receive the same
503 range (50–300 Hz). Therefore, so do all bands.

504 The range of parallel fibre rates is much smaller (shown for a randomly selected band
505 in Fig 4C and M). This is a result, by design, of the statistical effect of random sampling
506 by granule cells, and of regulation of the fraction of granule cells which fire (Fig 3),
507 because the number is controlled by an inhibitory threshold. This limits firing to cells
508 which receive rates at the higher end of the mossy fibre range.

509 The parallel fibre range is the same in all bands, so this is the range that stellate cells
510 all sample (Fig 4D and N). We show data for 16 stellate cells per sector (the number
511 that sit between two Purkinje cells), so this step generates around 7,200 values (16
512 stellate cells x 100 sectors x the weighted mean of distributed probabilities). Signals
513 received by each stellate cell are then passed through the stellate cell unit function.
514 The data range is further narrowed (Fig 4E and O), from an original range of 250 Hz
515 down to around 20 Hz. The outputs of the unit functions become the sample
516 distribution in the next step.

517 There is direct evidence of a linear relationship of granule cell input and stellate cell
518 rates [21] but as far as we know not confirmation, or data about the sensitivity or
519 latency of the response. However, fidelity of transmission may be assisted. GABA
520 spillover from stellate cell synapses enhances parallel fibre synaptic transmission to
521 stellate cells [57, 58]. 'It is well established that parallel fibers express functional
522 $GABA_A$Rs, that their activation depolarizes the axon, and that they increase release
523 probability at parallel fiber synapses' [59 p.16924, 60].

524

525 ## 3.2 Purkinje cells

526 In our simulation, a field contains 4 Purkinje cells. A Purkinje cell receives contact from
527 stellate cells from both of its flat sides, ∼16 each side, and a smaller number of stellate
528 cells in the two sagittally-neighbouring fields. Inhibition from neighbouring fields is
529 from a smaller number of stellate cells because not all of them extend their main axon
530 far enough to reach, and some of them extend their main axon the wrong way. We
531 represent those conditions as variables. The stellate cell main axon gives rise to
532 ascending and descending axon collaterals. The vertical range of collaterals and the
533 horizontal range of the main axon both increase with the depth of the cell body in the
534 molecular layer [49, 61]. Because of their more limited horizontal range, outer-level

535 stellate cells are excluded in the simulation from contacting Purkinje cells in
536 neighbouring fields.

537 The planar Purkinje cell arbour is large and branches profusely, and stellate cell contact
538 is dispersed across it. We divide the Purkinje cell arbour nominally into sixteen 50 x 75
539 µm compartments. Input to a compartment is not confined to (or even necessarily
540 received from all of) the nearest stellate cells. It is unknown how many stellate cells
541 innervate a Purkinje cell compartment. Most stellate cells afferent to a Purkinje cell are
542 in range of most compartments, but not all make contact, and fewer make significant
543 contact which dominates the postsynaptic effect. We assume six make significant
544 contact. To represent that in the model, each compartment receives contact from a
545 random sample of stellate cells which flank a Purkinje cell. Sample size is six. Sampling
546 is independent of sampling by other compartments, i.e., innervation of a compartment
547 by a stellate cell does not exclude innervation by the same cell of other compartments.

548 Here, a notional compartment is a computational unit, and the afferent data set is a
549 (random, effectively) subset of the local outputs of the stellate cell function. The unit
550 computation is again a linear function. Spatial and temporal integration of inputs are
551 absent from the computation because inputs are not discrete. Compartment
552 membrane voltage is a linear function of stellate cell rates afferent to that
553 compartment. In its physical form, the postsynaptic effect – the output of the

computation – is an analogue signal. In its computational form, a compartment continually recalculates inputs.

There is support for a linear effect. It has been reported that Purkinje cell firing is a linear function of synaptic excitation, with inhibition blocked in slices [19, 20]. In mice during locomotion, linear relationships are reported between granule cell rates and interneuron firing, the balance between excitation and inhibition of Purkinje cells and depolarising or hyperpolarising dendritic membrane voltage, and between membrane voltage and bidirectional modulation of Purkinje cell simple spike output [21]. The granule cell → interneuron → Purkinje cell pathway drives a fall in the firing rate of Purkinje cells as a linear function of the strength of input from granule cells [22].

To declutter the data, we show unit outputs for a single Purkinje cell in each sector (so 1,600 values) (Fig 4F and P). Data are now in a range of around 10 Hz on a normalised scale (switching calibration would add a layer of conjecture because the actual function is unreported).

A Purkinje cell receives inhibition from two stellate cell networks, one each side. We repeated the steps above for the other flanking network, then took the mean for each compartment (Fig 4G and Q).

