How does the cerebellum coordinate motor outputs? Lessons from the past

Short title: Cerebellar co-ordination of motor outputs

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

In most vertebrates, the cerebellum occupies 20–30% of the brain. It is critical for the smooth execution and coordinated timing of movement. It is modular – made up of repeating, anatomically overlapping networks. It is poorly understood how the cerebellar (or any) network codes and processes information, and little is known about the detailed topography of inputs and outputs of anatomically close and functionally related networks. Normally, network models are simplified to remove interference with theoretical computations. We propose the reverse, with evidence: neurophysiology itself controls for noisy variables to exploit statistical effects of running straightforward computations in parallel. With this foundation, we make a proposal in two parts. The first part is the network computation. We propose the cerebellar network computation is a passive effect of anatomy and random sampling by networked linear units, crucially mimicking replacement. This has the result that input to the cerebellum is converted to a hyper low-resolution internal code which spans the network, which is in turn progressively condensed to distil an output rate. The code is analogue and there is no algorithm. In the second part we use the network computation in a proposal of the principles of a wiring plan that extends a network outside the brain. We model undulating swimming. We propose that the cerebellum connects proprioception-motor loops and that data sharing between networks

automates coordination, on a plan that may have been conserved in its basics across vertebrate groups.

Keywords evolution, cerebellum, theory, microzone, network, code

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 main input to the cerebellum is received in the granular layer by granule cells, from mossy fibres. The molecular layer is divided functionally into long, thin strips called microzones [1-3], which lie at right angles to the axons of granule cells, which are called parallel fibres. Parallel fibres lie parallel to each other and the cerebellar surface. Tens of millions of parallel fibres pass through a microzone (see Fig 1 for cell morphologies, wiring and cell names).

Neurons code information in the rate at which they discharge action potentials (firing rates). It has been proposed that the granular layer converts high-dimensional mossy fibre input into a very low-resolution granule cell code [4]. The granule cell code is in units of information which have spatial dimensions that mirror microzones, so the granule cell code is striped. Information is coded in a form (collectively, in any snapshot of firing rates) that is independent of the pattern of active cells, so it is unnecessary to repeat the pattern to code the data (controversial). There is no smaller resolution of granule cell encoded information, and in that sense information has volume, like 3-D pixels. A pixel is the smallest unit of information. Below that scale, the

code is ‘homogenised’ – any random sample of granule cell signals (over a minimum) codes the same information (controversial).

If that is correct, it is unexpected, because it would mean there is a huge sacrifice of the resolution of input data. It is also contentious because it is difficult to reconcile with core assumptions that network models normally make. We will argue that organised data loss is a *function* and core attribute of the cerebellum. The first part of this paper focusses on what microzones do with the output of the granular layer conversion.

A microzone is the middle layer of the cerebellar network. The input layer is the region of the granular layer that provides parallel fibre input to it, and the output layer is the group of cells in the deep cerebellar nuclei that receives its output, which includes the output cells of the network (also the output cells of the cerebellum). Output cell firing rates receive powerful and bidirectional modulation by a coordinated change of Purkinje cell rates [5].

Firing rates are the biological form of numerical values that provide the inputs to network computations. Network models are normally simplified to remove noisy biological variables, enabling network computations to be performed and allowing them to be sophisticated. We argue the reverse: biology is highly adapted to control for biophysical and other noise in order to perform straightforward mathematical operations (but many of them). Neurobiology is awash with variables. As most code

nothing, there are strategies to isolate the effect of selected parameters. There are reported 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 [6-17].

There is growing evidence that rate information is linearly conserved in signalling between cerebellar neurons [6, 7, 12, 13], including synaptic signalling in the molecular layer [18-21], and that cells are adapted for faithful transmission [8, 9, 11, 16, 17, 22, 23]. Firing rates linearly code behavioural and task metrics 'at all levels of the cerebellar circuit' [24 p.239, citing 30 references].

The significance of the last two paragraphs is that neuroanatomy – normally seen as just a way to move data around – can be used as a computer, and we can model it without simplification and in detail because we have detailed information about anatomy. The mechanism is self-sufficient – it does not need outside help – and does not need to know anything about the information it is handling.

With this platform, it becomes possible to see cerebellar neurophysiology as the biological form of a network of linear units organised in layers, with direct correlation of units to the biological hardware. In its computational guise, a unit is a linear function. In its physical form, it can be a cell, part of a cell or some other structure. A unit randomly samples the afferent data layer (the outputs of the layer immediately upstream). Anatomy mimics random sampling of the whole afferent layer of firing

rates, with ‘replacement’, notwithstanding that a single unit receives unchanging contact from a tiny fraction of afferent-layer cells.

Sampling units, then, are not an abstraction or a simplification, but the physical hardware of a computational mechanism. This is key because it connects function to its physical form. The network model is derived from modern evidence, but comparative neuroanatomy [25-28] indicates that in its principles, main structures and cell types the brain has remained largely intact and unchanged through vertebrate evolutionary history. The hind brain, of which the cerebellum is part, may be the oldest part of the brain. In most vertebrates, the cerebellum is necessary for coordinated movements. In a marine environment, where it is not also necessary for skeletal muscles to resist gravity, this means almost exclusively the changing relative position of body parts. In the second half of the paper, we use the network computation to propose the principles that underpin the way a network maps to muscles – functional topography. We propose that networks are the junction of proprioception-motor loops, such that coordination of motor sequences is a passive but precise effect of data sharing by near-neighbouring networks. This can explain how the cerebellum emerged as a specialised structure to coordinate and propagate motor sequences, with little executive oversight, and on principles that may have been preserved in other vertebrate groups.

