THE CEREBELLUM CONVERTS INPUT DATA INTO HYPER LOW-RESOLUTION 3-D PIXELS. WHO KNEW?

Mike Gilbert[a] (corresponding author), Anders Rasmussen[b]

[a]School of Psychology, College of Life and Environmental Sciences, University of Birmingham B15 2TT, UK

ORCID: 0000-0003-4415-6915

[b]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

Our scope is the inner layer of the cerebellar cortex, the granular layer, where input to the cerebellum is received. We ask how input signals are converted into an internal code, what form that has, and how it solves some fundamental problems that come with using neurons to represent information. While there is a computational element, and the ideas are quantified with a computer simulation, the approach is primarily evidence-led and aimed at experimenters rather than the computational community. Network models are often simplified to provide a noiseless medium for sophisticated computations. We propose, with evidence, the reverse: physiology is highly adapted to provide a noiseless medium for straightforward computations, which are understandable with general science numeracy. We find that input data are converted to a hyper low-resolution, pixel-like internal code which spans the cerebellar network side-to-side. A pixel is the smallest unit of information but not the smallest unit of code – the same information is coded in any random sample of signals. Code in this form mitigates inconsistent performance by single neurons and synapses, explains how the cerebellum can work 'blind' (below a topographically-set scale threshold), and converts untidy input rates into coordinated firing of large cell groups – microzones – downstream.

Keywords: cerebellum; theory; model; granular layer; network; code

1. Introduction

It is generally assumed that brain-coded information is rich and replicable with computer-like high fidelity, and that those features make it precise and allow computations to be powerful. At least, that is not contentious. In many neural network learning models, including cerebellar models, computations are implemented by learned synaptic changes that selectively adjust synaptic transmission strength. This has been a mainstay of much modelling for 50 years. We propose that in the cerebellum, plausibly, all of that may not be true.

The inner layer of the cerebellar cortex, the granular layer, receives the main excitatory input to the cerebellum, from mossy fibres. Mossy fibres contact both granule cells and large inhibitory interneurons, Golgi cells. Golgi cells in turn inhibit granule cells. The granule cell axon rises into the outer layer, the molecular layer, where it divides in two to form parallel fibres, which lie parallel to the cerebellar surface and other parallel fibres. Parallel fibres make contact in passing on Purkinje cells, which carry the sole output of the cerebellar cortex.

We attempt to explain anatomy in the light of evidence of linear signalling. By anatomy we mean cell morphologies and connectivity, and bespoke adaptations of synaptic

contact. Inevitably, this involves attention to anatomical detail. The principles, however, address generic problems that come with using nerve cells to code information.

Speaking generally, neuron-to-neuron transmission involves the complex interaction of a large array of often co-dependent biophysical and other variables which are specific to the cell types making and receiving contact. With all the noise, it is difficult to imagine how information is represented or moved around without catastrophic corruption of data. Yet despite the seeming clamour, linear relationships are in fact conserved. There is now a substantial body of evidence that rate-coded information is linearly conserved in communication between cerebellar neurons [1-8], and that cells are adapted for faithful transmission [9-15]. Firing rates that linearly code behavioural and task metrics have been observed 'at all levels of the cerebellar circuit' [16 p.239, citing 30 sources]. There are a number of examples that the cerebellum is adapted to disengage or otherwise compensate for intrinsic non-linear properties of neurons and synapses [13, 14, 17, 18]. This is – we submit – part of a general strategy of isolating the effect of parameters that code information while eliminating or mitigating interference by other variables. Evidence of linear signalling in the granular layer is particularly rich. We include a review in the Methods section to which we cross refer from the main text.

We will argue that the granular layer is in effect the physiological form of a network of linear units which randomly sample firing rates in the afferent cell layer, within topographically-defined boundaries. As sampling units have identified anatomical counterparts, anatomy provides model parameters. We field evidence in the main text to explain the derivation of model parameters. Since cerebellar cell morphologies and connections are well studied, parameter values (usually a range) are available from the literature. We state in the next section the proposal that biology simulates a network of linear units. We expand on that in statements at or near the beginning of each section, headed 'Hypothesis', then describe (in each section) how it was tested, and the results.

Classically, neural network models solve a problem. The nature of the problem – and therefore the function of the network – is a proposal and becomes an assumption of the model. The solution is a computational mechanism; the mechanism is the model. The direction is function → mechanism (the first two steps of Marr's three levels of analysis; the third step is to consult the evidence to see if the known physiology could provide the physical hardware to implement the mechanism). We work in reverse. We propose first a mechanism that can explain the evidence, then infer function from the mechanism. High-level ideas come last. We use modelling in a different role, to quantify and test the ideas by simulating the mechanism, to verify that it works with physiological parameters. The direction is evidence → mechanism → function. This is simply a different approach and not in competition with other methods. However, it is

important (and we raise it here) because while there is a computational element, and the ideas are quantified with a computer simulation, it is not a computational paper. Simply, as firing rates represent values, we describe – what we think are – the code and computations, but as they are straightforward, they can be described in words and require only general science numeracy.

To aid navigation, our key findings are: **1.** Anatomy (cell morphologies and wiring) performs a passive computation unaided and unlearned. **2.** There is no algorithm. **3.** Physiology is the physical form of a network of linear units (meaning, output of a unit is a linear function of inputs). A unit can be a cell, part of a cell or some other structure, but not a single synapse. **4.** External signals are converted into a low-resolution internal code that occupies volume – 3-D pixels. Accordingly, internally-coded information has spatial dimensions, the dimensions of cell groups that code together (code groups). **5.** Code groups span the cerebellar network from side to side (sagittal direction, mirroring division of the outer layer of the cerebellar cortex into microzones). **6.** Information is coded in the combined activity of code-grouped cells, and code groups communicate by a facsimile of independent random sampling, for which physiology is highly adapted. **7.** The code is analogue, under modulation of continuous input during behaviour. **8.** Coding and data processing exploit statistical effects of random sampling. **9.** A pixel (the information represented by a code group) is the smallest unit of information. Inside pixel dimensions, information is indivisible – i.e., any smaller number of randomly selected cells (over a minimum number) still

represents the same information. **10.** By the same token, a pixel is not the smallest unit of code – the same information is coded in any random sample of firing rates in a code group. **11.** The form of the code is key to function. It has these results. (i) Neural representation of information does not depend on replicability of neural activity (a solution to the problem that input data are 'unreliable' in the sense that firing of single neurons may not code information consistently or precisely, and synaptic transmission involves random variables). (ii) The cerebellum can work 'blind' below a topographically-set scale threshold (because it does not need to differentiate between signals below pixel scale). (iii) It converts input rates – with any frequency distribution – into coordinated firing of large cell groups – microzones – downstream (because target cells can receive the same code without receiving the same signals).

2. Linear units

The cerebellum is the physical implementation of a computational network of linear units. Linear refers to the conversion function – output is a linear function of inputs, such as the weighted sum or mean of inputs. A unit can be a cell, part of a cell or some other structure, but never a single synapse. Units form layers that data are passed through. This happens continuously. Data are coded, at any instant, in a snapshot of the collective output of a layer. Where – as in most cases – units receive as inputs

values represented by firing rates in the afferent layer, they randomly sample afferent rates. Anatomy mimics random sampling of the whole cell layer afferent to a group (we will explain how). The result is a facsimile of random sampling by individual units of the entire afferent population of firing rates, with replacement, notwithstanding that a single unit receives unchanging contact from a tiny fraction of cells in the afferent layer. Cell morphologies and network geometry are explicable as bespoke adaptations for this purpose.

Random sampling, we will argue, is inside spatial dimensions that are functionally defined by anatomy, although they are not physically bounded. The result is that firing rates of successive layers are linearly related by rate statistics at collective level, at that scale. Signalling is analogue, so that signalling and computations are continuous during behaviour.

3. Mossy fibres

3.1 Anatomy, 'ranks' and 'files'

The molecular layer of the cerebellar cortex is divided functionally into long, thin strips called microzones [19]. Mossy fibres give rise to sagittally-extending collaterals, the direction of the long axis of microzones (Fig 1). These in turn give rise to multiple distal

branches, each ending in a cluster of terminals [20-23]. Terminal clusters are sagittally aligned and separated by a randomly variable distance typically not less than 200 µm [20, 21] measured from edge to edge of clusters (so they are separated by a greater distance measured from centre to centre). Terminals themselves are intimately and randomly intermingled with terminals of other mossy fibres.

Accordingly, mossy fibres do not terminate at a point but a region with dimensions: the field occupied by a terminal cluster. A single mossy fibre gives rise to several terminal branches and each of those to several terminals, such that a sagittal strip of the granular layer receives multiple copies of every signal at each of multiple locations.

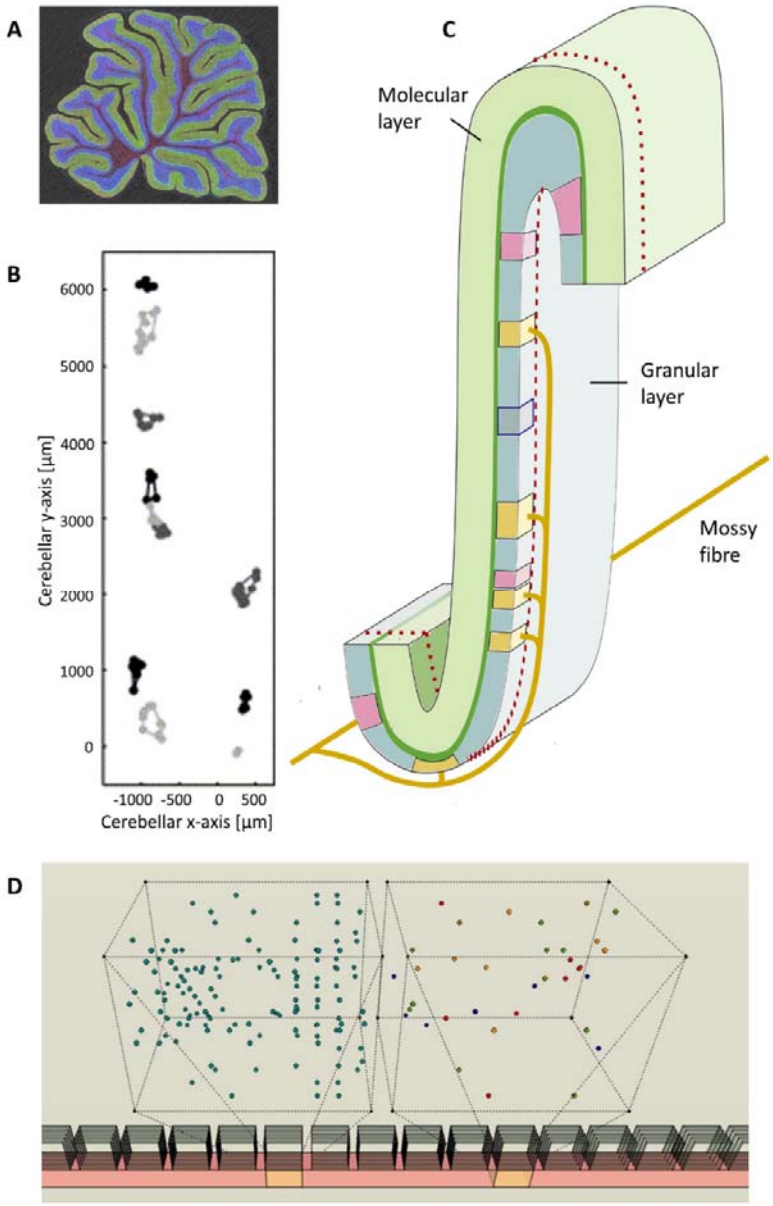
Cluster size is variable (because it depends on the number of terminals and how tightly clustered they are), averaging 200 µm sagittally x 150 µm mediolaterally [21]. We term the average dimensions of a cluster a 'field'. The granular layer is anatomically seamless. A field is simply a nominal division that we use for modelling. A sagittal row of 75–100 fields laid end to end has the area (viewed from the cerebellar surface) of a mid-range microzone [19, 24-26]. We term a row of fields a 'rank'. Ranks are functionally but not anatomically discrete. Ranks are parallel to microzones but they are divisions of the granular layer, while microzones are divisions of the molecular layer.