At the next step, dendritic signals are integrated at the Purkinje cell soma. Stellate cells contact Purkinje cells on smooth dendrites [49]. Smooth dendrites extend into all

compartments, mitigating the effect of filtering and distance on dendritic signalling. Recording from smooth dendrites, locomotion-dependent modulation of dendritic membrane voltage 'linearly transforms into bidirectional modulation of PC SSp [Purkinje cell simple spike] output' [21 p.9]. Dendritic signals are a sustained state (we moot). Passive modulation of somatic charge is fast and bidirectional, and reflects the average of dendritic states, which transforms into a linear effect on firing. To represent that, we took the mean of compartment membrane voltages for each Purkinje cell in all sectors, so that this step returned 400 values (4 x 100 Purkinje cells; Fig 4H, I, R and S). There is a further contraction of the data range. This then becomes the data set sampled by the next layer of computational units, with the physical form of a cell group in the deep cerebellar nuclei, which includes the output cells of the network (not part of the microzone computation but included in the Discussion).

The result is a progressive contraction of the data range received as inputs to the system, a cell network. At each step, data are coded in a frequency distribution. There is a linear relationship of the mean of each distribution with any other distribution. With physiological parameters, this is sufficient to distil a single firing rate even from an input range of 250 Hz. This result was independent of the input distribution (tested: uniform, normal and discontinuous).

4. Discussion

4.1 Summary of ideas

We make the assumptions stated earlier. A mechanism and evidence of the granular layer computation [34] and microzone memory [37] have appeared previously.

The microzone computation is a computational effect of anatomy. This is not incidental but a main driver of cell morphologies, the three-dimensional space they occupy and the scaffold-like geometry of the connections they make. Communication of cells is linear. Cells are highly adapted for faithful transmission [23-29]. Neurobiology is customised to control for biophysical and other noise. 'Although cerebellar neurons and synapses have nonlinear intrinsic properties [62-66], these properties are disengaged or compensated in ways that keep the cerebellar network operating in a linear regime [29, 67-70]' [30]. Anatomy and linearity combine. The result is a passive, unlearned computational effect of simply passing signals through the network. At each step, data are coded in a frequency distribution. Anatomy provides the physical hardware to randomly sample the afferent data layer. Sampling is a facsimile of sampling with 'replacement'. Samples are passed through the biological form of linear functions to generate the next data layer, which is in turn itself sampled by the next layer, and so on. Unit computations in a layer are made in parallel, continuously, so that unit outputs are constantly refreshed, i.e., analogue. The

result is a progressive reduction of the data range but preservation of a linear relationship of the means, by the central limit theorem.

4.2 Context related to theory

The dominant cerebellum model is the supervised learning model, a group of models and variants which share the idea that the cerebellum implements a learning algorithm which teaches memories that are stored as synaptic changes [5-7, 71]. The algorithm depends on the model. This idea has been widely influential on neural network modelling generally. Following training, a pattern of inputs that is received at training-modified synapses generates a learned response. Instruction is externally sourced, hence 'supervised'. The model does not need to conjecture how mossy fibre rates and granule cell rates are related because it is an assumption that the desired output of the system is what the algorithm teaches.

As noted, after a promising early start [72], there have been problems with the evidence [8-13]. It is academically possible but difficult to reconcile learning models with linear transmission because it is the purpose of teaching in learning models to *displace* a purely rate-coded response (and replace it with a learned one).

Our solution is to jettison the twin pillars of the traditional model – pattern memory stored as tuned synaptic weights, and a learning algorithm needed to teach them – and to propose instead that input and output of the system are connected by a chain of rate codes. Our unit functions are linear. Units are not an analogue of synapses. Output is an unlearned response to input rates and the mechanism of the network computation is anatomy. Many neural network learning models share the core assumptions of the traditional cerebellar model. We disagree with them also in these ways.

4.3 Context related to function

Network modelling is normally led by a high-level idea, which can be a function or problem for which the model is the mechanism or a solution. The order is function → model (the third step is to consult the evidence for support). These are the first two levels of Marr's levels of analysis. It is not contentious that a mis-step at level one is fatal. We have the reverse problem. We have a mechanism without a function. The high-level step is at the end, an inference from the mechanism.

The cerebellum is necessary for skilled execution and coordination of movement. We reason that in motor control, the cerebellum may process feedback to inform motor output. Cerebellospinal neurons provide 'a direct link between the cerebellum and

648 spinal substrates for motor coordination' [73]. The output pathway is reciprocated by
649 spinocerebellar neurons, providing a substrate for direct motor control, which spinal
650 locomotor networks alone cannot accomplish reliably or with finesse [74]. Neurons of
651 the spinocerebellar tracts are excited monosynaptically by group Ia and group Ib
652 afferents that originate in muscle spindles and tendon organs respectively [75].