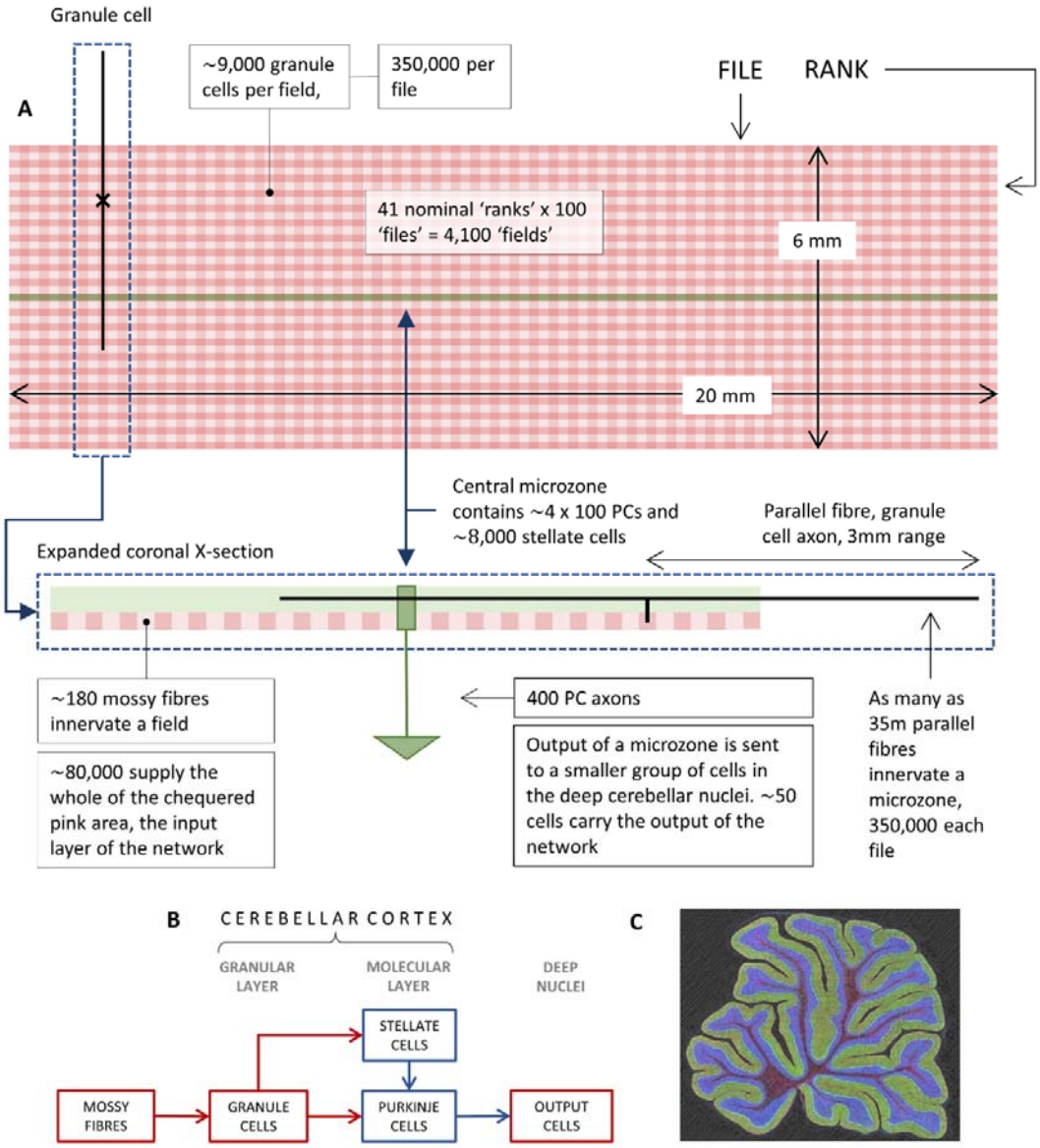


Fig 1. Graphic of the cerebellar network.

A View of a network from the cerebellar surface, with curvature and folding removed. Relative dimensions are preserved. Green line: a microzone (20 mm long x 150 µm wide x 300 µm deep, a functional division of the molecular layer of the cerebellar cortex and the middle layer of the cerebellar network). A microzone equally bisects the

region that provides parallel fibre input to it (hence 'middle microzone'). Chequered area: the region of the granular layer that provides parallel fibre input to the middle microzone (the input layer of the cerebellar network). A microzone has output to a discrete cell group in the deep cerebellar nuclei, which includes the output cells of the cerebellum (the output layer of the cerebellar network). Input to a microzone is from 40 'ranks' which lie parallel to it, 20 each side. Ranks are crossed at right angles by 75–100 'files' (100 in this graphic). Each intersection is a 'field'. Note that the granular layer is anatomically seamless; ranks, files and fields are nominal divisions. Networks overlap anatomically very substantially – neighbouring microzones receive almost all of the same parallel fibre input. PCs: Purkinje cells. **Inset** Cross-section: pink: granular layer; pale green: molecular layer; green: middle microzone. **B** Simplified cerebellar circuit wiring, highly preserved across vertebrates. Red: glutamate, an excitatory neurotransmitter; blue: GABA, and inhibitory neurotransmitter. The number of cells of each type in a network (given microzone length of 100 fields) is: 80,000 mossy fibres (not part of the network but supply input) → 35,000,000 granule cells (therefore the number of parallel fibres that innervate the middle microzone) → 8,000 stellate cells → 400 Purkinje cells → 50 output cells. **C** Reproduced with permission of Tom Deerinck. Section of rat cerebellum in the sagittal plane showing cortical layers. Green: molecular layer; blue: granular layer.

2. Terminology

A note on terminology. We divide the input layer of a network into nominal (200 x 150 µm) 'fields'. These are simply a device. They are not anatomically or indeed even functionally distinct or separate. A row of fields that spans the input layer sagittally is a

'rank' (a rank has the dimensions of a microzone seen from the cerebellar surface). A row that spans the input layer mediolaterally (therefore at right angles to ranks and microzones) is a 'file'. The microzone that equally bisects the input layer is the 'middle microzone'. Like fields, division into files is simply a device to quantify the ideas, and does not reflect physiological segregation of information. Figure 1 illustrates our terminology.

3. Input to a microzone

3.1 The physical form of code

As noted, we have previously modelled the granular layer computation. This generates the granule cell code [4]. Unlike, say, the wiring diagram of a house, space has a key role in the way information is coded in the cerebellum. Therefore, visualising the scaffold-like architecture of cerebellar cortex is essential. In this paper, for now, it is only necessary to visualise a microzone and parallel fibres that pass through it. A microzone is ~100 Purkinje cells long and 4 wide (microzone dimensions vary; 400 is consistent with reported convergence and divergence ratios onto 50 deep cerebellar nuclear cells [rats: 29]. Purkinje cells are severely flattened in the plane orthogonal to parallel fibres and to the cerebellar surface. 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) [30]. Microzones lie across the path of parallel fibres at right angles, so that parallel fibres pass through by the shortest distance.

Features of the granule cell code are:

(i) The same information is collectively coded in the firing rates of any random sample of active parallel fibres at the intersection with a sagittal plane. The intersection with a microzone approximates a sagittal plane.

(ii) The code spans a network so that a whole microzone receives, at any moment, at all locations, the same code. The form of the code is the frequency distribution of firing rates.

(iii) It is immaterial (codes nothing) which parallel fibres are active, in what pattern, or what particular rates they fire at individually.

(iv) Granule cells fire in short high-frequency bursts [20, 31, 32]. It is thought only a small fraction of parallel fibres is active at any time. But the constant high turnover of active cells is effectively an analogue code, because a Purkinje cell receives a large number of active inputs all the time.

(v) Active cells are randomly dispersed among the population, so activity has a uniform density but shifting pattern.

Key feature

The combination of this anatomy and the homogenised form of the granule cell code creates a facsimile of a statistical effect. The result is that it is as if all locations of a sagittal plane receive a random sample of the same distribution (frequency distribution of parallel fibre firing rates) with 'replacement'. Replacement is important so that sampling does not affect other samples – an example often used for illustration is that samples of coloured balls taken from a bag are replaced between samplings. Samples are said to be independent. The reason that is important is that there are statistical effects of sampling with replacement called the central limit theorem which do not otherwise occur. Obviously, replacement is not real because 'sampling' of parallel fibre firing rates by a plane, or by cells in a long row that approximates a plane, is simultaneous. It is also not physically possible because parallel fibre targets individually receive fixed contact from a tiny fraction of cells in the afferent cell layer (granule cells afferent to a microzone). Even a Purkinje cell receives contact from around only 0.5% of granule cells afferent to a microzone. However, the result mimics independent random sampling of the granule cell code because all locations of a microzone receive the same distribution.

3.2 The density of active parallel fibres is low and uniform

There is a long-standing prediction for theoretical reasons that the fraction of parallel fibres that are active at any time is regulated [33], fixed [34], and low (termed sparse code). We disagree with the traditional reasons, but agree the code is sparse, for these reasons. i) Low density is stable (Fig 2). Randomly fluctuating density would be noise. Stability may be enhanced by feedback via granule cell ascending axon contact on Golgi cell basal dendrites [22] and parallel fibre contact on Golgi cell apical dendrites, acting partly through gap junctions [35]. ii) Low density is energy efficient. Granule cells are very numerous and they are not myelinated (so spike propagation is costly). iii) The reported number of coincidentally active inputs to a stellate cell is in single figures [36], consistent with low density of co-active parallel fibres (see Table 1). iv) It is thought that the pattern of active and inactive parallel fibres, and the pattern in which firing rates are distributed among active cells, are decorrelated – effectively, quasi-randomised. Low density enhances practical decorrelation, that is, if decorrelation is not perfectly random – as it may not be [37] – low density increases the randomness of the pattern of synaptic activations received by targets.