Each terminal is ensheathed by a semi-permeable membrane which restricts neurotransmitter diffusion, a structure taken together termed a glomerulus. Excitation

172 of granule cells and Golgi cells by mossy fibres, and inhibition of granule cells by Golgi

173 cells, all take place here.

174



175

Fig 1. Mossy fibre termination. A Sagittal section of the rat cerebellum (obtained by multi-photon microscopy, courtesy Tom Deerinck). The outer layer of the cerebellar cortex, the molecular layer, is stained green and the inner layer, the granular layer, is blue. **B** Output of a procedure used to group mossy fibre terminals into clusters [21 fig 6], reproduced with permission. All terminals belong to one mossy fibre, which terminates in two strips. Clusters are shown in different shades of grey. **C** Schematic of an area of the folded cerebellar cortex. The strip inside the red dotted line is the width and depth of a microzone but only part of its length. Yellow zones are terminal locations of a single mossy fibre. Pink zones are terminal locations of a second mossy fibre. In reality, zones overlap extensively and terminals are intermingled. The zone enclosed by a blue line has the dimensions of a 'field', a nominal division of a rank with the average dimensions of a mossy fibre terminal cluster. **D** Section of a microzone viewed obliquely from the cerebellar surface, to illustrate the size and random intermingling of mossy fibre terminals. The wire frame boxes are an exploded view of the orange fields. The right box shows randomly distributed terminals of 5 mossy fibres (each a different colour). Proportions of terminal size to field size are preserved. The left box shows terminals of 20 mossy fibres (out of ~180 that contribute terminals to a field).

3.2 Hypothesis

Mossy fibre anatomy is an anatomical solution to a computational problem. The cerebellum exploits statistical effects of random sampling called the central limit theorem. For the effects to occur, samples must be independent, also termed sampling

199 ‘with replacement’ (in the illustration sometimes used, coloured balls selected at
200 random from a bag are put back in the bag between samplings). Clearly, physical
201 replacement is impossible because neither a terminal cluster nor a terminal can be in
202 more than one place at the same time. However, anatomy simulates replacement.
203 Terminal branching simulates sampling with replacement by fields in a rank of input
204 rates to the whole rank – the distributed probabilities of the number of terminal
205 branches per mossy fibre are the same as the probabilities of the number of repeat
206 samplings if rates received by a rank were in fact independently randomly sampled by
207 fields. Terminal clustering simulates sampling with replacement by cells in a field of
208 input rates to the field – the distributed probabilities of the number of terminals per
209 cluster are the same as the distributed probabilities of the number of repeat samplings
210 if rates received by a field were in fact randomly sampled with replacement by
211 individual target neurons. The combined effect is a biological facsimile of sampling
212 with replacement by individual neurons of firing rates received as input to a whole
213 rank, notwithstanding that they receive fixed contact from a tiny fraction of afferent
214 cells. Note that it is the population of *rates* that are hypothetically sampled with
215 replacement, and not *signals.* The population of signals includes copies of rates
216 because of branching and clustering.

217

218

3.3 Terminal branching

If each mossy fibre could terminate in only one field, a mossy fibre 'sampled' by one field clearly cannot also be sampled by any others. Termination by a single mossy fibre in a random number of terminal branches, at multiple locations, removes that restriction.

In theory, sampling with replacement should mean any number of fields can receive the same signal, up to all of them. Clearly this is not true – there is still an anatomical limit. But in practice even with real replacement the number of repeat samplings is constrained by a probability distribution. If, then, the number of terminal branches per mossy fibre has the same probability distribution, branching has the same result as if firing rates received by each field in a rank were in fact a random sample with replacement of firing rates received by the rank as a whole.

We can test that by calculating the probability distribution in the hypothetical case that mossy fibres did not branch terminally and rates were in fact sampled with replacement. The probability, p, that a given rate is selected k times in taking N samples, sample size x, assuming mossy fibres each give rise to an average of m terminal branches, is given by

$$\left[{Nx!}/{k!\,(Nx-k)!}\right] * \left[{1}/{(Nx/m)}\right]^{k} * \left[1-\left({1}/{(Nx/m)}\right)\right]^{Nx-k}$$

The results are shown in Table 1.

Table 1. Probability p of k repeat selections of a single mossy fibre

k	0	1	2	3	4	5	6	7	8	9	10	11
p	0.007	0.034	0.084	0.140	0.175	0.175	0.146	0.104	0.065	0.036	0.018	0.008

Although data are scarce, the number of repeat selections is consistent with what we know about the biological range of the number of terminal clusters per mossy fibre [21, 27]. Estimated by this method, the mean is 4–5.

3.4 Terminal clusters

The principle is the same here but at local, field scale. Mossy fibre terminal clustering, and random intermingling with terminals of other mossy fibres (Fig 1), means that contact on any particular cell does not mean the same mossy fibre cannot also contact cells somewhere else. Again, to approximate independence, it is only necessary for the number of repeat selections of a signal to have the same probability distribution as if there was actual replacement.

A practical problem with the strategy of clustering is that a target could receive the same signal twice (or more). The risk is mitigated by taking small samples. A sample

size of 4 is small enough for a low risk. Granule cells average 4 dendrites, each receiving contact from a single terminal, and we estimate that Golgi cells receive an average of 4 inputs to each of their basal dendrites (see derivation of parameters in the next section). Given m x 100 terminals and an independent and equal chance that contact is by any terminal, the probability of an nth contact by the same mossy fibre on a single dendrite is:

$$P_n = \left[\prod_{x=2}^{n}\left(\frac{m-x+1}{100m-x+1}\right)\right] \times \frac{y!}{x!\,(y-x)!}$$

where y is the number of inputs to a single dendrite and m is the number of terminals per cluster, provided by (for illustration) 100 mossy fibres. Calculated in this way, around one in every twenty samples contains a duplicate, and less than two in a thousand contain more than one. The actual fractions are smaller because the number of mossy fibres that innervate a field is higher, around 180 (see Methods, Derivation of mossy fibre rates).

Quasi-sampling by each field of mossy fibre firing rates received as input to the rank as a whole, and by cells in each field of rates received by that field, is a biological facsimile of simultaneous random sampling with replacement by single neurons of input rates to a whole rank.

4. Golgi cells

4.1 Hypothesis

Anatomical detail is critical to computational performance and the functional division of the granular layer into groups of cells that work together. Despite forming an anatomically seamless carpet of cells, Golgi cells work in groups: ensembles. An ensemble is the population of a sagittal row of three fields. Groups are not anatomically bounded or even functionally discrete, but function as a single unit for the anatomical reason that the axonal range of all Golgi cells in a 3-field row includes all or most of the middle field. Put another way, each field defines a functional population of Golgi cells – an ensemble – being those cells with convergent input to that field.

Ensembles perform a computational conversion of mossy fibre rates received by a three-field row into inhibition of granule cells. Say a population of values (with any frequency distribution) is randomly sampled. 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 untidy result. However, if

the new distribution is then itself randomly sampled, and the process is repeated, the results get tidier and narrower.

The ensemble computation is (chiefly) a statistical result of random sampling, in two steps. At the first step, Golgi cell basal dendrites receive contact from a random sample of mossy fibres that innervate their location (strictly speaking, they randomly sample terminals); and in the second step, middle-field glomeruli each receive innervation by a random sample of Golgi cells. The effect of the computation is to synchronise inhibition of granule cells in the middle field. The strength of inhibition is linearly related to the mean of mossy fibre firing rates received as input to the entire rank. Ensembles conflate sagittally anatomically and functionally. Ensembles in a sagittal strip all return the same output (because they all perform the same computation on input data sampled from the same distribution). Effectively: the Golgi cell population of a whole rank functions as a single unit.

4.2 The Golgi cell computation

We envisage that in motor circuits during behaviour signalling is analogue, avoiding the noise caused by discrete signals. So, for example, Golgi cell dendritic signalling is a sustained state under modulation by fluctuating mossy fibre rates. Similarly, GABA spillover into the intraglomerular space, almost exclusively the agent of granule cell

306 inhibition, becomes elevated by a build-up during behaviour, so that this too is at a
307 sustained level, controlled by afferent Golgi cell rates. Thus, at each step, data are
308 coded at any moment in a snapshot of a state of change.

309 We tested the hypothesis *in silico* by simulating a rank. Derivation of model
310 parameters is described in: Methods, Derivation of ensemble dimensions and model
311 parameters. Mossy fibre rates sampled by Golgi cells in each field were obtained by
312 sampling from a distribution representing signals received as input to a rank, which
313 were randomly generated subject to constraints on the shape of the distribution. This
314 was modified to take into account mossy fibre terminal branching and clustering,
315 random variation of the number of terminals per cluster and clusters per cell, and the
316 fact that most clusters fall partly inside and partly outside a field (see Methods,
317 Derivation of mossy fibre rates).

318 This gave the number of mossy fibre signals, firing rates and the number of copies of
319 each rate received by each field, and therefore input to an ensemble territory,
320 populated by 30 Golgi cells, 10 per field. There are ~700 mossy fibre terminals per
321 field. As far as we know (and we propose), contact by mossy fibres on Golgi cells is at
322 random at field scale, i.e., signals received by a single basal dendrite are provided by a
323 random sample of local terminals. The number of physiological samplings is equal to
324 the number of basal dendrites per field; sample size is the number of mossy fibres
325 which make contact on each dendrite. The effect is sustained dendritic depolarisation

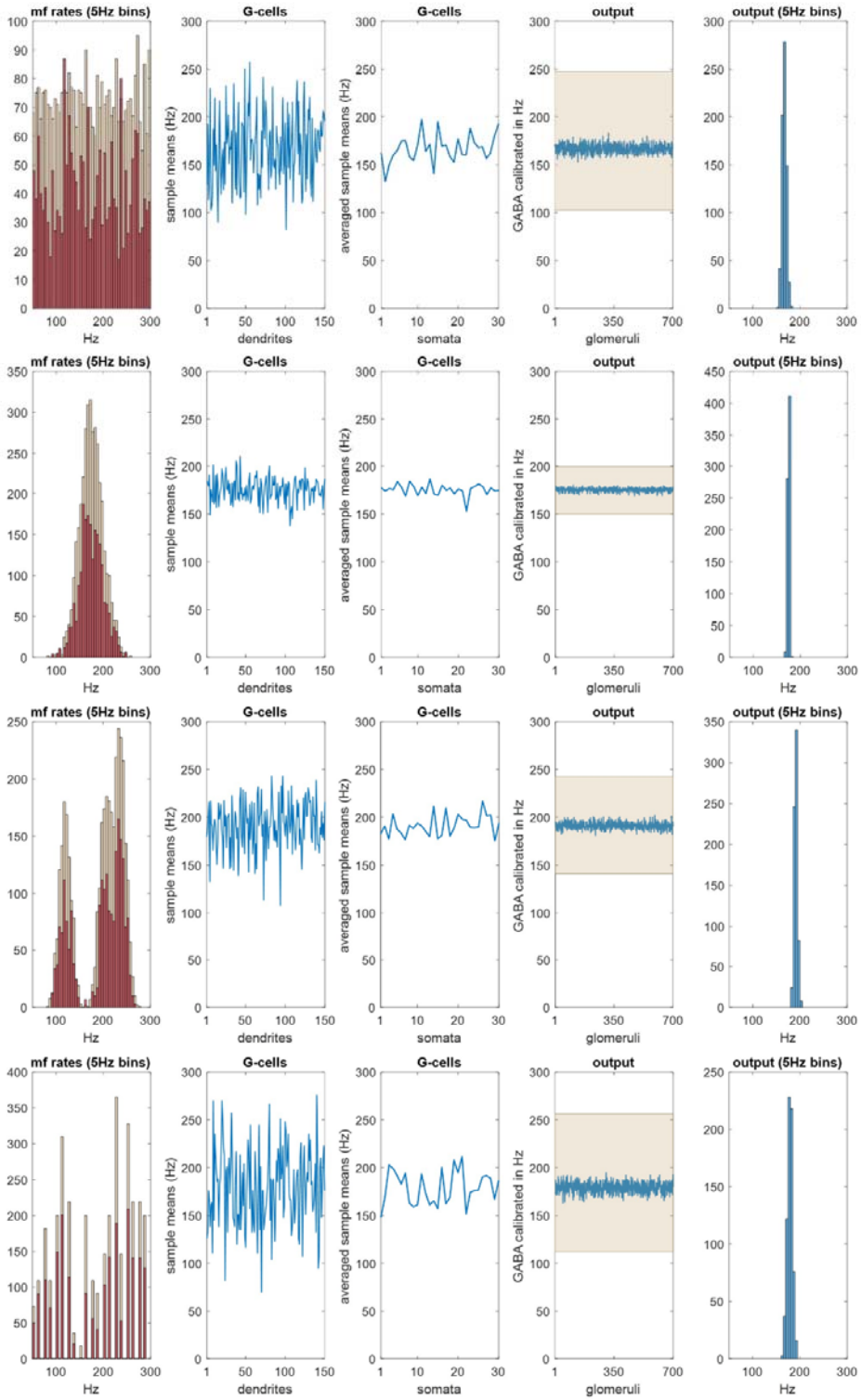
326 under modulation by a linear function of input rates (see Methods, Linear transmission
327 of mossy fibres to Golgi cells), which we take as the mean.