653 As noted, we have proposed that the granule cell code conflates mossy fibre rate
654 information in long thin strips, so that a unit of granule cell information has those
655 minimum dimensions. We must suppose, if we are correct, that disregard of resolution
656 below that scale is functional. To conjecture, the many-to-one conversion of input to
657 output rates may distil a precise readout from individually untidy data, which include
658 proprioceptive data. Biologically made components, including muscle fibres, stretch
659 receptors and neurons, are not precision-made standard units, as models tend to
660 assume. Speaking generally, it is well established that neuron performance is noisy –
661 there is inherent biological variability from cell to cell and even of firing of the same
662 cell at different times in the same conditions. Muscle fibres come in different types,
663 vary individually, and have different tensile properties. Synaptic transmission is
664 randomly variable, even at the same synapse, and so on. Virtually all recordings
665 provide examples.

666 The biological solution – we suggest – is to conflate the data using a statistical effect of
667 random sampling. Output of a muscle is pooled for this purpose. Input to the

cerebellum is accordingly handled with disregard for resolution below a topographically-defined scale. It may be that large cell groups in the cerebellum are a representation of multicellular structures, which include muscles, which themselves generate signals which together code indivisible information. Hyper-conflation of data is a *function* of cerebellar data processing.

Correspondence between a code group in the cerebellum and a muscle may not be limited to the input pathway. 'Motor neurons are clustered in columnar, spinal nuclei called motor neuron pools (or motor nuclei). All of the motor neurons in a motor neuron pool innervate a single muscle, and all motor neurons that innervate a particular muscle are contained in the same motor neuron pool. Thus, there is a one-to-one relationship between a muscle and a motor neuron pool.'[1] Spinal neurons terminate profusely in a motor neuron pool. In principle, this permits a one-to-one relationship of a cerebellar network and a motor pool, and through a motor pool, a one-to-one relationship with a muscle.

Graded control of muscle fibres is provided because motor neurons are different sizes (in mammals there are three sizes). Each step up in size selectively innervates more powerfully contractile (but more exhaustible) muscle fibres. Larger motor neurons require stronger depolarising current to fire, so weak input signals excite only small motor neurons and the least powerful muscle fibres. Stronger signals cause more

[1] Reproduced with permission by McGovern Medical School, University of Texas

motor neurons to be recruited, and therefore more muscle fibres, including fibres that contract with more force.

Credibly, these are the input and output pathways of a proprioception-motor loop and feedback about movement is processed to inform motor output. The effect and function of processing is to distil precision from individually imprecise data. Inform means input rates code and *directly drive* motor output. The appearance of the sacrifice of resolution is in fact what a collective code looks like – single signals have no meaning for the cerebellum, which is designed to conflate data below minimum dimensions. The cerebellum does not need to 'know' what data it is working with because code groups are defined by topography.

Code groups are strips in our terminology, but in complex animals there is no need to think that strips in a network all have the same width, or that networks are all the same size. The idea that a muscle may drive motor output to another muscle is a candidate to explain where drive comes from, if it is not central pattern generators. In higher vertebrates, it is likely that this simple principle has acquired complex organisation. But it is readily imagined in a simpler system.

The vertebrate cerebellum first appeared in a marine environment. It is probably not contentious that early vertebrate swimming, before efficient teleosts, was by undulating body movements (anguilliform). Almost the entire musculature of fish is made up of a bilaterally symmetrical row of paired axial muscle segments, termed

myomeres. Movement through the water is produced by a repeating sequence of muscle contractions which cause lateral body flexion. In anguilliform swimming, alternating left and right head turns set up a herring-bone pattern of high pressure which the animal rides by making itself an undulating shape. A sagittally-flattened tail may contribute thrust.