3.3 The density of active parallel fibres is stable

We modelled regulation of parallel fibre density (Fig 2). Previously, we modelled a single rank [4]. The model simulates – in close anatomical detail – the conversion of signals received as input to a rank from around 2,000 mossy fibres into the number of granule cells that fire in each field and the rates they fire at. Here, we extend that to the whole input layer of a network, based on a 3 mm range of parallel fibres in each direction from the granule cell soma [38, 39] and a microzone dimensions of 150 µm x 2,000 µm viewed from the cerebellar surface (estimates and probably actual dimensions vary [2, 3, 40]). Otherwise, we used the same methods, described in Gilbert and Rasmussen (2024) [4]. A region of that size is crossed by 41 ranks, 20 either side of the middle microzone and one directly underneath (Fig 1), and at right angles by 100 files. The number of active granule cells in each field gives the number in each file. The number in a file is equal to the number of active parallel fibres at the intersection of that file with the middle microzone. Local regulation of the number of active granule cells in each field becomes still tighter regulation of the number of active parallel fibres in a file, a statistical effect of cerebellar cortical architecture (Fig 2, 3rd and 6th columns). Density is independent of the frequency distribution of mossy fibre rates. At low density, the regulated density of active parallel fibres is hyper focussed.

Local feedback via the granule cell ascending axon [22, 41] may condense the physiological range of field-to-field variation. Feedback via parallel fibres [34, 35] would (additionally) reduce the range of file-to-file variation. These are not represented in the simulation.

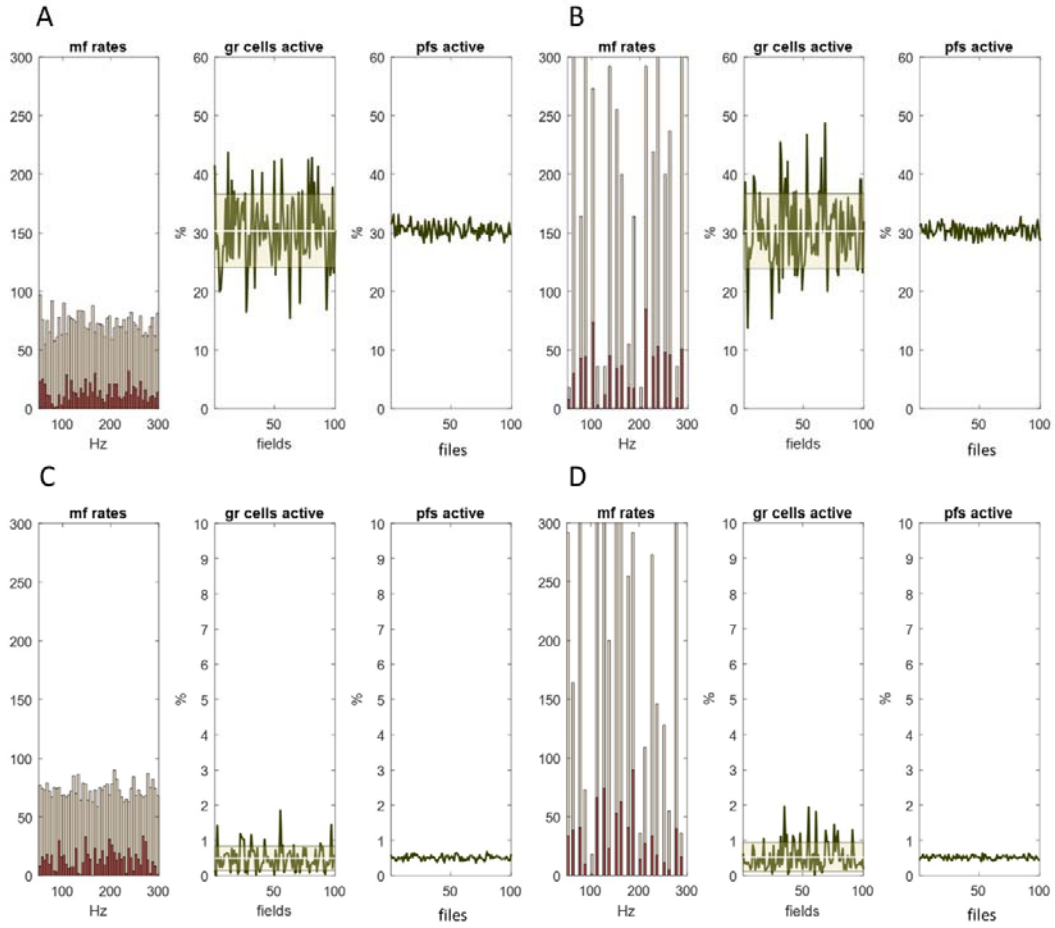


Fig 2. Recoding in the cerebellar granular layer regulates the fraction of active parallel fibres.

We simulated the conversion of input to the cerebellar network into granule cell signals using the method described here: [4]. The input layer of the network is crossed by 41 ranks (see Fig 1). The first and fourth columns show the frequency distribution of input rates to a single rank, selected at random. Input to each rank was generated independently. The conversion regulates the number of granule cells that are active at any time. The data in the second and fifth columns are the percentage of granule cells that fire in each field of a single rank, again selected at random. The data in the third and sixth columns are the percentage that fire in each file. This is therefore the percentage of active parallel fibres in each file that pass through the middle microzone. The shape of the sampled distribution (columns 1 and 4) does not affect regulation. 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 range of the percentage of granule cells active in each field (columns 2 and 5) is further narrowed when it is pooled with other fields in a file to give the percentage of active parallel fibres that converge on a microzone (columns 3 and 6).

4. The microzone computation

4.1 A statistical effect of random sampling

In our proposal, the cerebellum exploits statistical effects that depend on the independence of random samples. Say a population of values (with any frequency distribution) is randomly sampled with replacement. 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. Modest values of n or m give an untidier result. However, if the new distribution is then itself randomly sampled, and the process is repeated, the results get tidier and narrower. We propose that the neurophysiology of a microzone is the physical form of a network of sampling units organised in layers, and a unit is the physical form of a linear function, and that this set up exploits the statistical effect of random sampling with (a biological facsimile of) replacement. Layers are the length of the middle microzone and accordingly span a network from side to side. The result is to distil from the granule cell code – which is itself derived from 80,000 mossy fibre inputs to the network – a tightly focused effect on Purkinje cells in the middle microzone, which we infer causes firing of Purkinje cells to be closely aligned.

4.2 Stellate cells

4.2.1 Derivation of sample size

To test (by simulating it) if anatomy could in theory have the computational effect we propose, we first need the number of inputs to a stellate cell that are active at any time: sample size. The stellate cell territory is sagittally flattened in the plane orthogonal to parallel fibres. Stellate cells form planar networks [42] that occupy the

space (about 40 µm) between Purkinje cells [43]. 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 [43 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 [44]), 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 [43, 45], 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' [36 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

4.2.2 Sampling of parallel fibre rates by middle microzone stellate cells

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. 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 [46], meaning 80 stellate cells per field (there are four Purkinje cells per field 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 stellate cell in each middle-microzone field (Fig 3D).

Molecular layer interneurons reflect 'granule cell input with linear changes in firing rate' [20 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

we do not need it to demonstrate a statistical relationship of data layers, we use a normalised scale at this and in later steps of the computation. The result is a new data set with a smaller range and the same mean as granule cell rates. The true relationship of the means is the same as the function that relates the stellate cell firing rate to the mean of input rates.