328 Taking the sample means generates a new set of values (Fig 2 column 2) with a smaller
329 range and centred on the mean, which is linearly related to the mean of the sampled
330 distribution. Basal dendritic charge transfer is passive and little filtered [11, 28].
331 Integration of dendritic states at the Golgi soma is represented by taking the mean of
332 dendritic charge, for each simulated cell. Golgi cell firing rates have a linear
333 relationship with the amplitude of depolarising current [29], suggesting (and we
334 assume) a linear conversion of somatic membrane potential into the Golgi cell firing
335 rate.

336 The output is a third data set, Golgi cell firing rates (Fig 2 column 3), on the same
337 normalised scale. Each glomerulus in the middle field receives innervation from a
338 random subset of Golgi cells afferent to the field, and accordingly a random sample of
339 Golgi cell firing rates. Golgi cells inhibit granule cells in the glomerulus. A granule cell
340 extends a single dendrite into a glomerulus, typically receiving synaptic contact from
341 two or three Golgi cells [30]. However, inhibition of granule cells by Golgi cells is
342 almost entirely by GABA spillover into the intraglomerular space [5], via extrasynaptic
343 receptors [31, 32] (see Methods, Linear transmission of Golgi cells to granule cells).
344 This effectively increases the Golgi cell to granule convergence ratio, since the number
345 of Golgi cells that are functionally afferent to a granule cell is not the number that

make direct contact, but the number afferent to a glomerulus. The membrane around a glomerulus restricts GABA diffusion so that signalling is by modulation of the intraglomerular concentration of GABA, which in turn reflects the rate it is released. The rate of release by a single cell is proportional to the firing rate of that cell. Intraglomerular GABA concentration is under continuous modulation at a sustained level which we take as a linear function of the mean of afferent Golgi cell rates (the sum also works). Sample size is provided by the Golgi cell to glomerulus convergence ratio. The number of Golgi cells afferent to a glomerulus is unreported, as far as we know. We take it as a random number in the range 8–12, generated for each glomerulus. We tested other ratios to confirm that the ratio does not materially affect the ensemble computation (Supplementary Materials). The number of samples is the number of glomeruli per field (~700, the same as terminals)

This generates a final population of values (Fig 2 columns 4 and 5), the output of the ensemble conversion. There are 700 outputs, representing glomeruli in the middle field. All granule cells that extend a dendrite into the same glomerulus receive equal inhibition (of that dendrite).


mf rates (5Hz bins)
G-cells
G-cells
output
output (5Hz bins)
Hz
dendrites
somata
glomeruli
Hz
sample means (Hz)
averaged sample means (Hz)
GABA calibrated in Hz

363

364 **Fig 2. Golgi cell ensemble conversion. Column 1** We randomly generated (3,645)
365 values representing mossy fibre rates with a uniform, normal, twin-peaked or
366 discontinuous distribution, simulating input to a row of 100 fields. Rate data were
367 randomly sampled to derive input rates to each field, and copies added to reflect the
368 fact that each afferent cell contributes a randomly variable number of terminals. Red
369 data represent copies-included signals received by a row of three fields. See Methods
370 for derivation of rates. **Column 2** The first-column red data were randomly sampled
371 150 times, sample size 4, representing mossy fibre input to a 3-field population of 30
372 Golgi cells, each with 5 basal dendrites (the mean) which each receive contact from 4
373 mossy fibres. Data are the sample means, representing dendritic depolarisation.
374 **Column 3** We next randomly sampled column 2 data *without* replacement 30 times,
375 sample size 5, and took the sample means, representing somatic integration of
376 dendritic signals. **Column 4** Column 3 data were randomly sampled 700 times, sample
377 size 8–12, reflecting convergent input from 3 fields of Golgi cells onto 700 glomeruli in
378 the middle field (which receives the output of a three-field ensemble), and took the
379 sample means, representing GABA spillover that is a linear function of the mean of
380 afferent rates. Inhibition of granule cells is almost exclusively mediated by spillover.
381 **Column 5** The column 4 data are presented as a histogram. **All rows** Column 1 mossy
382 fibre data are converted to narrowly-focussed inhibition of granule cells. The mean of
383 the column 4/5 data is linearly related to the mean of the mossy fibre distribution. This
384 result is independent of the shape of the original distribution.

385

386

387

4.3 Results

4.3.1 An ensemble

There is a progressive narrowing of the data range at each step. Inputs to an ensemble (mossy fibre signals received by a row of three fields) are converted to narrowly-grouped and normally-distributed outputs (representing GABA concentration in middle-field glomeruli). The narrow output range indicates that inhibition is locally synchronised, i.e., granule cells in the middle field receive co-modulated inhibition. Inhibition has a linear relationship with mossy fibre rates, acting through the ensemble conversion. These results are independent of the frequency distribution of mossy fibre rates – input to the system may be in any permutation of firing rates, with the same result.

4.3.2 A rank

We next used the ensemble conversion to derive inhibition of granule cells in a sagittal row of 100 fields, a rank. The results are shown in Figure 3. As ensembles all perform the same computation and effectively randomly sample (with replacement) the same data (the distribution of mossy fibre input rates to the whole rank), they all return the same output, or nearly the same output. There are 700 outputs per field, representing glomeruli, as before. The mean of outputs, taken for each field, has a narrow range centred on the mean of mossy fibre input rates to the parent rank. Focus varies with

the bandwidth of the sampled distribution (mossy fibre rates). A normal mossy fibre distribution with a range of over 100 Hz is converted to a blue range of ~3 Hz.

In Figure 3 rows A and B, blue data in the bottom right panel represent the mean, for each field, of glomerulus data. The range of the blue data is centred on the grey dashed line – the mean of mossy fibre input rates received by the whole rank – and varies independently of the red data. That is, field-to-field variation of output appears to be inherent in the mechanism – i.e., caused by inherent uncertainty of the outcome of random sampling. However, it is not compounded – random variation of the red data doesn't flow through to amplify random variation of the blue data (in fact the blue data have a slightly smaller range).

Conclusion: Golgi cells in a rank act as a single functional unit. Ensembles conflate in the sagittal direction functionally as they do anatomically. As a result, the whole granule cell population *of a rank* receive synchronised inhibition which is linearly related to, and varies with, the mean of mossy fibre rates.

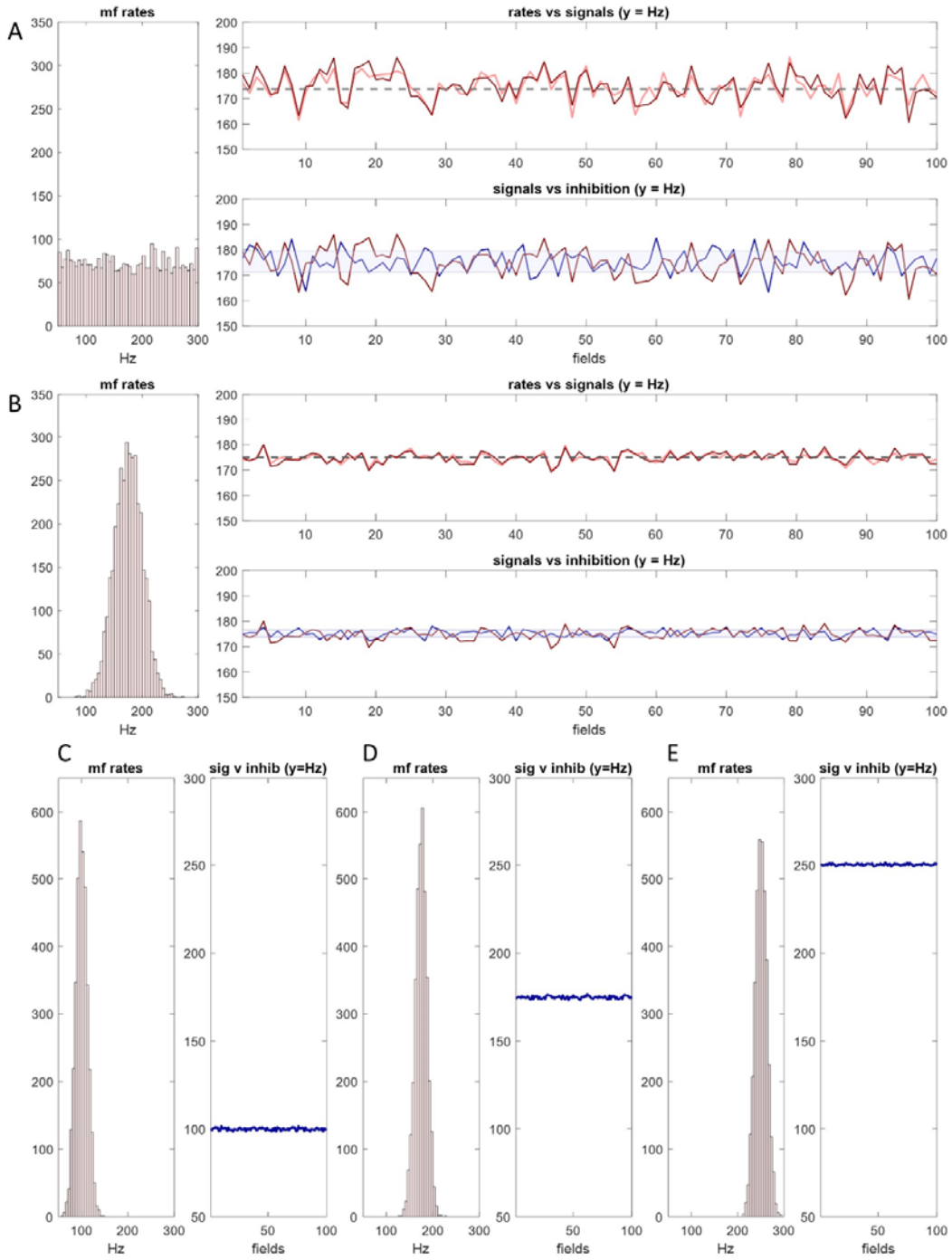

A
mf rates
Hz
rates vs signals (y = Hz)
signals vs inhibition (y = Hz)
fields
B
mf rates
Hz
rates vs signals (y = Hz)
signals vs inhibition (y = Hz)
fields
C
mf rates
Hz
sig v inhib (y=Hz)
fields
D
mf rates
Hz
sig v inhib (y=Hz)
fields
E
mf rates
Hz
sig v inhib (y=Hz)
fields

Fig 3. The Golgi cell population of a rank functions as a single unit. A Simulation of Golgi cells in a sagittal row of 100 fields, taking the mossy fibre distribution on the left as input (see Fig 2 and Methods for derivation of mossy fibre rates). **Top right**: pink: the mean of mossy fibre input rates to each field; red: the mean with duplicate signals added; dashed grey line: the mean of input rates to the whole rank. **Bottom right**: red: same as top right; blue: mean, for each field, of glomerulus data derived in the same way as Figure 2 columns 4 and 5 (the 'inhibitory field mean'); pale blue: SD of glomerulus data. **B** Same as (A) but sampling a different mossy fibre distribution. **A and B** Red data closely track the pink data in top right panels, indicating that duplicate signals (because the mossy fibre axon ends in a cluster of terminals) do not skew the mean of firing rates received by a field. Inhibitory field means are tightly grouped: inhibition of granule cells is synchronised in a rank. **C** The same simulation. Left: the sampled distribution (mossy fibre input rates to a rank); right: the inhibitory field means (we use the means to compare fields). The y axis of the right panel has the same range as the x axis of the left panel. Panels are tall to improve visual comparison to other rows. **D** and **E** The same, moving the sampled distribution to the right. Inhibition of granule cells is precisely and linearly coupled to the mean of mossy fibre input rates to the whole rank.