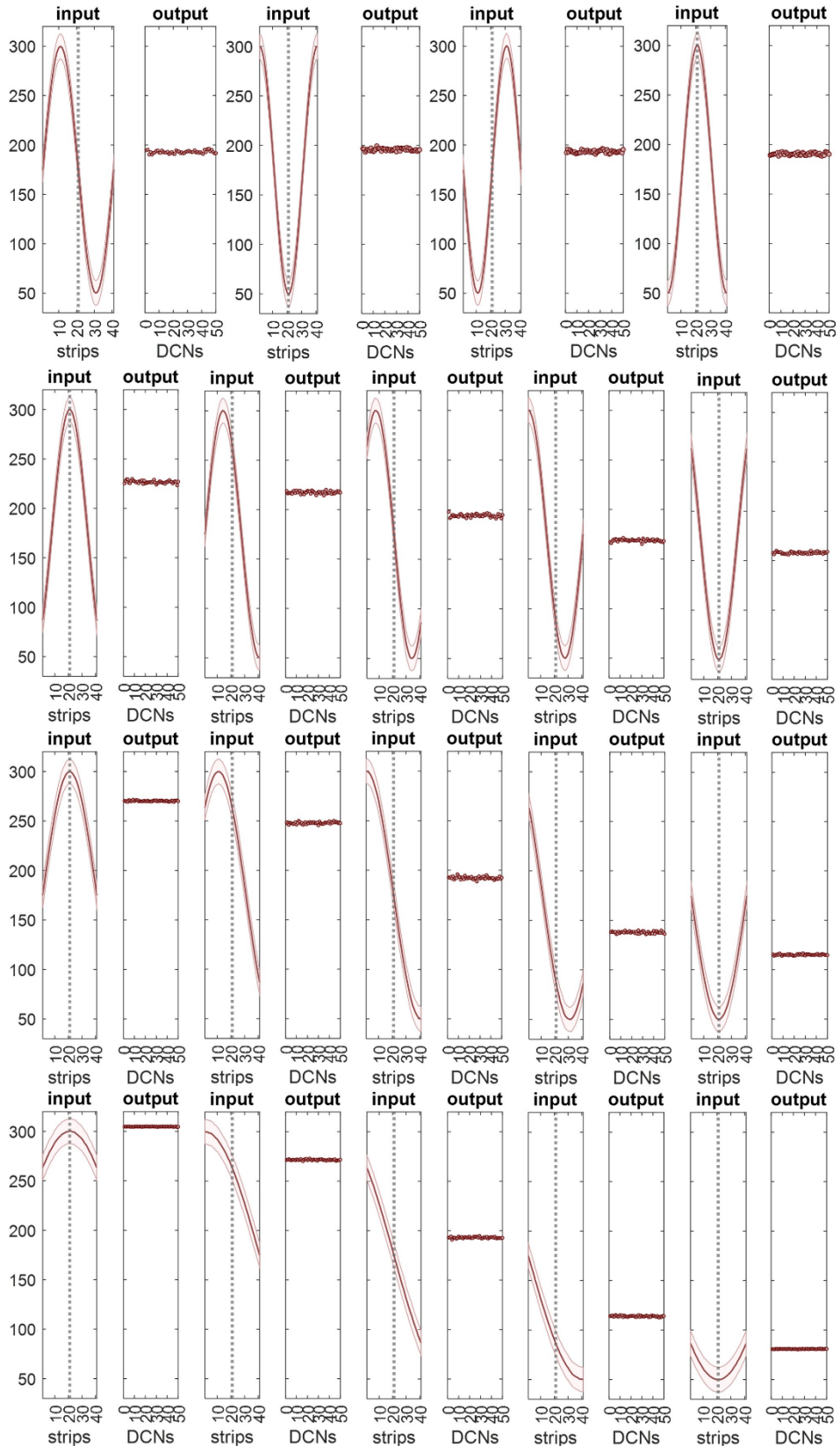
Modern muscles generate information constantly, in muscle spindles, specialised muscle fibres. Group II afferent neurons signal muscle length. Group Ia afferent neurons fire at a rate that varies with muscle length but in addition varies proportionally to the rate of muscle extension and (more modestly) inversely proportionally to the rate of muscle contraction. We observe simply that these are the makings of self-drive. Say that in a vertebrate ancestor, input to a row of topographically-defined strips of the granular layer arose from a corresponding row of contralateral myomeres, with neighbouring strips receiving the output of neighbouring myomeres. Say also that these provide drive via parallel fibres to the microzone which equally bisects the group, and that this controls motor output to the ipsilateral muscle segment opposite the middle of the row providing input. A network is therefore the junction of multiple input pathways and a single output pathway.[2] Passive muscle extension might in this way be used to drive muscle contractions on the other side of

[2] To elaborate, neighbouring networks connected by parallel fibres send output to neighbouring myomeres. This connectivity is mirrored in the contralateral cerebellum, which has motor output to the other half of the body.

725 the body. Contractions would be automatically phase-locked and wavelength sensitive.
726 The reason that can't be mediated by spinal reflexes is that it would drive a positive
727 feedback loop. To address that, it is routed through the cerebellum. The solution is
728 parallel fibres.

729 Parallel fibres add a second step to the sacrifice of resolution. Strip data only retain
730 their integrity – i.e., are only segregated – in granule cell ascending axons, while signals
731 travel towards the cerebellar surface. Beyond that, after the granule cell axon divides
732 in two, and parallel fibre signals are randomly mixed up, within modal strata [49, 53,
733 76]. On the face of it, this only seems to spectacularly compound the corruption of
734 clean data available downstream to microzones. However, it is able to code phase and
735 wavelength because of topography. By averaging strip codes to generate output, the
736 amplitude of output is made subject to constraints that it can't exceed. The cap varies
737 with the number of muscle segments in a wavelength of the body and the number of
738 strip codes that are averaged to generate an output signal.