4.3 Purkinje cells

In our simulation, a field contains 4 Purkinje cells. A Purkinje cell receives contact from stellate cells from both of its flat sides, ~16 each side, and a smaller number of stellate cells in the two sagittally-neighbouring fields. Inhibition from neighbouring fields is from a smaller number of stellate cells because not all of them extend their main axon far enough to reach, and some of them extend their main axon the wrong way. We represent those conditions as variables. The vertical range of stellate cell axon collaterals and the horizontal range of the main axon both increase with the depth of the cell body in the molecular layer [43, 47]. Because of their more limited horizontal range, outer-level stellate cells are excluded in the simulation from contacting Purkinje cells in neighbouring fields.

The planar Purkinje cell arbour is large and branches profusely, and stellate cell contact is dispersed across it. We divide the Purkinje cell arbour nominally into sixteen 50 x 75 µm compartments. For morphological reasons, input to a compartment is not confined

to (or even necessarily always received from) the nearest stellate cells, but each compartment is in range of most of the afferent network. As a result, each compartment receives contact from a random sample of stellate cells afferent to a Purkinje cell.

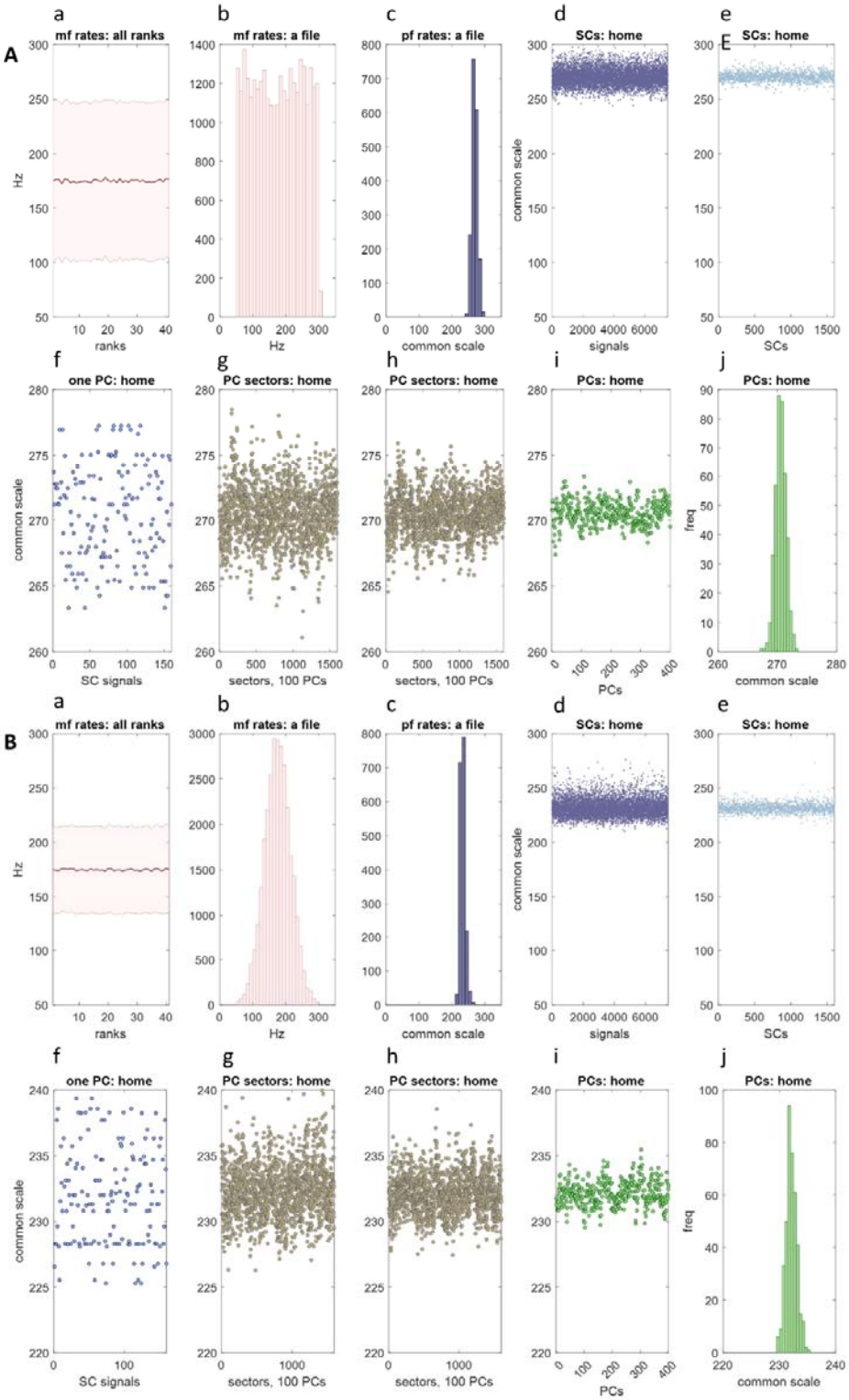
It is unknown how many stellate cells innervate a Purkinje cell compartment. We assume 5, ~20% of the afferent network, selected to be conservative (and therefore hypothesis neutral). Sampling is independent of sampling by other compartments, i.e., innervation of a compartment by a stellate cell does not exclude innervation by the same cell of other compartments. We represented Purkinje cell compartments by randomly sampling data for stellate cells afferent to a Purkinje cell.

Interneuron firing rates linearly reflect granule cell input [20]. Disynaptic inhibition via interneurons causes the firing rate of Purkinje cells in slices to decrease as a linear function of the strength of granule cell firing [21]. As there is also a linear relationship of Purkinje cell firing and direct synaptic excitation (with inhibitory synaptic transmission blocked) [18, 19], we took compartment membrane voltage as a linear function of stellate cell rates afferent to that compartment (derived by sampling). This created the next data layer, 6,400 values made up of 400 data sets, one for each Purkinje cell in the middle microzone (Fig 3G).

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 3H).

Stellate cells contact Purkinje cells on smooth dendrites [43]. 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' [20 p.9]. We represented Purkinje cell firing rates as the mean of compartment membrane voltages, generating the next-downstream data layer.

We have previously shown that mossy fibre rates are linearly coded in granule cell rates [4]. We show here that granule cell rates are linearly coded in Purkinje cell rates. If we are correct, a short sequence of computational steps is sufficient to focus an input range of hundreds of Hz down to a few Hz which is precisely predicted by mossy fibre rates upstream (Fig 3I and J).


A
a
mf rates: all ranks
Hz
ranks
b
mf rates: a file
Hz
c
pf rates: a file
common scale
d
SCs: home
common scale
signals
e
SCs: home
SCs
f
one PC: home
common scale
SC signals
g
PC sectors: home
sectors, 100 PCs
h
PC sectors: home
sectors, 100 PCs
i
PCs: home
PCs
j
PCs: home
freq
common scale
B
a
mf rates: all ranks
Hz
ranks
b
mf rates: a file
Hz
c
pf rates: a file
common scale
d
SCs: home
common scale
signals
e
SCs: home
SCs
f
one PC: home
common scale
SC signals
g
PC sectors: home
sectors, 100 PCs
h
PC sectors: home
sectors, 100 PCs
i
PCs: home
PCs
j
PCs: home
freq
common scale

380

Fig 3. The microzone computation.