5. Granule cells

5.1 The fraction of active granule cells is reliably regulated by feed-forward inhibition at low density

5.1.1 Hypothesis

Anatomy simulates independent random sampling by a single granule cell of the population of mossy fibre input rates to the whole rank (section 3). There is a competition for influence between mossy fibre excitation and Golgi cell inhibition of granule cells. The competition takes place in the glomerulus. During behaviour, this takes place ceaselessly. The strength of inhibition is synchronised between glomeruli that populate a rank (section 4), so that all mossy fibre signals compete, at any moment, against the same inhibitory threshold. The outcome in each glomerulus is independent of the outcome at other terminals, including other terminals that innervate the same granule cell. To fire, a granule cell must receive a minimum number of mossy fibre inputs that prevail in the glomerular competition. The probability that enough prevail depends, at any moment, on the strength of inhibition. Since inhibition is synchronised, the probability is the same for all granule cells at any given instant in time. The number that fire is very near the expected proportion (the number predicted by probability) because granule cells are numerous. The probability

465 is also the same at all times because inhibition has a linear relationship with mossy
466 fibre firing rates (section 4). This is always conserved (by the Golgi cell ensemble
467 conversion) regardless of the mossy fibre distribution (section 4). So, the probability is
468 invariant, and therefore so is the proportion of active granule cells.

469 Note: the idea of an all-or-nothing result of the intraglomerular 'balance' of excitation
470 and inhibition has appeared previously [33].

471 5.1.2 Anatomy

472 Granule cells receive input to each dendrite from a single mossy fibre. Granule cells
473 have 3–5 dendrites, averaging 4 (the simulation assumes 4, but the result is the same
474 with 3–5). Each dendrite of a single granule cell extends into a different glomerulus.
475 Estimates vary of the number of granule cells that receive contact from a single mossy
476 fibre terminal. 50 is mid-range [34, 35], but estimates vary significantly [2]. 50 is
477 consistent with our estimates of the number of terminals and granule cells in a field.
478 The granule cell sampling unit is the whole cell since it receives only a single input rate
479 to each dendrite. Sample size is therefore 4, and the number of samples per field is the
480 number of granule cells, which we estimate at 8,750 (see Methods, Derivation of
481 mossy fibre rates).

482 Excitation and inhibition compete for influence in a glomerulus. It is unclear if there is
483 always an emphatic winner, but a swing in the dominant direction may be amplified by

484 positive feedback, acting through presynaptic receptors. Activation of presynaptic
485 $GABA_B$ receptors on mossy fibres inhibits glutamate release [36], while activation of
486 presynaptic mGluR2 receptors on Golgi cells inhibits GABA release [37]. We postulate
487 that only mossy fibre signals which prevail in the glomerular competition contribute
488 significantly to granule cell depolarisation.

489 5.1.3 Feed-forward inhibition regulates the probability that granule cells 490 fire

491 The probability that a given granule cell receives three competitive signals and
492 therefore fires is a function of the percentage of mossy fibre signals that are
493 competitive. The expected number of granule cells that fire in a field is given by

494

$$E_{fld} = T \sum_{x=j}^{n} \left[\frac{n!}{(n-x)!x!} [P(e > i)]^x [1 - P(e > i)]^{n-x} \right]$$

495

496 where T is the total number of granule cells per field, n is the number of dendrites per
497 granule cell, and j is the minimum number of dendrites that must receive dominant
498 excitation, $e > i$. The proportion of granule cells that fire is near the expected number
499 by the law of large numbers. The percentage of mossy fibre signals that are
500 competitive is always the same because inhibition scales with the mean of mossy fibre

501 rates (because Golgi cells translate mossy fibre firing rates into proportional
502 inhibition). As the percentage is unvarying, so is the probability.

503 5.1.4 Does physiology have that outcome?

504 To test that, we incorporated the Golgi cell ensemble conversion into a model that
505 recodes mossy fibre signals into granule cell signals. Probability is an abstraction with
506 no direct physiological counterpart. The model does not include a binomial function
507 but derives output from input by comparing excitatory to inhibitory input to each
508 glomerulus, the way we claim it is derived physiologically. The equation in the previous
509 section is simply shorthand – biology exploits chance and large numbers to generate
510 the same result as if inputs were passed through the equation.

511 The model compares mossy fibre signal frequency with inhibition at each of 4
512 dendrites of each of 8,750 granule cells per field, in each of 100 fields. Mossy fibre
513 input rates to each field are derived per section 4 and Methods, so taking into account,
514 *inter alia*, that a field receives a randomly variable number of copies of most signals,
515 and most terminal clusters lie only partly inside a field. Inhibition – outputs of the
516 ensemble conversion, 700 per field – is provided by the Golgi cell model. Excitation and
517 inhibition received by each glomerulus are randomly paired. Pairing is used to identify
518 the subset of granule cells that fire. We assume that, to fire, a granule cell must
519 receive at least 3 mossy fibre signals which exceed the inhibitory threshold. In
520 experimental conditions, using stimulation that excites very high mossy fibre rates,

521 granule cells can be induced to fire in response to only one input [1]. However, other
522 accounts suggest one is too few [5, 38, 39], and three are needed [40, 41]. Output of
523 the model is the number of granule cells that fire in each field, the mean of mossy fibre
524 rates they each receive (the sample means), and the mean of the sample means in
525 each field ('field means'). The results are shown in Figure 4.

526 ### 5.1.5 Results

527 Feed-forward inhibition – mossy fibre → Golgi cell → granule cell – reliably and closely
528 regulates the fraction of granule cells that are active at low density. Field-to-field
529 random variability of the number is inversely proportional to the fraction that are
530 active, suggesting that, if primary regulation is provided by feed-forward inhibition, the
531 fraction is low. We would nonetheless expect to see greater precision, suggesting the
532 physiological presence of a supplementary mechanism(s) that we have not
533 represented in the model.

534 A local feedback mechanism may be provided by granule cell ascending axon contact
535 on Golgi cells. The ascending portion of the granule cell axon makes synaptic contact
536 on local Golgi cells, on basal dendrites [11]. A Golgi cell receives ~400 functional
537 synaptic contacts from this source, around 25% of the granule cells that a Golgi cell
538 receives contact from in total. During sustained mossy fibre discharge, synapses from
539 local granule cells contribute equivalent charge to Golgi cells as mossy fibre synapses

540 and ‘display similarly fast kinetics’ [42]. Short-latency inhibitory feedback may add
541 homeostatic fine adjustment to feed-forward regulation at local level, we propose.

542 It is a long-standing idea [43] that granule cell activity may be regulated by feedback
543 via parallel fibres, which contact Golgi cell apical dendrites. We have previously
544 proposed that gap junctions between apical dendrites may provide the means for
545 Golgi cells to detect the density of parallel fibre activity [44]. Feedback by this pathway
546 is not sensitive to local fluctuation of granule cell active numbers, and may be the
547 weakest regulator.

548 The results are the same with any (tested) distribution of mossy fibre rates and also
549 whether mossy fibres are all active or some are active and others are not. Active cells
550 can be in any random permutation.

551

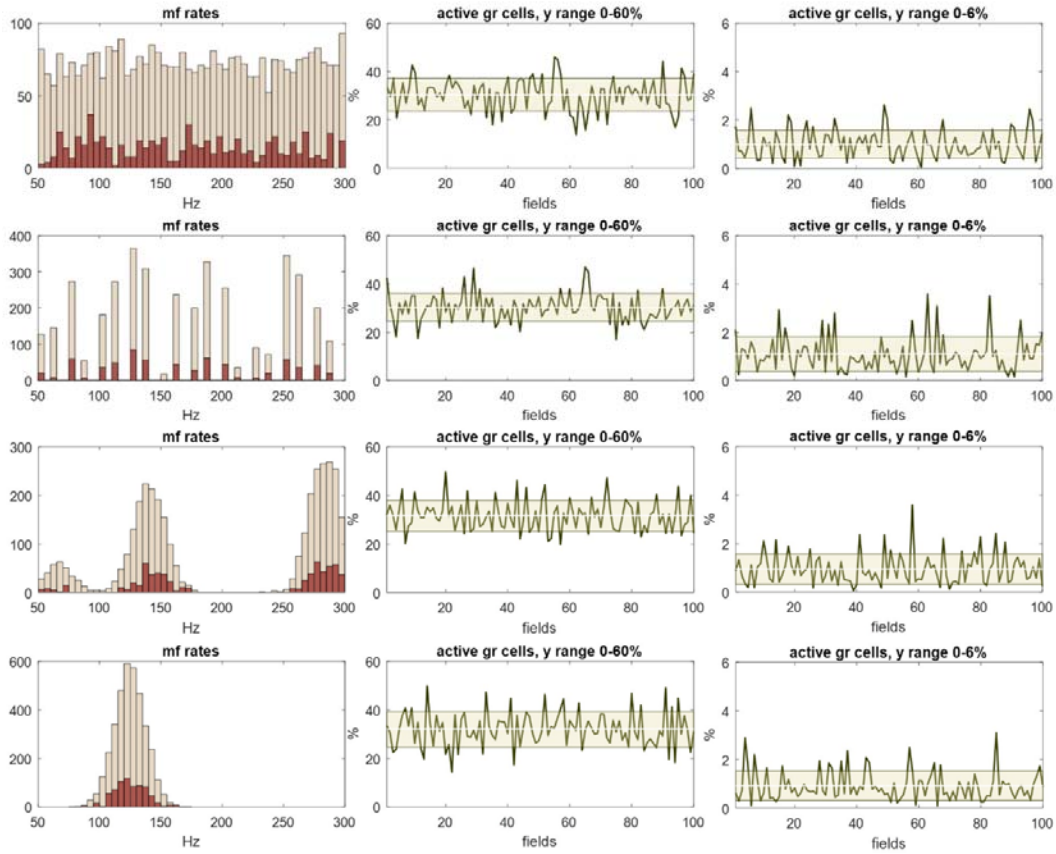


Fig 4. Regulation of granule cell activity by the feed-forward pathway (mossy fibres→Golgi cells→granule cells) We incorporated the Golgi cell ensemble conversion into a simulation of the conversion of mossy fibre signals into granule cell signals (recoding). One of the outputs of recoding is regulation of the percentage of active granule cells. We tested if the recoding computation causes the percentage to be stable, i.e., in the same narrow range in all fields regardless of mossy fibre rates. **Column 1** Mossy fibre firing rates were randomly generated with a uniform, discontinuous, three-peak or normal distribution, to represent input to a row of 100 fields. Rates, and rates received by each field, were derived in the same way as earlier Figures. Red data are rates received by a single randomly-selected field, including duplicate signals. y axes are scaled to the data – all distributions contain the same

number of signals. **Column 2** The percentage of granule cells that fire in each field. The Golgi cell computation includes no adjustment to sparsen granule cell activity. The white line and shaded area are mean and SD. **Column 3** The same, with an adjustment (a co-efficient) to sparsen granule cell activity. y axes have 1/10 the range of column 2. The percentage is confined to a narrow range. This result is independent of the sampled (i.e., mossy fibre) distribution, suggesting that regulation is independent of fluctuation of mossy fibre rates. Nonetheless, field-to-field variation relative to the mean is significant. Homeostatic local fine tuning may be provided physiologically by feedback via granule cell ascending axon contact on Golgi cells, not represented in the simulation.