739 We can visualise this by varying input rates to each strip. In our illustration, the
740 number of strip codes is constant and the number of muscle segments in a body
741 wavelength is varied (Fig 5). Decreasing the amount of a body wavelength represented
742 in signals that are received as input to a network, increases output amplitude (and
743 sensitivity), and vice versa. Note that output is phase sensitive because it is coded in
744 input.



745

746 **Fig 5. In theory, parallel fibres allow mutual exchange of drive between antagonist**
747 **muscles without positive feedback**

748 **A** Input rates to a network were varied from strip to strip, so that a plot of the mean and
749 SD of input rates to each strip is sine-wave-like. The frequency distribution received by
750 strips individually is normal and narrow (SD = 12.5 Hz, all strips). The central strip is
751 indicated by the vertical dashed line. We imagined that in a locomotor cycle, a single
752 strip receives input which varies sinusoidally at these rates. All strips receive the same
753 cycle. Next-neighbouring strips are a step out of phase. Each pair of panels – 1 and 2, 3
754 and 4 and so on – represents a moment in time, but at a different phase of the cycle,
755 advanced 90 degrees each time. Input to the network is shown in the left panel of a pair.
756 This was passed through the network computation to derive output. Output of the
757 computation is shown in the right panel. This represents firing of deep cerebellar
758 neurons (DCNs), the network output cells. Output is the same in all phases. Parallel fibres
759 perfectly dampen output. **B** The same but input rates represent 3/4 of a wavelength and
760 the phase shift between panel pairs is 45 degrees. Output has become sensitive to cycle
761 phase but is still dampened. **C** Input rates represent 1/2 a wavelength. Amplitude of the
762 output range is now around twice the range in (B). **D** Input rates represent 1/4 of a
763 wavelength. The output range increases again. Amplitude now approaches the input
764 range (and remains phase sensitive) but can't exceed a maximum which is set by a linear
765 relationship with input rates. END CAPTION

766 There are two physiological adaptations (not included in the simulation) that may
767 mean the response receives phase-biased weighting. Parallel fibres make about half as
768 many synapses distally than proximally, and swellings where contact is made contain
769 fewer synaptic densities [77]. As a result, there is a stronger influence of nearer
770 granule cells. Second, Purkinje cells receive significant granule cell ascending axon
771 contact [78]. Ascending axon synapses are resistant to long-term depression, and have
772 low uncertainty of release probabilities [79]. Accordingly, excitation of Purkinje cells
773 receives phase weighting from ascending axons of directly underlying granule cells.

774 We observe finally that this suggests how motor outputs may be coordinated by
775 topography of proprioception-motor loops which are connected by parallel fibres.
776 Immediately neighbouring microzones receive almost all the same parallel fibre input,
777 made up of all the same strip codes except one. Input to their next nearest neighbours
778 differs by two strip codes, and so on. The strip code is a frequency distribution.
779 Compounding strip codes as parallel fibre signals creates a frequency distribution
780 centred on the mean of the strip code means. So, microzones receive drive which is
781 automatically coordinated with their near neighbours because it has the same range
782 which is centred on the mean of most of the same signals. The relationship is closest
783 with close neighbours and weaker with further microzones, diminishing with distance.
784 Because firing of Purkinje cells – and therefore in turn output cells – is coded in parallel
785 fibre rates, output of neighbours is automatically coordinated by the relationship of
786 inputs.