Mossy fibre rates are converted into synchronised firing of middle microzone Purkinje cells. mf: mossy fibre; pf: parallel fibre; SC: stellate cell; PC: Purkinje cell. **A. a** We generated a uniform random distribution representing mossy fibre input rates to the entry layer of a cerebellar network using the method described here: [4]. The range, 50–300 Hz, is the typical physiological range for mossy fibres [48]. The range received by a network at any given time is probably smaller, and perhaps much smaller. Input to each rank was independently generated. Rate data were then randomly sampled, rank by rank, to derive input rates to each field, and copies added to reflect variable size of mossy fibre terminal clusters, and that most clusters straddle field boundaries, so lie partly inside and partly outside a field. Red line: mean; pink area: SD. **b** Frequency distribution of mossy fibre rates received by a randomly selected file. As all files cross all ranks, they all receive approximately the same frequency distribution of mossy fibre rates. **c** Parallel fibre rates in a randomly selected file. We used the granular layer model [4] to derive the number of granule cells that fire in each field and the rates they each fire at. As all files receive the same frequency distribution of mossy fibre rates, recoding in all files returns the same distribution of granule cell/parallel fibre rates. 0.5% of parallel fibres are active. 'Common scale' means normalised, discussed in the main text. **d** Parallel fibre rates in each file were randomly sampled 80 times, to represent sampling by middle microzone stellate cells. Sample size was a randomly generated (each time a sample was taken) number with Table 1 probabilities, giving around 7,200 values (16 stellate cells x 100 fields x the weighted mean of distributed probabilities). 'Home' in panel titles refers to the middle microzone. **e** Molecular layer interneurons are reported to reflect 'granule cell input with linear changes in firing rate' [20 p.6]. We took stellate cell firing rates to be the mean of input rates to each cell (hence 1,600 data points). There is a further contraction of the data range. **f** Data represent stellate cell input rates to a single Purkinje cell. The y-axis range (in this

panel and in g–i) is shortened (from 250 down to 20). **g** 'Sectors' in the panel title refers to the nominal division of the Purkinje cell dendritic arbour into 16 compartments, each receiving convergent input from a random sample of stellate cells afferent to the cell. Data are the sample means, shown for a single Purkinje cell per field in the middle microzone (so there are 1,600 values, 16 compartments x 1 Purkinje cell x 100 fields). **h** A Purkinje cell receives inhibition from two planar networks of stellate cells, one each side. Step G was repeated, and the data averaged for each compartment. **i** Mean of the compartment means calculated for each Purkinje cell in the middle microzone (so 400 data points, 4 Purkinje cells per field x 100 fields), representing somatic integration of dendritic compartment data. **j** The same data displayed as a histogram (note the x axis range is the same as the y axis in panels f to i, so much shorter than b and c). A uniform distribution of mossy fibre rates with a range of 50–300 Hz in panel a, becomes a very narrow Purkinje cell range with a normal distribution. **B.** We repeated A but with normally distributed mossy fibre rates (SD 40), with the same result. We also obtained the same results (not shown) with a discontinuous mossy fibre distribution (range 50–300 Hz).

4.4 Interim conclusion

Here, control of Purkinje cells is by the granule cell code, and therefore ultimately by mossy fibre rate information received as input to the network. The microzone computation is a passive effect, unaided and unlearned, of cellular and network anatomy.

In the classic cerebellar learning model [33, 34, 49], the chief instrument of the microzone computation is the learned modification of parallel fibre synaptic weights. Training causes Purkinje cells to return a learned response to inputs in a pattern received by trained synapses, displacing the naive response to firing rates. A granule cell code is effectively unnecessary.

We propose an alternative, where control of Purkinje cells is by a chain of rate codes passed in stages through a statistical effect of independent random sampling by each layer of the afferent data layer. There is no algorithm. We note also this. It is generally assumed that the representation of information in the brain is in high resolution and replicable with computer-like high fidelity, and that those features make it precise and allow neural network computations to be powerful. In our proposal, external signals are converted to a hyper low-resolution internal code which does not need to be replicable.

5. Parallel fibres coordinate movement

5.1 Proprioception-motor loops

Input drives output. In the traditional cerebellar model, it is unspecified where input is from but implicitly relegated to low importance because the response to input rates is

replaced with a learned response. The reason that linear units tend to be unpopular with modellers is that there is an assumption that network function is to bring about a computational transformation, so that, viewed in isolation, it is not obvious what a linear network would accomplish. We will argue that topography can explain it.

Cerebellospinal neurons provide 'a direct link between the cerebellum and spinal substrates for motor coordination' [50]. The output pathway is reciprocated by spinocerebellar neurons, providing a substrate for direct motor control, which spinal locomotor networks alone cannot accomplish reliably or with finesse [51]. Neurons of the spinocerebellar tracts are excited monosynaptically by group Ia and group Ib afferents that originate in muscle spindles and tendon organs respectively [52].

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). Cyclostomes – jawless fish – the nearest now-living relative of a common vertebrate ancestor, swim in this way. Modern cyclostomes have at best a rudimentary analogue of the cerebellum (hagfish) or, arguably, don't have one (lampreys), suggesting that a cerebellum might be exclusive to jawed vertebrates, and appeared after divergence of cyclostomes from the main vertebrate line. However, gene expression patterns which control the development of the embryonic neural tube – which develops into the vertebrate brain – have been

found for the cerebellum in both hagfish and lampreys [53-55], evidence that they were present in a common ancestor.

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. Thrust from a sagittally-flattened tail is absent or secondary.

Muscles generate information constantly, in muscle spindles, specialised muscle fibres. There are three types of muscle spindle fibres, two firing at rates proportional to muscle length, and the third signalling the rate of change of muscle length. Group II afferent neurons innervate the first two so they signal muscle length. Group Ia afferent neurons innervate all three, so they 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. Lateral flexion of the trunk during swimming causes muscle fibre length to vary sinusoidally, and therefore generates sinusoidally varying signal frequencies. The rate of change of muscle fibre length is therefore also sinusoidal since it is the derivative (the slope of a tangent to a sinusoidal waveform, derived by differentiation, is itself a sine wave).

We propose that in a vertebrate ancestor, input to a row of topographically-defined strips of the granular layer (ranks) arose from a corresponding row of contralateral myomeres, with neighbouring ranks receiving the output of neighbouring myomeres. During locomotion, at any moment, a graph of mossy fibre input rates to the input layer of a network – 40 or so ranks in our model – is therefore a sine wave or part of a sine wave (Figs 4 and 5). In the next moment, the wave advances slightly. The size of the portion of a sine wave represented by a snapshot of input to a network depends on the number of muscle segment pairs that fit into a full body wave. Moment by moment, input to a single rank passes through the cyclic output of a muscle, which is received continuously. Input to immediately neighbouring ranks is a little output of phase, and to next neighbouring ranks is a little more out of phase, and so on.

Output of a network is ipsilateral and to a single myomere, opposite the middle of the row that provides input. A network is therefore the junction of multiple input pathways and a single output pathway. 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.