5.2 Granule cells encode mossy fibre rate information collectively in rank-sized strips

5.2.1 Hypothesis

Anatomy simulates independent random sampling by a single granule cell of the population of mossy fibre input rates to the whole rank (section 3). The granule cell firing rate is a linear function of depolarising somatic charge, in turn a linear function of the mean (or sum) of excitatory input rates that prevail against inhibition. We postulate that the linear relationship is undiluted by inhibition because dominant signals displace a significant effect of inhibition. Derived in this way, granule cell rates are collectively well predicted by mossy fibre rates, by the central limit theorem. This

586 provides the substrate for granule cells to collectively encode mossy fibre rate
587 information. Granule cell code groups are functionally defined by the topography of
588 mossy fibre input to the cerebellum, as Golgi cell networks are. The output of a rank is
589 the smallest unit of granule cell encoded information. The permutation of active mossy
590 fibres, and of granule cells, and how firing rates are distributed among active cells, all
591 code nothing.

592 5.2.2 Do granule cell firing rates – the output of a rank – have the 593 hypothesised features?

594 We ran the same simulation as section 5.1 to test whether, and how, mossy fibre rates
595 are related to granule cell rates, and if fields in a rank all have the same output. We
596 take granule cell firing rates to be a linear function of the mean of dominant input
597 rates to each active cell, which we take as the mean itself. The results are shown in
598 Figures 5 and 6. We present evidence relating to mossy fibre → granule cell linear
599 transmission in the Methods section (Methods, Linear transmission of mossy fibres to
600 granule cells).

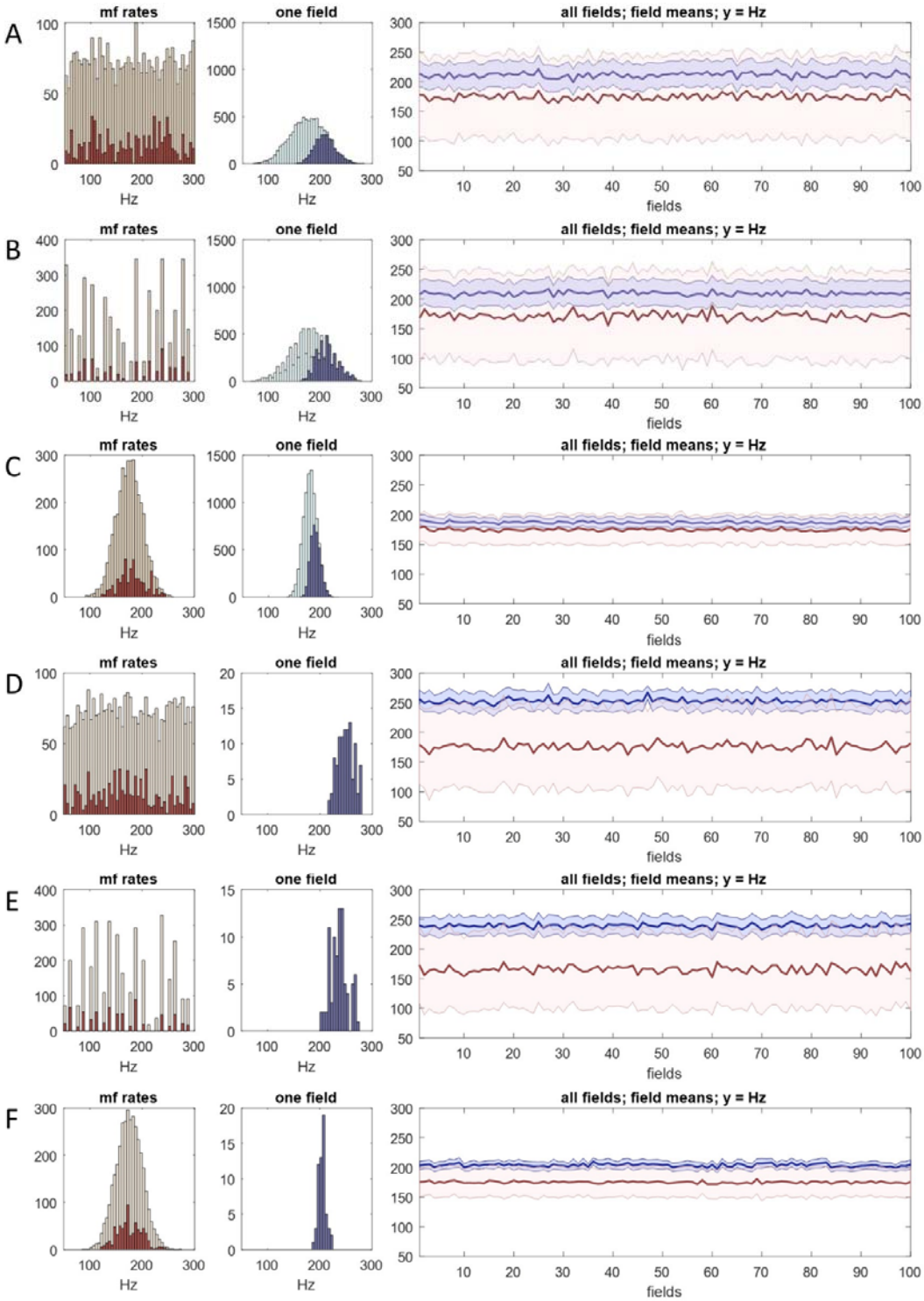
601 The narrowing and predicted shape of the granule cell frequency distribution is evident
602 in all conditions. The field means (representing the mean of granule cell rates in each
603 field) approach a straight line, the test that a rank operates as a single functional unit.
604 When mossy fibre rates are normally distributed (SD 12.5 Hz), the granule cell

605 distribution is very narrow (and normal if rates are counted in narrower bins) and the
606 field means are very close to a straight line. The field means follow the mean of mossy
607 fibre rates at a precisely fixed distance. In all tested conditions, granule cell rates have
608 a narrowed range, which is the same in all fields, clustered around the mean, and
609 attracted to a normal distribution (more evident with narrowly and/or normally
610 distributed mossy fibre rates). This holds even though a single granule cell receives a
611 tiny fraction of mossy fibre rates, and each field receives only a small fraction of mossy
612 fibre signals received by a rank, and regardless of the shape of the mossy fibre
613 distribution.

614 We conclude that, plausibly, granule cell firing conserves mossy fibre rate information,
615 coded in the output of a whole rank. This is therefore the minimum resolution of
616 granule cell information. However, it is not the minimum unit of code. The same
617 information is coded in any random sample of granule cell signals, over a minimum
618 number. It is immaterial (i.e. codes nothing) which particular cells are active or which
619 cells fire at what rates, or even what rates those are, provided any random sample of
620 rates is as if drawn from the same frequency distribution, with replacement. The
621 output of a rank can be likened by analogy to a pixel: it is only divisible into smaller
622 pieces of the same information.

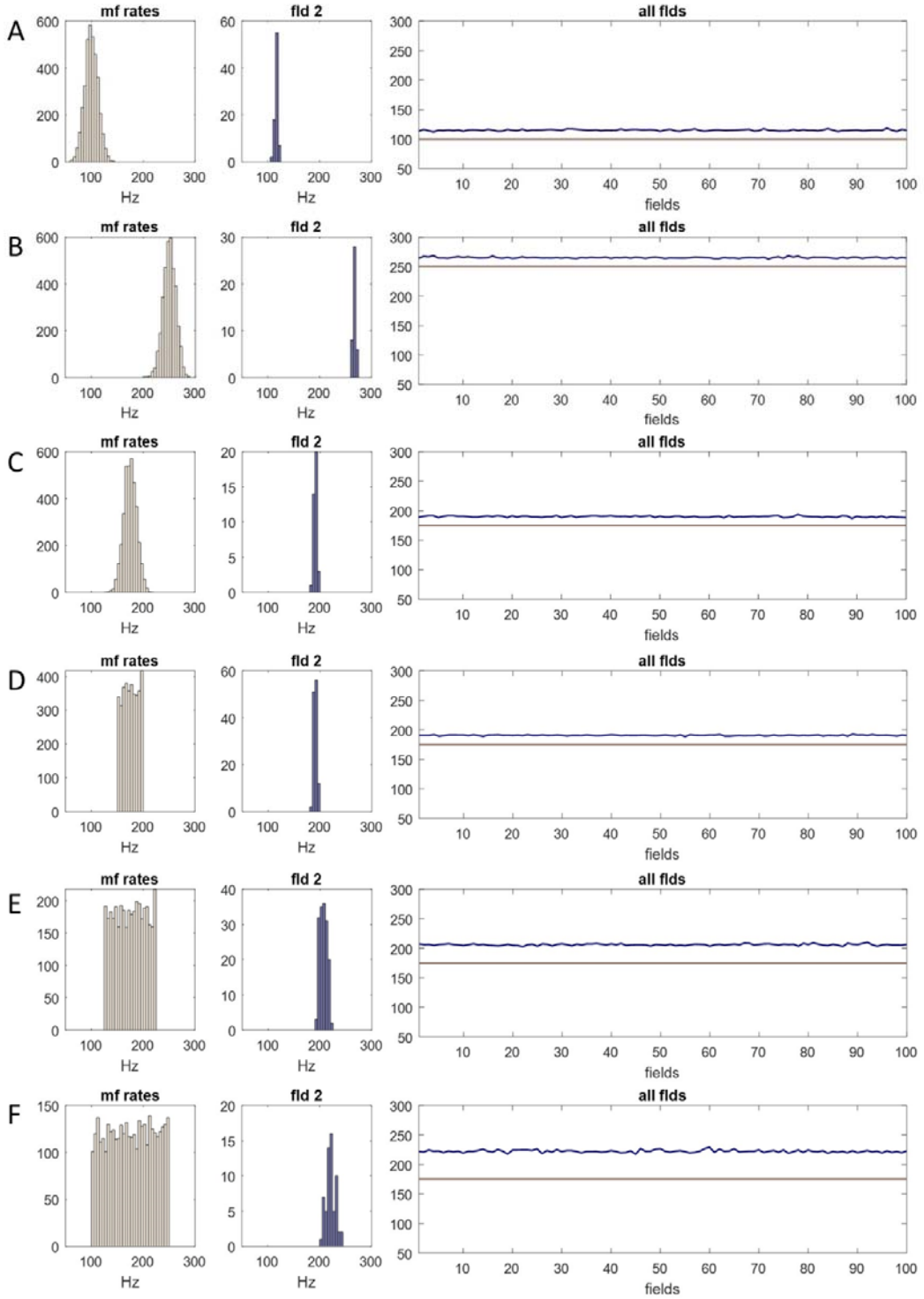
623

624


A
B
C
D
E
F
mf rates
one field
all fields; field means; y = Hz
Hz
fields

625

Fig 5. A rank functions as a single unit. The same simulation as Figure 4 but data are granule cell firing rates or derived from granule cell firing rates. **Column 1** Mossy fibre rates were generated in the same way as earlier figures (the number of signals is the same in all rows). **Column 2** Pale green: sample means – the mean of input rates to each granule cell in a single field selected at random; blue: sample means of the subset of granule cells which meet the conditions to fire. In rows A–C, ~30% fire. In rows D–F, ~1% fire, y axes are rescaled and green data are omitted. **Column 3** Red: mean of mossy fibre input rates to each field in a rank, including duplicate signals; pink: SD of mossy fibre input rates to each field; blue: field means (i.e., the mean of granule cell firing rates in each field); pale blue: SD of granule cell firing rates in each field. Regardless of the shape and large range of the mossy fibre distribution, the field means always converge towards a straight line which is a fixed distance from the overall mean of mossy fibre rates. The amount and direction of field-to-field random fluctuation of the blue data is independent of local fluctuation of the red data, and in a smaller range. Uncertainty inherent in random sampling – first by single fields of input rates to a rank, and then by Golgi cells and granule cells of input rates to a field – is not compounded.



644

Fig 6. Granule cells collectively code mossy fibre rate information. Column 1 Mossy fibre rates were generated in the same way as earlier Figures, with either a normal (SD 12.5) or uniform distribution representing input to a row of 100 fields, with copies added – the sampled distribution. **Column 2** Histogram of granule cell firing rates in a randomly selected field, obtained in the same way as in the bottom 3 rows of Fig 5. **Column 3** Grey line: the mean of the sampled distribution; blue line: field means. **A–C** The fields means are precisely and linearly related to the mean of mossy fibre rates. Blue data approach a straight line which is parallel to the grey line at a mean distance of 15 Hz (SD 1.21 Hz). **D–F** If the shape and mean of a uniform sampled distribution are unchanging, the distance between the blue and grey lines is linearly related to the range.