Acknowledgements

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

Author contributions

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

REFERENCES

1. Garwicz, M., C.F. Ekerot, and H. Jorntell, *Organizational principles of cerebellar neuronal circuitry*. News in Physiological Sciences, 1998. **13**: p. 28-34.
2. Oscarsson, O., *Functional units of the cerebellum - sagittal zones and microzones*. Trends in Neurosciences, 1979. **2**: p. 143-145.
3. Ozden, I., et al., *Reliable coding emerges from coactivation of climbing fibers in microbands of cerebellar Purkinje neurons*. J Neurosci, 2009. **29**(34): p. 10463-73.
4. Pedroarena, C.M. and C. Schwarz, *Efficacy and short-term plasticity at GABAergic synapses between Purkinje and cerebellar nuclei neurons*. J Neurophysiol, 2003. **89**(2): p. 704-15.
5. Albus, J.S., *A Theory of Cerebellar Function*. Mathematical Biosciences, 1971. **10**(1-2): p. 25-61.
6. Marr, D., *A theory of cerebellar cortex*. J Physiol, 1969. **202**(2): p. 437-70.
7. Fujita, M., *Adaptive filter model of the cerebellum*. Biol Cybern, 1982. **45**(3): p. 195-206.
8. Welsh, J.P., et al., *Normal motor learning during pharmacological prevention of Purkinje cell long-term depression*. Proc Natl Acad Sci U S A, 2005. **102**(47): p. 17166-71.
9. Schonewille, M., et al., *Reevaluating the role of LTD in cerebellar motor learning*. Neuron, 2011. **70**(1): p. 43-50.
10. ten Brinke, M.M., et al., *Evolving Models of Pavlovian Conditioning: Cerebellar Cortical Dynamics in Awake Behaving Mice*. Cell Rep, 2015. **13**(9): p. 1977-88.
11. Boele, H.-J., et al., *Impact of parallel fiber to Purkinje cell long-term depression is unmasked in absence of inhibitory input*. Science Advances, 2018. **4**(10).
12. Isope, P. and B. Barbour, *Properties of unitary granule cell-->Purkinje cell synapses in adult rat cerebellar slices*. J Neurosci, 2002. **22**(22): p. 9668-78.
13. Ekerot, C.F. and H. Jörntell, *Parallel fibre receptive fields of Purkinje cells and interneurons are climbing fibre-specific*. Eur J Neurosci, 2001. **13**(7): p. 1303-10.
14. Hesslow, G., et al., *Classical conditioning of motor responses: what is the learning mechanism?* Neural Netw, 2013. **47**: p. 81-7.
15. Rancz, E.A., et al., *High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons*. Nature, 2007. **450**(7173): p. 1245-8.
16. Ritzau-Jost, A., et al., *Ultrafast action potentials mediate kilohertz signaling at a central synapse*. Neuron, 2014. **84**(1): p. 152-163.
17. Duguid, I., et al., *Control of cerebellar granule cell output by sensory-evoked Golgi cell inhibition*. Proc Natl Acad Sci U S A, 2015. **112**(42): p. 13099-104.
18. Duguid, I., et al., *Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex*. J Neurosci, 2012. **32**(32): p. 11132-43.
19. Walter, J.T. and K. Khodakhah, *The linear computational algorithm of cerebellar Purkinje cells*. J Neurosci, 2006. **26**(50): p. 12861-72.
20. Walter, J.T. and K. Khodakhah, *The advantages of linear information processing for cerebellar computation*. Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4471-6.
21. Jelitai, M., et al., *Dendritic excitation-inhibition balance shapes cerebellar output during motor behaviour*. Nat Commun, 2016. **7**: p. 13722.

22. Park, S.M., E. Tara, and K. Khodakhah, *Efficient generation of reciprocal signals by inhibition*. J Neurophysiol, 2012. **107**(9): p. 2453-62.
23. Delvendahl, I. and S. Hallermann, *The Cerebellar Mossy Fiber Synapse as a Model for High-Frequency Transmission in the Mammalian CNS*. Trends Neurosci, 2016. **39**(11): p. 722-737.
24. Saviane, C. and R.A. Silver, *Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse*. Nature, 2006. **439**(7079): p. 983-7.
25. Cesana, E., et al. *Novel granule cell-Golgi cell excitatory input in the cerebellar granular layer*. in *FENS Abstr*. 2010.
26. Moore, J.W. and D.E.J. Blazis, *11 - Simulation of a Classically Conditioned Response: A Cerebellar Neural Network Implementation of the Sutton–Barto–Desmond Model*, in *Neural Models of Plasticity*, J.H. Byrne and W.O. Berry, Editors. 1989, Academic Press. p. 187-207.
27. Sargent, P.B., et al., *Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse*. J Neurosci, 2005. **25**(36): p. 8173-87.
28. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptic Specializations Support Frequency-Independent Purkinje Cell Output from the Cerebellar Cortex*. Cell Reports, 2016. **17**(12): p. 12.
29. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptotagmin 7 confers frequency invariance onto specialized depressing synapses*. Nature, 2017. **551**: p. 503.
30. Raymond, J.L. and J.F. Medina, *Computational Principles of Supervised Learning in the Cerebellum*. Annu Rev Neurosci, 2018. **41**: p. 233-253.
31. DiGregorio, D.A., et al., *Desensitization properties of AMPA receptors at the cerebellar mossy fiber granule cell synapse*. J Neurosci, 2007. **27**(31): p. 8344-57.
32. Mapelli, L., S. Solinas, and E. D'Angelo, *Integration and regulation of glomerular inhibition in the cerebellar granular layer circuit*. Front Cell Neurosci, 2014. **8**: p. 55.
33. Galliano, E., P. Mazzarello, and E. D'Angelo, *Discovery and rediscoveries of Golgi cells*. J Physiol, 2010. **588**(Pt 19): p. 3639-55.
34. Gilbert, M. and A. Rasmussen, *The cerebellum converts input data into a hyper low-resolution granule cell code with spatial dimensions: a hypothesis*. bioRxiv, 2024: p. DOI: 10.1101/2023.07.14.548987 ver 9, https://www.biorxiv.org/content/10.1101/2023.07.14.548987v9.
35. Hansel, C., D.J. Linden, and E. D'Angelo, *Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum*. Nat Neurosci, 2001. **4**(5): p. 467-75.
36. Gao, Z., B.J. van Beugen, and C.I. De Zeeuw, *Distributed synergistic plasticity and cerebellar learning*. Nature Reviews Neuroscience, 2012. **13**(9): p. 619-635.
37. Gilbert, M., *Gating by Memory: a Theory of Learning in the Cerebellum*. The Cerebellum, 2021. **21**(6): p. 926-943.
38. De Gruijl, J.R., et al., *Inferior Olive: All Ins and Outs*, in *Handbook of the Cerebellum and Cerebellar Disorders*, S.J.D. Manto M., Rossi F., Gruol D.L., Koibuchi N., Editor. 2012, Springer. p. 1013-1058.