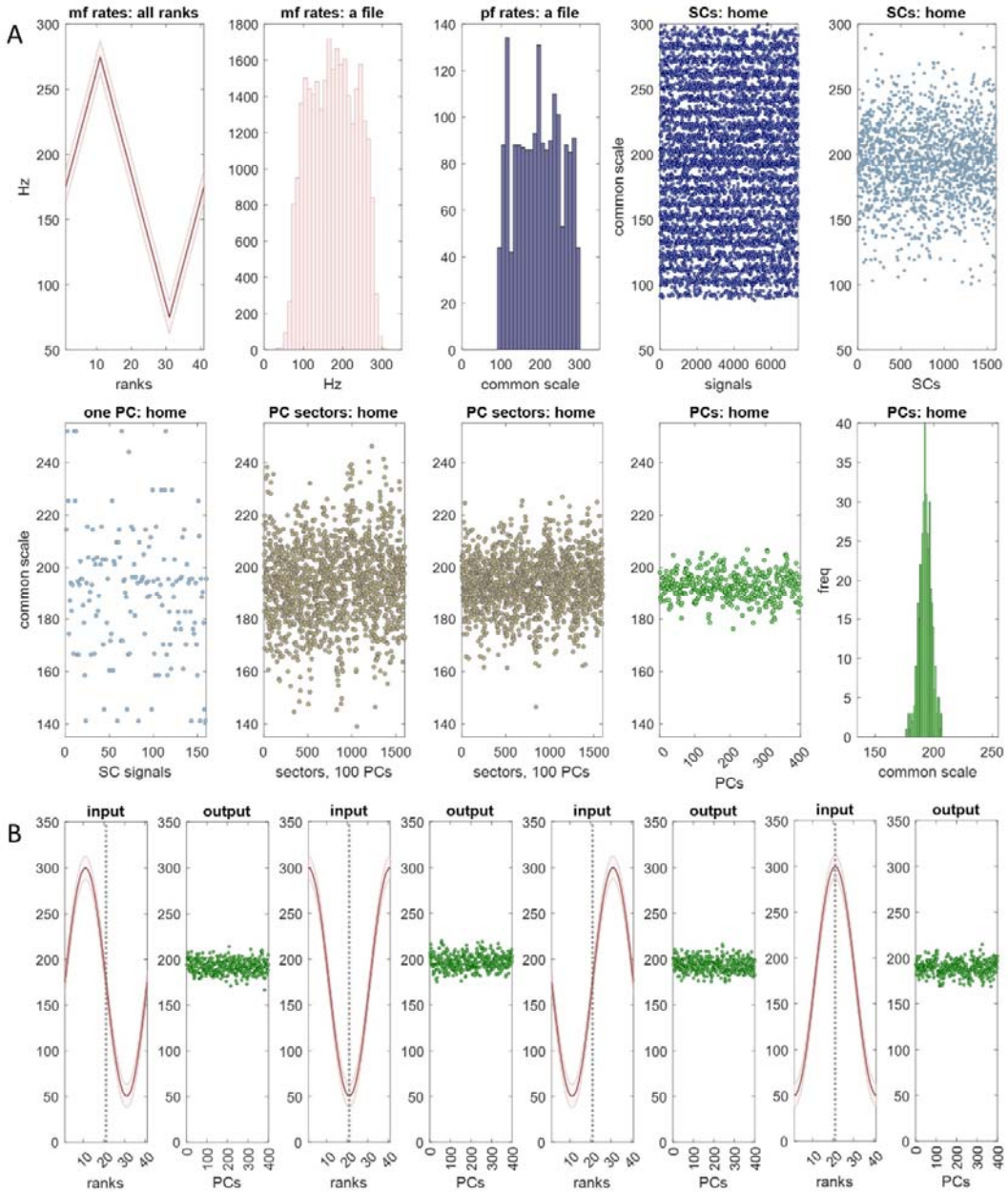


Fig 4. Information is coded in the 'shape' of mossy fibre input rates to a network, but not in all shapes.

A The same as Fig 3 except that the mean of mossy fibre input rates to each rank is varied from rank to rank, so that it has a zig-zag pattern displayed as a graph. The frequency distribution received by ranks individually is normal and narrow (SD = 12.5 Hz, all ranks). The frequency distribution of mossy fibre rates in a file that crosses at right angles is the same in all files (because they all cross all the same ranks) but not the same as ranks. That impacts on the spread of parallel fibre rates in top row panel 3 and therefore the spread of data in all later panels, so that Purkinje cell rates (middle row panels 4 and 5) are poorly focussed. The mean of the green data – the output of the microzone computation – still precisely tracks the mean of input rates to the network, as in Fig 3, but, critically, has no relationship with the pattern of input signals frequencies: the zigzag pattern of input frequencies cannot be inferred from green data. The significance is that the network computation cannot 'decode' the striped pattern of mossy fibre input rates. Note: banding in panel 4 is an artefact of using discrete data sets as input to ranks. Physiological ranks do not have clean boundaries. **B** We selected the zig-zag pattern as an illustration. However, the physiological pattern may be sine wave-like, or part of a sine wave, during a locomotor cycle. The vertical, dashed line is the position of the middle microzone. Green data are firing rates of Purkinje cells in the middle microzone. Performance does not improve, meaning: Purkinje cell firing rates do not code the striped pattern of mossy fibre firing rates. We advanced the phase of the cycle by 90 degrees each time, but with no effect on Purkinje cell rates.

5.2 Why would that generate smoothly modulated motor outputs with exactly coordinated timing?

The cerebellum contributes smooth execution and precise coordination to movements. The granular layer computation converts high-dimensional mossy fibre input to the cerebellum into a very low-resolution granule cell code [4]. Granule cell-encoded information is indivisible in ranks – it can only be divided into smaller pieces of the same information – so that information is striped viewed from the cerebellar surface. In that sense the internal code is pixellated. The striped pattern is literally short-lived, however. Data only retain their integrity – i.e., are only segregated – in granule cell ascending axons, while signals travel towards the cerebellar surface. Beyond that, after the granule cell axon divides in two, parallel fibre signals are randomly mixed up, within modal strata [31, 43, 56].

On the face of it, this only seems to spectacularly compound the loss of resolution available downstream to microzones. However, it is able to code phase and wavelength because of topography. [We propose that hyper-conflation of data is a *function* of data processing. There is fossil evidence of a proto-chordate with almost vertical muscle segments [57].] Modern adult fish have nested W-shaped myomeres with the centre pointing forwards [58]. The W shape spreads the force transmitted to the skeleton by muscle contractions. It also helps to reduce the forces acting on collagen membranes called myosepta which separate myomeres. Myomeres, like

skeletal muscle generally, are made up of muscle fibres. Muscle fibres are anchored to myosepta so that concerted contraction bends the body. Fibres are angled towards horizontal so that the force acting perpendicular to myosepta is less than the force of the contraction, without weakening the force that bends the body [59]. Analysis of the 3-dimensional pattern of fish muscle fibre angles is complex [60]. One interpretation is that they are arranged so that they contract 'to a fairly similar extent when the fish bends', delivering more power than contraction by different amounts [61].

Speaking generally, it is well established that neuron performance is noisy – there is inherent biological variability from cell to cell and even of firing of the same cell at different times in the same conditions. We might suppose, therefore, that firing rates which code the length of muscle fibres, and the rate they change length, are to some degree randomly variable. Also, muscle fibres that are differently angled, or different lengths, may not change length by the same amount or at the same rate.

The biological solution – we suggest – is to conflate the data using a statistical effect of random sampling. Output of a myomere is pooled for this purpose. This can't be done at spinal level, so it goes through the cerebellum. Input to the cerebellum from muscle fibres in a myomere, received by a granular layer rank, is accordingly handled with disregard for resolution below a topographically-defined scale, and converted to the pixel code. Pixel codes are in turn conflated by the microzone computation. The

network computation distils a single output rate, or output that approaches a single rate.

Output is sent to motor neurons. '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] Axons of spinal neurons that innervate and excite a motor neuron pool branch terminally, making contact indiscriminately on neurons inside.

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 of a network and a myomere. 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 motor neurons to be recruited, including larger ones, and therefore more muscle fibres, including fibres that contract with more force.