6. Discussion

6.1 Conclusions

Network models are often a simplified representation of physiology to provide a substrate for sophisticated computations. We find the reverse: physiology is highly adapted to provide a substrate for straightforward computations. In summary we find:

1. There is a computational effect, unaided and unlearned, of the combination of cell morphologies, network architecture and linear units, for which the granular layer is highly adapted. There is no algorithm.

2. Physiology is the physical form of a network of linear units. A unit may be a cell, part of a cell or some other structure. Units randomly sample afferent activity. Code groups communicate by a facsimile of independent random sampling. Anatomy replicates the same result as if single neurons independently randomly sample the entire population of firing rates in the afferent cell layer, notwithstanding that single cells receive fixed contact from a tiny fraction of afferent cells.

3. Data processing exploits statistical effects of random sampling, such that i) inhibition of granule cells in a rank is synchronised and tracks mossy fibre rates; and ii) granule cell rates are in a condensed range, reliably (because statistically) well-predicted at collective level, and again linearly related to mossy fibre rates.

4. High-dimensional mossy fibre information recodes as a very low-resolution granule cell code. Large granule cell groups function as a single unit, coding information in their collective activity.

5. At any instant, a whole group codes the same information. Accordingly, information has spatial dimensions, which can be likened to 3-dimensional pixels. Cell groups take their dimensions from a combination of topography and cell type morphologies, rather than having physical boundaries.

6. Pixels span the cerebellar network in the sagittal direction, from side to side, mirroring microzonal organisation of the molecular layer. Accordingly, granule cell encoded information has a striped pattern viewed from the cerebellar surface.

7. While firing of a whole code group is the minimum unit of information, it is not the minimum unit of code. Inside code group boundaries the same information is coded in any random sample of granule cell firing rates. We might say for convenience of reference the code is homogenised.

8. The shape of data and homogenised code are key to function, having results discussed below.

6.2 Control of Purkinje cells

There is a huge sacrifice of data resolution on entry to the cerebellum, counter to the general expectation that data processing is sophisticated. This is functional. Coded in this form, granule cell data can drive coherent firing of large cell groups downstream. Microzones lie at right angles to parallel fibres and parallel to code groups. The statistical form of the granule cell code means the same information is received by all locations of sagittal plane, so that microzone-grouped Purkinje cells can receive the same information without receiving the same signals. An estimated 350,000 parallel fibres pass through a Purkinje cell dendritic territory [45-48]. If even only a small fraction of parallel fibres is active (0.5%), and allowing for the fact that only one in two make contact, and that a majority of synapses are severely depressed or silent [49, 50], a single Purkinje cell still receives around 200 parallel fibre signals at working synapses

704 at any time. A granule cell code in this form provides a means to modulate firing of
705 Purkinje cells in unison by a chain of rate codes: mossy fibre → granule cell → Purkinje
706 cell. There is evidence that Purkinje cells linearly code the strength of granule cell
707 inputs [6-8] and linearly code behavioural metrics [3, 51-54].

708

709 6.3 Replicability

710 A statistical code may be a physiological strategy to mitigate variable performance of
711 single neurons and synapses, a long-known impediment to precise and replicable
712 biological information capture. The traditional cerebellar learning model [43, 55, 56]
713 relies on (and assumes) replicability of patterns of active and inactive parallel fibres,
714 because pattern replication is necessary for signals to match up with the correct
715 training-modified synaptic weights. Learned changes of synaptic weights are also a
716 core feature of artificial neural network learning models.

717 However, this may not be safe. First, it is not clear that input to the cerebellum is ever
718 replicated. Around 4,000 mossy fibres innervate the region that supplies parallel fibre
719 input to a single Purkinje cell. It may not be a safe assumption that conditions
720 upstream are routinely repeated with precision sufficient to consistently (or perhaps
721 ever) generate a repeat of the same inputs. Second, single neurons and synapses are
722 not reliably consistent performers. Neurons even of the same type are by no means

replicas, and synaptic transmission involves random variables. Third, the granular layer is *designed* to amplify the effect on the pattern of active parallel fibres if the pattern of active mossy fibres should vary, even only by a modest number of cells. Indeed, even that is not necessary – the same pattern of active mossy fibres but firing in a different permutation of rates also generates a different pattern of active granule cells, and therefore parallel fibres. Small differences of input (the pattern of active mossy fibres and/or permutation of rates they fire at) are converted by decorrelation to wholesale changes to output (the pattern of active parallel fibres). Possibly, and we submit, the cerebellum actively eschews reliance on pattern replication in favour of group code which gives a more (because statistically) reliable outcome.

6.4 Working 'blind'

A perennial issue for network theory is that biological networks do not seem to have much information about the signals they receive. Input to the cerebellum is eclectic and in multiple modalities. In some cases, a single granule cell can receive multi-modal input. In the traditional cerebellar model, supervision solves this problem. It is unnecessary for the cerebellum itself to discriminate because it is an assumption of the model that externally-sourced instruction signals train the correct response.

741 We propose instead the solution that resolution of high-dimensional input data is
742 sacrificed. The cerebellum disregards source and modality data below a physical scale
743 threshold. It is therefore unnecessary to discriminate below that scale. Mapping
744 studies show that mossy fibre signals (evoked by cutaneous stimulation) share the
745 topographical organisation of climbing fibres [57-59]. At sub-topographical
746 dimensions, signal source and type are simply discounted. Therefore, networks do not
747 need to be adapted to the source or type of input they receive – anatomically,
748 networks are interchangeable, so network design can be modular. This feature may
749 have been necessary to allow basic cerebellar wiring to survive almost unchanged
750 through dramatic upheavals of the vertebrate phenotype [60, 61].

751

752 **METHODS**

753 **DERIVATION OF MOSSY FIBRE RATES**

754 **The number of mossy fibres that innervate a rank**

755 As far as is known, fields receive a random sample mossy fibre input rates to a rank. To
756 simulate that, we need the size of the sampled population – the number of mossy fibre signals
757 received by a rank – and sample size, the number received by a field.

758 Despite the repeating 'crystalline' architecture of the cerebellar cortex, at local level there are
759 anatomical random variables. Variables are: i) the number of mossy fibres that innervate a
760 field, ii) the number of terminals they each end in, and iii) the number of those that are inside
761 field limits.

762 The number of mossy fibres afferent to a field has been estimated at 100 [21]. However, that
763 number is a field-sized fraction of the general population. The actual number is larger because
764 mossy fibres end in a dispersed group (cluster) of terminals, and most groups straddle field
765 limits.

766 We work from an estimate of the total number of terminals per field, which is, in turn, derived
767 from convergent estimates of the number of granule cells per field. (Estimate 1) There are an
768 estimated 1.92×10^6 granule cells per µl [45]. The number of fields which fit in that volume is
769 $(1{,}000/150)^2 \times (1{,}000/200) = \sim 222$, assuming a field depth of 150 µm [21]. The number of
770 granule cells in a field is, therefore 1,920,000/222 = 8,649. (Estimate 2) An estimated 350,000
771 parallel fibres pass through a single Purkinje cell's territory [45]. Parallel fibres run about 3 mm
772 in each direction from the granule cell soma [62, 63]. The Purkinje cell dendritic arbour fills the
773 molecular layer vertically and averages ~200 µm wide. Therefore, 6,000/150 = 40 fields fit in a
774 row that provides parallel fibres that pass through a Purkinje cell dendritic territory, giving
775 350,000/40 = 8,750 granule cells per field. Estimates vary of the number of granule cells which
776 extend a dendrite into a single glomerulus. 50 is mid-range. (8,750/50) x 4 = 700 glomeruli per
777 field.

Granule cells have 3–5 dendrites, averaging 4, and receive input to each of them from a single mossy fibre (so an average granule cell receives 4 inputs). Estimates of divergence of a single mossy fibre terminal onto granule cells vary [2]. It is unclear how much this reflects biological diversity versus differences in opinion and methods. Here we use a mid-range estimate of 1:50 [34, 35]. Assuming a mossy fibre terminal receives a single granule cell dendrite from each of d granule cells, and given an average of 4 dendrites per granule cell and z granule cells per field, the total number of mossy fibre terminals in a field is $N = (z \times 4)/d = 700$ given $z = 8{,}750$ and $d = 50$. A convergent estimate is provided by multiplying the per field number of mossy fibres by the average number of terminals per cluster.

Cluster size varies. We take the range as 4–12 terminals per cluster [21]. We assume the number varies at random and cluster sizes occur with equal probability. A maximum of 7 fits in a field (because a field shares the dimensions of the area that encloses a mean cluster). Fields are not anatomically bounded, so many clusters straddle field limits. To reflect that, we assume there is a 1/7 probability that a 7-terminal cluster (or larger) has 7 terminals inside a field and also the same probability of any other number. A 6-terminal cluster has a 2/7 chance that all 6 terminals are inside and a 1/7 chance of any other number, and so on.

Table 1 shows the derivation of the expected (i.e., predicted by probability) ratios of terminals contributed to a field by each cluster size (R1), not counting terminals that lie outside when a cluster straddles field boundaries; and the ratios of mossy fibres contributing n = 1–7 terminals (R2), regardless of cluster size. R1 ratios are given by the sum of numbers in each row. So, for example, the ratio of cells that end in 8 terminals to cells that end in 5 is 28:25. R2 ratios are given by the number of appearances of each number in the grid – so the expected ratio of cells which contribute 7 terminals to the number which contribute 4 is 6:12. The ratios effectively give the distributed probability of possible states. For example, there is a probability of $10/(6 + 8 + 10 + 12 + (3 \times 9)) = 0.159$ that a mossy fibre contributes 5 terminals to a field (regardless of cluster size). Similarly, there is a probability of $((6 \times 28)/((6 \times 28) + 27 + 25 + 22))/7 = 0.099$ that a given cell contributes 7 terminals to a field.

807 **Table 1. Derivation of expected ratios**

size	number of terminals inside							R1
12	7	6	5	4	3	2	1	**28**
11	7	6	5	4	3	2	1	**28**
10	7	6	5	4	3	2	1	**28**
9	7	6	5	4	3	2	1	**28**
8	7	6	5	4	3	2	1	**28**
7	7	6	5	4	3	2	1	**28**
6	6	6	5	4	3	2	1	**27**
5	5	5	5	4	3	2	1	**25**
4	4	4	4	4	3	2	1	**22**
R2	6	8	10	12	9	9	9	

808

809 If R1 ratios are denoted by a series $x_4, x_5 \ldots x_{12}$, and a given element in the series is x_j, the
810 fraction of terminals provided by a given cluster size is $x_j/(x_4 + x_5 + \cdots + x_{12})$, and this
811 fraction is multiplied by N to give the number of terminals. The number of terminals is divided
812 by the expected mean number of terminals that fall inside field limits (for that cluster size) to
813 give the expected number of mossy fibres that provide them. The number of j-terminal mossy
814 fibres that innervate a field is therefore given by

$$N\left[{x_j}\Big/{\textstyle\sum_{i=1}^{n} x_i}\right]\Big/\left[{x_j}/{7}\right]$$

815 The results are shown in Table 2. The equal number of mossy fibres (mfs) in each column is
816 what we would expect, because we made it an assumption that all sizes occur with equal
817 probability, and mf values are expected values – i.e., in proportions predicted by their

818 probability. The mf total is the expected number of mossy fibres afferent to a field (which
819 contribute at least one terminal).

820 **Table 2. The expected number of mossy fibres which innervate a field**

cluster size	12	11	10	9	8	7	6	5	4	total
terminals	81	81	81	81	81	81	78.1	72.3	63.6	700
mfs	20.25	20.25	20.25	20.25	20.25	20.25	20.25	20.25	20.25	182.25

821 R2 ratios – the relative incidence of afferent mossy fibres that contribute 7 terminals to a field,
822 6 terminals, 5 and so on, down to 1 – and the expected number of mossy fibres that contribute
823 them, are shown in Table 3.