39. Ju, C., et al., *Neurons of the inferior olive respond to broad classes of sensory input while subject to homeostatic control.* J Physiol, 2019. **597**(9): p. 2483-2514.
40. Gilbert, M. and A. Rasmussen, *Gap Junctions May Have A Computational Function In The Cerebellum: A Hypothesis*. The Cerebellum, 2024. **23**(5): p. 1903-1915.
41. Jörntell, H. and C.F. Ekerot, *Receptive field plasticity profoundly alters the cutaneous parallel fiber synaptic input to cerebellar interneurons in vivo*. J Neurosci, 2003. **23**(29): p. 9620-31.
42. Wilms, C.D. and M. Häusser, *Reading out a spatiotemporal population code by imaging neighbouring parallel fibre axons in vivo.* Nature Communications, 2015. **6**: p. 6464.
43. Brand, S., A.L. Dahl, and E. Mugnaini, *The length of parallel fibers in the cat cerebellar cortex. An experimental light and electron microscopic study.* Exp Brain Res, 1976. **26**(1): p. 39-58.
44. Mugnaini, E., *The length of cerebellar parallel fibers in chicken and rhesus monkey*. J Comp Neurol, 1983. **220**(1): p. 7-15.
45. Ramirez, J.E. and B.M. Stell, *Calcium Imaging Reveals Coordinated Simple Spike Pauses in Populations of Cerebellar Purkinje Cells.* Cell Reports, 2016. **17**(12): p. 3125-3132.
46. Cesana, E., et al., *Granule cell ascending axon excitatory synapses onto Golgi cells implement a potent feedback circuit in the cerebellar granular layer*. J Neurosci, 2013. **33**(30): p. 12430-46.
47. Napper, R.M. and R.J. Harvey, *Number of parallel fiber synapses on an individual Purkinje cell in the cerebellum of the rat*. J Comp Neurol, 1988. **274**(2): p. 168-77.
48. Häusser, M. and B.A. Clark, *Tonic synaptic inhibition modulates neuronal output pattern and spatiotemporal synaptic integration.* Neuron, 1997. **19**(3): p. 665-78.
49. Palay, S.L. and V. Chan-Palay, *Cerebellar cortex : cytology and organization*. 1974, Berlin: Springer.
50. Harvey, R.J. and R.M. Napper, *Quantitative study of granule and Purkinje cells in the cerebellar cortex of the rat.* J Comp Neurol, 1988. **274**(2): p. 151-7.
51. Carter, A.G. and W.G. Regehr, *Quantal events shape cerebellar interneuron firing.* Nat Neurosci, 2002. **5**(12): p. 1309-18.
52. Llinás, R.R. and M.N. Negrello. *Cerebellum*. Scholarpedia, 10(1):4606 2015 [cited 2017 06/01]; Available from: http://www.scholarpedia.org/article/Cerebellum.
53. Jörntell, H. and C.F. Ekerot, *Properties of somatosensory synaptic integration in cerebellar granule cells in vivo.* J Neurosci, 2006. **26**(45): p. 11786-97.
54. van Beugen, B.J., et al., *High Frequency Burst Firing of Granule Cells Ensures Transmission at the Parallel Fiber to Purkinje Cell Synapse at the Cost of Temporal Coding.* Frontiers in Neural Circuits, 2013. **7**: p. 95.
55. Person, A.L. and I.M. Raman, *Synchrony and neural coding in cerebellar circuits*. Frontiers in Neural Circuits, 2012. **6**: p. 97.
56. van Kan, P.L., A.R. Gibson, and J.C. Houk, *Movement-related inputs to intermediate cerebellum of the monkey.* J Neurophysiol, 1993. **69**(1): p. 74-94.
57. Stell, B.M., et al., *Activation of presynaptic GABA(A) receptors induces glutamate release from parallel fiber synapses.* J Neurosci, 2007. **27**(34): p. 9022-31.