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

In anguilliform swimming, a wave of muscle extension passes down each side of the body, 180 degrees out of phase. Signals representing muscle fibre length and rate of extension are streamed to the contralateral cerebellum, where they are received by a corresponding set of ranks, pixellated, blended in parallel fibre modal strata, and passed through the middle microzone computation. Each microzone sends its output to an ipsilateral motor neuron pool. By this pathway, contralateral muscle fibre extension drives proportional ipsilateral muscle contractions. Sensory and motor signals are a response to each other in a spiral that causes a wave of contractions to pass backwards along the body. Alternating left and right head turns initiate backwards propagating waves 180 degrees out of phase, assisted by passive undulation.

5.3 Parallel fibres

It is a long-standing puzzle why neighbouring microzones share such a high percentage of parallel fibre inputs. Next door neighbouring microzones receive almost *all* of the same parallel fibre inputs (~97.5% assuming a microzone span of 150 µm). In the present proposal, this has central importance. It means motor output of immediately neighbouring microzones is very precisely coordinated because they receive almost all the same parallel fibre input, and control of output is by a chain of rate codes and not displaced by learning. Coordination is in this way automated, simply the outcome of anatomy. The same relationship exists with more distant neighbours of a network in both directions to a degree that is inversely related to the distance. Output is

streamed, so coordination of output rates provides coordination of the timing of motor outputs and automates smooth execution of performance.

Pooling information has a cost. For precision, input to a microzone must be exactly phase-sensitive. The input layer of the cerebellar network receives inputs from a row of contralateral muscle segments, diluting phase-specificity. Taken to an extreme, if the number of segments providing signals is the same as the number in a full wave of the body, the information is perfectly imprecise – there is no phase information at all (Fig 4). However, if muscle segments are thin, so that signals informing motor output originate in a smaller length of the body relative to wavelength, phase information emerges. The lower the ratio (of sector length to wavelength), the more precise it is (Fig 5).

There are in addition at least two strategies that help focus a phase-specific response. In the first, there is a gradient. Parallel fibres make about half as many synapses distally than proximally, and swellings where contact is made contain fewer synaptic densities [62]. As a result, there is a stronger influence of nearer granule cells, and therefore mossy fibre signals originating in nearer myomeres.

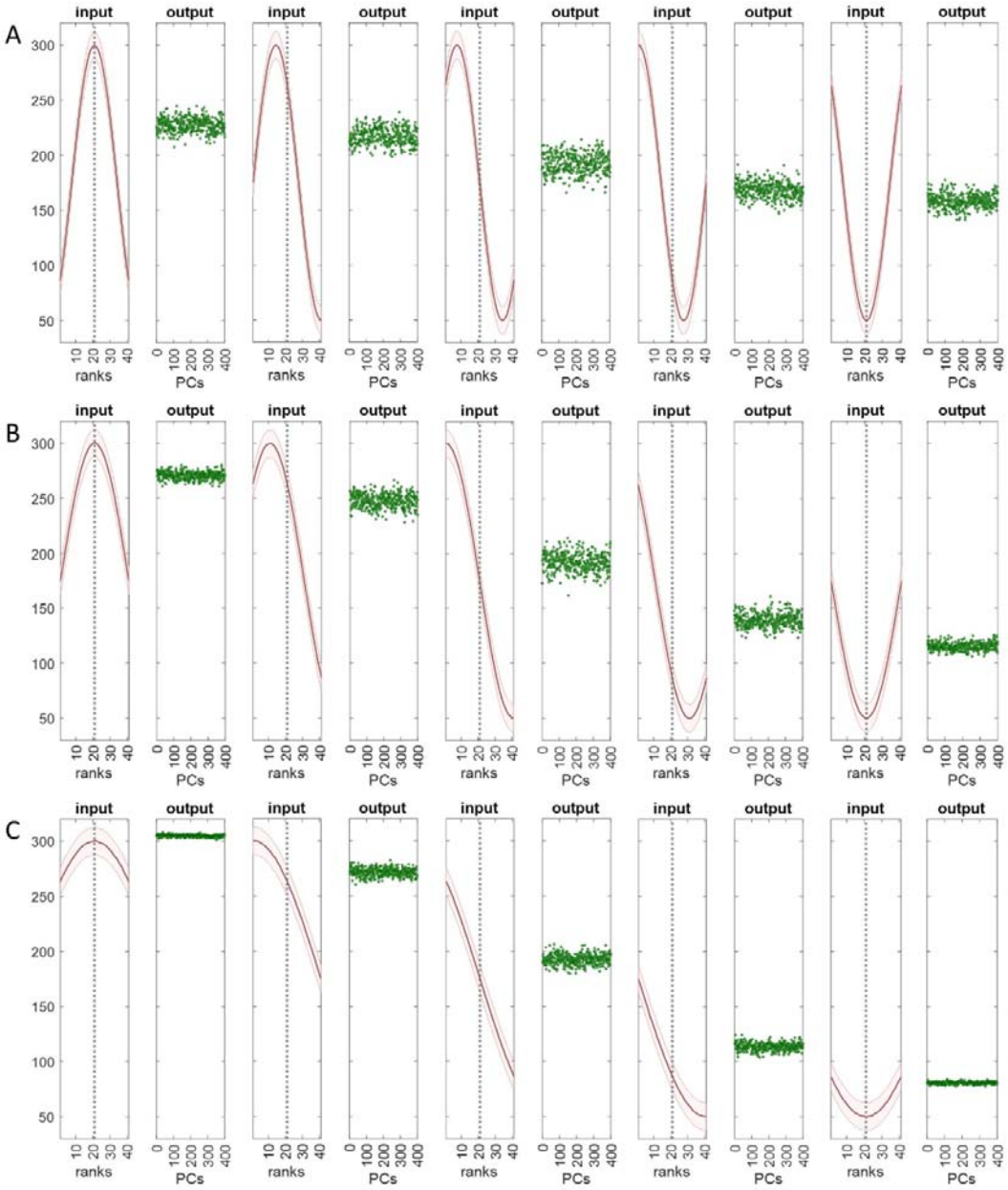


Fig 5. How does the middle microzone read the striped pattern of mossy fibre input rates to a network?

Phase information becomes readable if input to a network is confined to only part of a cycle. **A** The same as Figure 4B but mossy fibre rates represent 3/4 of a wavelength, phase shifted 45 degrees between the first pair of panels and the second pair, the second pair and the third pair, and so on. Output now reflects cycle phase, but the response is insensitive, i.e., output has a modest range compared to the range of mossy fibre rates. Also, Purkinje cell firing is not focussed. **B** The same, but mossy fibre rates represent 1/2 a wavelength. The size of the output range is larger, ~150 normalised 'Hz', twice the range in A. Purkinje cell data are more focussed. **C** Mossy fibre rates represent 1/4 of a wavelength. The bandwidth of the output range increases again, to almost the mossy fibre range. That is, output reliably and precisely codes phase information. Purkinje cell firing rates are further focussed. The smaller the fraction of a wavelength that mossy fibre signals code, the more sensitive and precise the information coded collectively by Purkinje cells.

The second is a result of granule cell T-shaped morphology. Purkinje cells receive significant granule cell ascending axon contact [63]. Ascending axon synapses are resistant to long-term depression, and have low uncertainty of release probabilities [64]. Accordingly, excitation of Purkinje cells receives phase weighting from ascending axons of directly underlying granule cells.

A feature we like is that parallel fibres of standard length maintain the amplitude of body undulations as they pass backwards. This is efficient because the whole body rides the herringbone pattern of high pressure created by head turns, not achieved by caudally diminishing amplitude of passive undulations.