824 **Table 3. The number of mossy fibres which contribute *n* = 1–7 terminals to a field**

n	7	6	5	4	3	2	1	total
R2	6	8	10	12	9	9	9	63
R3	42	48	50	48	27	18	9	
mfs	17.34	23.15	28.92	34.72	26.04	26.04	26.04	182.25
%age	9.52	12.69	15.87	19.05	14.29	14.29	14.29	

825 R3 is *n* x R2, giving the ratio of terminals provided by mossy fibres with *n* terminals inside a
826 field. %age is given by (R2/63) x 100. this is the percentage of mossy fibres with *n* terminals
827 inside a field.

828 From here, the number of mossy fibres that innervate a rank is given by 182.25 x 100 divided
829 by the mean number of clusters per mossy fibre per rank. If a rank receives, on average, five
830 clusters per mossy fibre, ~3,645 mossy fibres innervate a rank. This is the estimate used in the
831 simulation unless stated otherwise.

832 *Assumptions.* Cluster sizes occur with equal probability. The probabilities of each outcome
833 when a cluster straddles field boundaries (how many terminals are inside and how many
834 outside) are as stated. A single mossy fibre ends in an average of 5 terminal clusters per rank.

835 **Mossy fibre input rates to a field**

Terminal clustering means a field receives multiple copies of most signals. The proportions of mossy fibres afferent to a simulated field that contribute each possible number of terminals (1–7) are generated in simulated steps. In the first step, the simulation randomly generates, for each field, the cluster size ratio. This means, the relative proportions of afferent mossy fibres (which contribute at least 1 terminal to a field) that end in 4 terminals, 5, 6 and so on up to 12, assuming they occur with equal probability, disregarding how many terminals fall inside and outside field boundaries. In step two, the ratio of cluster sizes is weighted to reflect the greater probability that large clusters contribute more terminals to a field. In step 3, ratios are converted to the fraction of terminals contributed by each cluster size. When multiplied by the total number of terminals, this gives the number of terminals contributed by each cluster size. In step four, the values obtained in step three are divided up between cells that contribute a single terminal, 2 terminals, 3 and so on, to reflect straddling of field limits. The proportions give the number of mossy fibres afferent to a simulated field that contribute a single terminal, two, three and so on up to seven, the maximum.

There is no field-level effect of mossy fibre terminal branching. Terminal clusters arising from the same mossy fibre are separated by a minimum distance, so that a field cannot contain terminals of more than one cluster per mossy fibre.

Terminal clustering means that fields contain a randomly variable number of copies of most signals. In the simulation, copies are added after sampling of mossy fibre rates to obtain rates received by a field. An alternative order would be to add copies to the sampled population and then sample that. We follow anatomy: the number of *rates* received by a field is equal to the number of afferent mossy fibres; the number of *signals* contains duplicate rates.

DERIVATION OF ENSEMBLE DIMENSIONS AND MODEL PARAMETERS

4 Purkinje cells fit in a field, based on field size, thickness of the Purkinje cell dendritic plate – 9 µm [64] – and the distance between them – ~40 µm [65]. ‘In man, the ratio [of Purkinje cells to Golgi cells] is 1:1.5, while in the monkey and cat it is almost 1:1.9 and in the rat 1:3.3’ [66].

863 Possibly, estimates are low because smaller Golgi cells at deeper level were not counted [65
864 p.121]. We take it that there are 10 Golgi cells per field, a ratio of 1:2.5 with 4 Purkinje cells.

865 Golgi cell morphology creates ensemble dimensions. The Golgi cell axon ramifies profusely,
866 giving rise to a dense plexus. The axonal field is sagittally elongated – that is, in the direction of
867 the long axis of a microzone (mean range 650 +/- 179 µm by 180 +/- 40 µm [in mice: 67]) – and
868 is the depth of the granular layer, which it fills vertically. Fine, beaded axon fibres enter
869 glomeruli, where they inhibit granule cells [68]. Because of the dimensions and orientation of
870 the Golgi cell plexus, all Golgi cells in a sagittal row of three fields inhibit substantially the
871 whole of the middle field. Each field defines a functional population of Golgi cells – an
872 ensemble – being those cells with convergent input to that field. A single Golgi cell contacts
873 thousands of granule cells [68, 69].

874 Contact by mossy fibres on a Golgi cell is on basal dendrites. 4–6 basal dendrites arise from the
875 Golgi cell soma. We assume 5, so a population of ensemble-grouped Golgi cells has a
876 combined total of 30 x 5 = 150 basal dendrites. The number of mossy fibres which contact a
877 Golgi cell is unreported. Dendrites radiate from the cell body. A single dendrite has a range of
878 50–100 µm [65]. The longest dimension of a field – which contains ~700 mossy fibre terminals
879 – is 200 µm. We estimate the mean number of mossy fibres that contact a single dendrite is 4,
880 so 20 per cell, or ~3% of the population of terminals (or 0.6% per dendrite).

881 Glomeruli individually receive innervation from a random sample of Golgi cells afferent to a
882 field. It is unknown how many Golgi cells innervate a glomerulus.

883

884 **LINEAR TRANSMISSION OF MOSSY FIBRES TO GOLGI CELLS**

885 It is part of our proposal that Golgi cell basal dendritic signals linearly reflect mossy fibre rates,
886 so that mossy fibre rate information is conserved proportionally in Golgi cell firing rates. Golgi
887 cells are large interneurons whose cell bodies and basal dendrites lie in the granular layer.
888 Mossy fibres directly contact Golgi cell basal dendrites. Golgi cell firing frequency increases
889 linearly with the amplitude of depolarising current [29 p.845]. 'Golgi cells can follow peripheral

890 signals in a continuous fashion, modulating their frequency with the intensity of the stimulus
891 [citing [70, 71]]' [72]. As a result, during movement, the 'output of Golgi cells varies smoothly
892 as a function of input.…They are capable of modulating their output to reflect their input with
893 little noise or signal distortion' [12]. 'Sensory-evoked Golgi-cell inhibition scales proportionally
894 with the level of mossy fiber excitatory synaptic input', such that inhibition reliably conveys
895 mossy fibre rate information [4]. On this evidence: mossy fibre rate information is conserved
896 proportionally in Golgi cell firing rates.

897 Golgi cells extend their basal dendrites into glomeruli [68, 73]. Contact on them by mossy
898 fibres is multi-synaptic [65], contributing to a reliable, rapid (submillisecond) response. The
899 fast rise time and weak distance dependence of the amplitude and timing of Golgi cell EPSCs
900 evoked by mossy fibre signals, suggests that basal dendrites introduce relatively little filtering
901 [11, 28], as expected for large-diameter dendrites. We note that the pause in Golgi cell firing
902 following a discrete stimulus under anaesthesia [4, 74, 75] disappears from Golgi cell
903 recordings during locomotion [71].

904 During behaviour in freely-moving animals, mossy fibre activity is dense: a high fraction of
905 mossy fibres are active [76]. Both mossy fibres [77-79] and Golgi cells [71] have been reported
906 to fire with a sustained, time-varying signature. In the behaving animal, Golgi cell basal
907 dendritic membrane potential is a continuous variable under modulation by sustained inputs,
908 we submit. As charge transfer to the soma is passive, polarisation of the soma is likewise
909 sustained, under modulation by dendritic states.

910 It is worth noting that very high mossy fibre rates can be generated in experimental conditions.
911 However, the typical physiological range is 50–300 Hz [78]. In the simulations, we use the
912 physiological rates, within stated constraints on the shape of the distribution.

913 **Other control of Golgi cells?** Do Golgi cells receive a significant influence of input from other
914 sources? The evidence is incomplete but currently suggests other influence is absent or weak.
915 Contrary to early reports, neither Purkinje cells nor climbing fibres contact Golgi cells [80, who
916 give references], and only a modest minority of inhibitory inputs to Golgi cells (if any) are from
917 molecular layer interneurons [81], which generate weak synaptic currents [82], consistent with

918 either extremely weak or wholly absent innervation [83]. There is conflicting evidence whether
919 Golgi cells make inhibitory synaptic contact on each other [83, 84]. Our simulation does not
920 include an effect on Golgi cell firing by other sources of input.

921

922 **LINEAR TRANSMISSION OF GOLGI CELLS TO GRANULE CELLS**

923 'In the adult rat cerebellum, each granule cell dendrite receives 2.6 ± 0.55 synaptic contacts
924 from Golgi axon terminals' [30] citing [35].[1] However, the large majority of inhibition (98%) is
925 by spillover [5], where neurotransmitter released into the synaptic cleft spills out into the
926 glomerular space. Golgi cells release GABA, an inhibitory neurotransmitter. This is detected by
927 high-affinity $GABA_A$ receptors located perisynaptically and extrasynaptically on granule cells
928 [31, 32]. Even synaptically-received signals have most of their effect (i.e., the large majority of
929 charge transfer is mediated) via spillover.[2]

930 Golgi cells fire spontaneously [85] at a time-varying rate in the behaving animal, so that
931 downstream, a glomerulus receives continuous convergent input. As a result, there is a
932 sustained build-up of glomerular GABA during behaviour [86] at an adjustable concentration
933 controlled by the mean of afferent Golgi cell firing rates [87]. A glomerulus randomly samples
934 Golgi cells afferent to a field. Golgi cells provide the only inhibitory input to granule cells.

935 Signalling by spillover is sometimes assumed to be ambient and slow. However, the action of
936 glomerular GABA spillover has a fast phasic component – not as fast as synaptic transmission
937 (~1 ms) but with a rise time of only a few milliseconds [87]. Unlike the spiky appearance of
938 synaptically-induced inhibitory postsynaptic currents, spillover [4] generates a sustained

---

[1] Data are sparse. The original data are an estimate derived from visual inspection of slices thirty years ago so far without later corroboration, except from Rossi and Hamann 1998, who derive a lower estimate but propose the disparity is because their estimate is too low.

[2] Because of the long duration of IPSCs generated by spillover (recorded in slices), the total charge carried is three times that of IPSCs generated by directly connected terminals. 30. Rossi, D.J. and M. Hamann, *Spillover-mediated transmission at inhibitory synapses promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry.* Neuron, 1998. **20**(4): p. 783-95.

939 outward current. The response time to multiple inputs is likely faster than the experimentally
940 isolated response to a discrete signal (because the combined rate of GABA release is higher).

941 **Golgi cell oscillations.** Golgi cells fire autonomously. Under anaesthesia, firing falls into a slow
942 oscillating pattern [88]. Under excitatory synaptic input, however, this disappears [84].
943 Anaesthesia 'has a strong influence on spontaneous activity of Golgi cells' [80]. Discharge at
944 low rates with irregular timing under anaesthesia [85] is replaced by higher rates with more
945 regular timing without anaesthesia [70, 71, 89]. (This paragraph is a cautionary note about
946 evidence obtained under anaesthesia, which has been used to argue that oscillations provide a
947 form of signalling.)

948

949 **LINEAR TRANSMISSION OF MOSSY FIBRES TO GRANULE CELLS**

950 There is significant evidence that granule cell firing rates are linearly related to input rates, but
951 also evidence thought to be unfavourable to that conclusion. We take them in turn.

952 The short and equal length of granule cell dendrites would suggest light and equal filtering.
953 The mossy fibre-granule cell connection has a range of adaptations which support high-fidelity
954 transmission of high frequency signals across a wide bandwidth [9]. Faithful transmission of
955 rate information is reported [1, 2]. Vesicle release and replenishment are fast [10, 13] and
956 postsynaptic AMPA receptors operate in their linear range [13], where they are resistant to
957 desensitisation [90]. Multiple contacts are made by a mossy fibre with each granule cell [35].
958 Conversion of depolarising somatic charge into granule cell firing is a simple function, such that
959 granule cells 'have a relatively linear and uncomplicated conversion of depolarisation level to
960 spike rate' [39 p.2393, citing Jorntell and Ekerot 2006 and D'Angelo et al 1998].

961 Glutamate spillover enhances precision and reliability of transmission [13, 91]. During
962 behaviour, at sustained physiological mossy fibre rates, there may be a build-up of
963 intraglomerular glutamate, increasing the relative influence of spillover in a balance with
964 synaptic transmission, assisted by short-term synaptic depression. In these conditions,
965 spillover may dominate, as it dominates inhibitory transmission [5, 87]. Spillover from multiple

966 release sites (200-400 at a single mossy fibre terminal) is plausibly sufficient to mitigate
967 variability of vesicular size, release probability, and the number of release points at synapses
968 made on any individual cell, increasing fidelity and equality of excitation of granule cells.