58. Dellal, S.S., R. Luo, and T.S. Otis, *GABAA receptors increase excitability and conduction velocity of cerebellar parallel fiber axons.* J Neurophysiol, 2012. **107**(11): p. 2958-70.
59. Pugh, J.R. and C.E. Jahr, *Activation of Axonal Receptors by GABA Spillover Increases Somatic Firing.* The Journal of Neuroscience, 2013. **33**(43): p. 16924-16929.
60. Pugh, J.R. and C.E. Jahr, *Axonal GABAA receptors increase cerebellar granule cell excitability and synaptic activity.* J Neurosci, 2011. **31**(2): p. 565-74.
61. Sultan, F. and J.M. Bower, *Quantitative Golgi study of the rat cerebellar molecular layer interneurons using principal component analysis.* J Comp Neurol, 1998. **393**(3): p. 353-73.
62. Hoebeek, F.E., et al., *Differential olivo-cerebellar cortical control of rebound activity in the cerebellar nuclei.* Proc Natl Acad Sci U S A, 2010. **107**(18): p. 8410-5.
63. Ishikawa, T., M. Shimuta, and M. Häusser, *Multimodal sensory integration in single cerebellar granule cells in vivo.* eLife, 2015. **4**: p. e12916.
64. Jahnsen, H., *Electrophysiological characteristics of neurones in the guinea-pig deep cerebellar nuclei in vitro.* J Physiol, 1986. **372**: p. 129-47.
65. Loewenstein, Y., et al., *Bistability of cerebellar Purkinje cells modulated by sensory stimulation.* Nat Neurosci, 2005. **8**(2): p. 202-11.
66. Person, A.L. and I.M. Raman, *Purkinje neuron synchrony elicits time-locked spiking in the cerebellar nuclei.* Nature, 2012. **481**(7382): p. 502-505.
67. Alvina, K., et al., *Questioning the role of rebound firing in the cerebellum.* Nat Neurosci, 2008. **11**(11): p. 1256-8.
68. Rössert, C., et al., *Model cerebellar granule cells can faithfully transmit modulated firing rate signals.* Front Cell Neurosci, 2014. **8**: p. 304.
69. Schonewille, M., et al., *Purkinje cells in awake behaving animals operate at the upstate membrane potential.* Nat Neurosci, 2006. **9**(4): p. 459-61; author reply 461.
70. van Vreeswijk, C. and H. Sompolinsky, *Chaos in neuronal networks with balanced excitatory and inhibitory activity.* Science, 1996. **274**(5293): p. 1724-6.
71. Albus, J.S. *The Marr and Albus theories of the cerebellum-two early models of associative memory*. in *Digest of Papers. COMPCON Spring 89. Thirty-Fourth IEEE Computer Society International Conference: Intellectual Leverage*. 1989.
72. Ito, M. and M. Kano, *Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex.* Neurosci Lett, 1982. **33**(3): p. 253-8.
73. Sathyamurthy, A., et al., *Cerebellospinal Neurons Regulate Motor Performance and Motor Learning.* Cell Rep, 2020. **31**(6): p. 107595.
74. Parker, D., *Complexities and uncertainties of neuronal network function.* Philos Trans R Soc Lond B Biol Sci, 2006. **361**(1465): p. 81-99.
75. Sengul, G. and C. Watson, *Chapter 8 - Ascending and Descending Pathways in the Spinal Cord*, in *The Rat Nervous System (Fourth Edition)*, G. Paxinos, Editor. 2015, Academic Press: San Diego. p. 115-130.

76. Zhang, W. and D.J. Linden, *Calcium influx measured at single presynaptic boutons of cerebellar granule cell ascending axons and parallel fibers*. Cerebellum, 2012. **11**(1): p. 121-31.
77. Pichitpornchai, C., J.A. Rawson, and S. Rees, *Morphology of parallel fibres in the cerebellar cortex of the rat: an experimental light and electron microscopic study with biocytin*. J Comp Neurol, 1994. **342**(2): p. 206-20.
78. Gundappa-Sulur, G., E. De Schutter, and J.M. Bower, *Ascending granule cell axon: an important component of cerebellar cortical circuitry*. J Comp Neurol, 1999. **408**(4): p. 580-96.
79. Sims, R.E. and N.A. Hartell, *Differences in transmission properties and susceptibility to long-term depression reveal functional specialization of ascending axon and parallel fiber synapses to Purkinje cells*. J Neurosci, 2005. **25**(12): p. 3246-57.