5.4 Conclusion

A striking feature of vertebrate brain evolution is that it seems to have changed much less than body form and physiological strategies that were necessary to colonise new habitats [25, 26, 53-55]. Complex divisions of the vertebrate brain appeared more than 500 million years ago. It remains poorly understood (i.e., not understood) how the cerebellum controls movement. We don't have a wiring diagram that shows how it is connected around the body. But the highly conserved organisation of the brain makes it a useful investigative tool. Modern data provide a potential line of enquiry into the early brain and, conversely, evolutionary theory can provide insight into modern motor control.

Originally, head turns to and fro may have caused passive undulation of a flexible pre-chordate body (modern anguilliform swimming remains in part passively powered). Head turns may have been chemotactically-motivated, with net movement towards a gradient of increasing chemical concentration. Contractile fibres behind the head aggregated into neck muscles to increase force. Muscles were bilaterally paired, restricting bending to a single plane to exploit pressure gradients (flagellate wriggling is inefficient). Extension of a muscle was used to power contraction of its antagonist, so that it did not need to be directly powered by sensory information. Reflex arcs were unsuitable to process the traffic because the messy data coming out of a muscle needed to be cleaned up. An enlargement of the dorsal brain stem formed near the

cervical muscles that were sending and receiving the data. Muscle data were segregated – communication between a pair of muscles was in a dedicated rostro-caudal strip – but inside a strip it was pooled to distil a clean signal. Distillation was the passive effect of anatomy and linear functions, as described. Purkinje cells were flattened to fit. In a later development, parallel fibres extended outside. Data sharing via parallel fibres improved coordination. The availability of hind brain processing allowed more muscle pairs to be added, extending backwards, because lengthening of parallel fibres caused neck bends to propagate.

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.

Declaration of interests

The authors declare no competing interests.

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. Gilbert, M. and A. Rasmussen, *The cerebellum converts input data into hyper low-resolution 3-D pixels. Who knew?* bioRxiv, 2024: p. DOI: 10.1101/2023.07.14.548987 version 8, https://www.biorxiv.org/content/10.1101/2023.07.14.548987v8.
5. 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.
6. Rancz, E.A., et al., *High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons.* Nature, 2007. **450**(7173): p. 1245-8.
7. Ritzau-Jost, A., et al., *Ultrafast action potentials mediate kilohertz signaling at a central synapse.* Neuron, 2014. **84**(1): p. 152-163.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. Duguid, I., et al., *Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex.* J Neurosci, 2012. **32**(32): p. 11132-43.
14. 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.
15. Galliano, E., P. Mazzarello, and E. D'Angelo, *Discovery and rediscoveries of Golgi cells.* J Physiol, 2010. **588**(Pt 19): p. 3639-55.
16. 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.
17. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptotagmin 7 confers frequency invariance onto specialized depressing synapses.* Nature, 2017. **551**: p. 503.

18. Walter, J.T. and K. Khodakhah, *The linear computational algorithm of cerebellar Purkinje cells.* J Neurosci, 2006. **26**(50): p. 12861-72.
19. 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.
20. Jelitai, M., et al., *Dendritic excitation-inhibition balance shapes cerebellar output during motor behaviour.* Nat Commun, 2016. **7**: p. 13722.
21. Park, S.M., E. Tara, and K. Khodakhah, *Efficient generation of reciprocal signals by inhibition.* J Neurophysiol, 2012. **107**(9): p. 2453-62.
22. Cesana, E., et al. *Novel granule cell-Golgi cell excitatory input in the cerebellar granular layer*. in *FENS Abstr*. 2010.
23. 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.
24. Raymond, J.L. and J.F. Medina, *Computational Principles of Supervised Learning in the Cerebellum.* Annu Rev Neurosci, 2018. **41**: p. 233-253.
25. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from Myxinoids through Birds*. 1967: The University of Minnesota Press.
26. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from Monotremes through Apes*. 1970: The University of Minnesota Press, Minneapolis.
27. Bell, C.C., *Evolution of cerebellum-like structures.* Brain Behav Evol, 2002. **59**(5-6): p. 312-26.
28. Bell, C.C., V. Han, and N.B. Sawtell, *Cerebellum-like structures and their implications for cerebellar function.* Annu Rev Neurosci, 2008. **31**: p. 1-24.
29. Person, A.L. and I.M. Raman, *Synchrony and neural coding in cerebellar circuits.* Frontiers in Neural Circuits, 2012. **6**: p. 97.
30. 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.
31. 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.
32. 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.
33. Marr, D., *A theory of cerebellar cortex.* J Physiol, 1969. **202**(2): p. 437-70.
34. Albus, J.S., *A Theory of Cerebellar Function.* Mathematical Biosciences, 1971. **10**(1-2): p. 25-61.
35. Gilbert, M. and A. Rasmussen, *Gap Junctions May Have A Computational Function In The Cerebellum: A Hypothesis.* The Cerebellum, 2024.
36. 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.
37. 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.

38. 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.
39. Mugnaini, E., *The length of cerebellar parallel fibers in chicken and rhesus monkey.* J Comp Neurol, 1983. **220**(1): p. 7-15.
40. 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.
41. 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.
42. Häusser, M. and B.A. Clark, *Tonic synaptic inhibition modulates neuronal output pattern and spatiotemporal synaptic integration.* Neuron, 1997. **19**(3): p. 665-78.
43. Palay, S.L. and V. Chan-Palay, *Cerebellar cortex : cytology and organization*. 1974, Berlin: Springer.
44. 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.
45. Carter, A.G. and W.G. Regehr, *Quantal events shape cerebellar interneuron firing.* Nat Neurosci, 2002. **5**(12): p. 1309-18.
46. 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.
47. 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.
48. 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.
49. Fujita, M., *Adaptive filter model of the cerebellum.* Biol Cybern, 1982. **45**(3): p. 195-206.
50. Sathyamurthy, A., et al., *Cerebellospinal Neurons Regulate Motor Performance and Motor Learning.* Cell Rep, 2020. **31**(6): p. 107595.
51. Parker, D., *Complexities and uncertainties of neuronal network function.* Philos Trans R Soc Lond B Biol Sci, 2006. **361**(1465): p. 81-99.
52. 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.
53. Sugahara, F., et al., *Reconstructing the ancestral vertebrate brain.* Dev Growth Differ, 2017. **59**(4): p. 163-174.
54. Sugahara, F., et al., *Genetic Mechanism for the Cyclostome Cerebellar Neurons Reveals Early Evolution of the Vertebrate Cerebellum.* Frontiers in Cell and Developmental Biology, 2021. **9**.
55. Sugahara, F., et al., *Evidence from cyclostomes for complex regionalization of the ancestral vertebrate brain.* Nature, 2016. **531**(7592): p. 97-100.
56. 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.

57. Morris, S.C. and J.B. Caron, *Pikaia gracilens Walcott, a stem-group chordate from the Middle Cambrian of British Columbia.* Biol Rev Camb Philos Soc, 2012. **87**(2): p. 480-512.
58. Tytell, E., *Fish muscle is still helical.* Journal of Experimental Biology, 2003. **206**(15): p. 2530-2530.
59. Lacalli, T., *The Middle Cambrian fossil Pikaia and the evolution of chordate swimming.* Evodevo, 2012. **3**(1): p. 12.
60. Gemballa, S. and F. Vogel, *Spatial arrangement of white muscle fibers and myoseptal tendons in fishes.* Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 2002. **133**(4): p. 1013-1037.
61. Alexander, R.M., *The orientation of muscle fibres in the myomeres of fishes.* Journal of the Marine Biological Association of the United Kingdom, 1969. **49**(2): p. 263-290.
62. 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.
63. 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.
64. 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.