969 Against that, there are reported to be heterogeneous synaptic weights. Different strength and
970 short-term dynamics of the mossy fibre to granule cell connection have been reported in
971 vestibular slices, with $GABA_A$ receptors blocked. The amplitude of the response, measured at
972 the soma, is reported to be statistically correlated to the source of the mossy fibre that is
973 stimulated [92], suggesting that the strength of the connection depends on the source.

974 However, the presence of weights does not necessarily mean that the function is to modulate
975 firing rates. Rather, it is suggested that the function may temporally code the source of inputs
976 because different combinations have different response onset times. An alternative, which
977 they also mention, is that some inputs may be in a supporting role, reminiscent of 'driver' and
978 'modulatory' signals in the thalamus and cortex [93]. Feasibly, signals in a supporting role must
979 be present for the postsynaptic granule cell to fire, but do not affect the granule cell firing rate.

Grants

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).

Acknowledgement

Dr Mike Evans, School of Mathematics, University of Leeds, LS2 9JT, UK, read an early draft of the paper and provided feedback.

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. Rancz, E.A., et al., *High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons.* Nature, 2007. **450**(7173): p. 1245-8.
2. Ritzau-Jost, A., et al., *Ultrafast action potentials mediate kilohertz signaling at a central synapse.* Neuron, 2014. **84**(1): p. 152-163.
3. Jelitai, M., et al., *Dendritic excitation-inhibition balance shapes cerebellar output during motor behaviour.* Nat Commun, 2016. **7**: p. 13722.
4. 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.
5. Duguid, I., et al., *Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex.* J Neurosci, 2012. **32**(32): p. 11132-43.
6. Walter, J.T. and K. Khodakhah, *The linear computational algorithm of cerebellar Purkinje cells.* J Neurosci, 2006. **26**(50): p. 12861-72.
7. 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.
8. Park, S.M., E. Tara, and K. Khodakhah, *Efficient generation of reciprocal signals by inhibition.* J Neurophysiol, 2012. **107**(9): p. 2453-62.
9. 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.
10. 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.
11. Cesana, E., et al. *Novel granule cell-Golgi cell excitatory input in the cerebellar granular layer*. in *FENS Abstr*. 2010.
12. 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.
13. 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.
14. 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.
15. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptotagmin 7 confers frequency invariance onto specialized depressing synapses.* Nature, 2017. **551**: p. 503.
16. Raymond, J.L. and J.F. Medina, *Computational Principles of Supervised Learning in the Cerebellum.* Annu Rev Neurosci, 2018. **41**: p. 233-253.
17. Alvina, K., et al., *Questioning the role of rebound firing in the cerebellum.* Nat Neurosci, 2008. **11**(11): p. 1256-8.
18. 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.

1037 19. Oscarsson, O., *Functional units of the cerebellum - sagittal zones and microzones.*
1038 Trends in Neurosciences, 1979. **2**: p. 143-145.
1039 20. Wu, H.S., I. Sugihara, and Y. Shinoda, *Projection patterns of single mossy fibers*
1040 *originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei.* J
1041 Comp Neurol, 1999. **411**(1): p. 97-118.
1042 21. Sultan, F. and D. Heck, *Detection of sequences in the cerebellar cortex: numerical*
1043 *estimate of the possible number of tidal-wave inducing sequences represented.* J
1044 Physiol Paris, 2003. **97**(4-6): p. 591-600.
1045 22. Shinoda, Y. and I. Sugihara, *Axonal Trajectories of Single Climbing and Mossy Fiber*
1046 *Neurons in the Cerebellar Cortex and Nucleus*, in *Handbook of the Cerebellum and*
1047 *Cerebellar Disorders*, S.J.D. Manto M., Rossi F., Gruol D.L., Koibuchi N., Editor. 2013,
1048 Springer, Dordrecht.
1049 23. Shinoda, Y., et al., *The entire trajectory of single climbing and mossy fibers in the*
1050 *cerebellar nuclei and cortex.* Prog Brain Res, 2000. **124**: p. 173-86.
1051 24. Garwicz, M., C.F. Ekerot, and H. Jorntell, *Organizational principles of cerebellar*
1052 *neuronal circuitry.* News in Physiological Sciences, 1998. **13**: p. 28-34.
1053 25. Ozden, I., et al., *Reliable coding emerges from coactivation of climbing fibers in*
1054 *microbands of cerebellar Purkinje neurons.* J Neurosci, 2009. **29**(34): p. 10463-73.
1055 26. Dean, P., et al., *The cerebellar microcircuit as an adaptive filter: experimental and*
1056 *computational evidence.* Nat Rev Neurosci, 2010. **11**(1): p. 30-43.
1057 27. Eccles, J.C., M. Ito, and J. Szentágothai, *The cerebellum as a neuronal machine*. 1967,
1058 Berlin, New York etc.: Springer-Verlag. 335 p.
1059 28. Kanichay, R.T. and R.A. Silver, *Synaptic and cellular properties of the feedforward*
1060 *inhibitory circuit within the input layer of the cerebellar cortex.* J Neurosci, 2008.
1061 **28**(36): p. 8955-67.
1062 29. Dieudonné, S., *Submillisecond kinetics and low efficacy of parallel fibre-Golgi cell*
1063 *synaptic currents in the rat cerebellum.* J Physiol, 1998. **510 (Pt 3)**: p. 845-66.
1064 30. Rossi, D.J. and M. Hamann, *Spillover-mediated transmission at inhibitory synapses*
1065 *promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry.*
1066 Neuron, 1998. **20**(4): p. 783-95.
1067 31. Nusser, Z., W. Sieghart, and P. Somogyi, *Segregation of different GABAA receptors to*
1068 *synaptic and extrasynaptic membranes of cerebellar granule cells.* J Neurosci, 1998.
1069 **18**(5): p. 1693-703.
1070 32. Brickley, S.G., S.G. Cull-Candy, and M. Farrant, *Development of a tonic form of synaptic*
1071 *inhibition in rat cerebellar granule cells resulting from persistent activation of GABAA*
1072 *receptors.* J Physiol, 1996. **497 (Pt 3)**: p. 753-9.
1073 33. Gilbert, M. and R.C. Miall, *How and Why the Cerebellum Recodes Input Signals: An*
1074 *Alternative to Machine Learning.* Neuroscientist, 2021a. **DOI:**
1075 **10.1177/1073858420986795**.
1076 34. Gao, Z., et al., *Excitatory Cerebellar Nucleocortical Circuit Provides Internal*
1077 *Amplification during Associative Conditioning.* Neuron, 2016. **89**(3): p. 645-57.
1078 35. Jakab, R.L. and J. Hámori, *Quantitative morphology and synaptology of cerebellar*
1079 *glomeruli in the rat.* Anat Embryol (Berl), 1988. **179**(1): p. 81-8.

36. Mitchell, S.J. and R.A. Silver, *GABA spillover from single inhibitory axons suppresses low-frequency excitatory transmission at the cerebellar glomerulus.* J Neurosci, 2000a. **20**(23): p. 8651-8.
37. Mitchell, S.J. and R.A. Silver, *Glutamate spillover suppresses inhibition by activating presynaptic mGluRs.* Nature, 2000b. **404**(6777): p. 498-502.
38. Chadderton, P., T.W. Margrie, and M. Hausser, *Integration of quanta in cerebellar granule cells during sensory processing.* Nature, 2004. **428**(6985): p. 856-60.
39. Bengtsson, F. and H. Jörntell, *Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs.* Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2389-94.
40. Billings, G., et al., *Network structure within the cerebellar input layer enables lossless sparse encoding.* Neuron, 2014. **83**(4): p. 960-74.
41. 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.
42. 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.
43. Albus, J.S., *A Theory of Cerebellar Function.* Mathematical Biosciences, 1971. **10**(1-2): p. 25-61.
44. Gilbert, M. and A. Rasmussen, *Gap Junctions May Have A Computational Function In The Cerebellum: A Hypothesis.* The Cerebellum, 2024.
45. 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.
46. Harvey, R.J. and R.M. Napper, *Quantitative studies on the mammalian cerebellum.* Prog Neurobiol, 1991. **36**(6): p. 437-63.
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. Napper, R.M. and R.J. Harvey, *Quantitative study of the Purkinje cell dendritic spines in the rat cerebellum.* J Comp Neurol, 1988. **274**(2): p. 158-67.
49. 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.
50. 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.
51. Dugué, G.P., et al., *Cerebellar re-encoding of self-generated head movements.* Elife, 2017. **6**.
52. Chen, S., G.J. Augustine, and P. Chadderton, *The cerebellum linearly encodes whisker position during voluntary movement.* eLife, 2016. **5**: p. e10509-e10509.
53. Herzfeld, D.J., et al., *Encoding of action by the Purkinje cells of the cerebellum.* Nature, 2015. **526**(7573): p. 439-42.
54. Hong, S., et al., *Multiplexed coding by cerebellar Purkinje neurons.* Elife, 2016. **5**.
55. Fujita, M., *Adaptive filter model of the cerebellum.* Biol Cybern, 1982. **45**(3): p. 195-206.
56. Marr, D., *A theory of cerebellar cortex.* J Physiol, 1969. **202**(2): p. 437-70.
57. Apps, R. and M. Garwicz, *Anatomical and physiological foundations of cerebellar information processing.* Nat Rev Neurosci, 2005. **6**(4): p. 297-311.

1124 58. Apps, R. and R. Hawkes, *Cerebellar cortical organization: a one-map hypothesis.* Nat
1125 Rev Neurosci, 2009. **10**(9): p. 670-81.
1126 59. Ji, Z. and R. Hawkes, *Topography of purkinje cell compartments and mossy fiber*
1127 *terminal fields in lobules ii and iii of the rat cerebellar cortex: Spinocerebellar and*
1128 *cuneocerebellar projections.* Neuroscience, 1994. **61**(4): p. 935-954.
1129 60. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from Myxinoids*
1130 *through Birds*. 1967: The University of Minnesota Press.
1131 61. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from*
1132 *Monotremes through Apes*. 1970: The University of Minnesota Press, Minneapolis.
1133 62. Mugnaini, E., *The length of cerebellar parallel fibers in chicken and rhesus monkey.* J
1134 Comp Neurol, 1983. **220**(1): p. 7-15.
1135 63. Brand, S., A.L. Dahl, and E. Mugnaini, *The length of parallel fibers in the cat cerebellar*
1136 *cortex. An experimental light and electron microscopic study.* Exp Brain Res, 1976.
1137 **26**(1): p. 39-58.
1138 64. Hámori, J., *Anatomy and neurochemical anatomy of the cerebellum*, in *Cerebellar*
1139 *degenerations : clinical neurobiology*, A. Plaitakis, Editor. 1992, Kluwer Academic.
1140 65. Palay, S.L. and V. Chan-Palay, *Cerebellar cortex : cytology and organization*. 1974,
1141 Berlin: Springer.
1142 66. Lange, W., *Regional differences in the distribution of golgi cells in the cerebellar cortex*
1143 *of man and some other mammals.* Cell Tissue Res, 1974. **153**(2): p. 219-26.
1144 67. Barmack, N.H. and V. Yakhnitsa, *Functions of interneurons in mouse cerebellum.* J
1145 Neurosci, 2008. **28**(5): p. 1140-52.
1146 68. Hámori, J. and J. Szentágothai, *Participation of Golgi neuron processes in the cerebellar*
1147 *glomeruli: an electron microscope study.* Exp Brain Res, 1966. **2**(1): p. 35-48.
1148 69. D'Angelo, E., et al., *The cerebellar Golgi cell and spatiotemporal organization of*
1149 *granular layer activity.* Front Neural Circuits, 2013. **7**: p. 93.
1150 70. Miles, F.A., et al., *Long-term adaptive changes in primate vestibuloocular reflex. III.*
1151 *Electrophysiological observations in flocculus of normal monkeys.* J Neurophysiol, 1980.
1152 **43**(5): p. 1437-76.
1153 71. Edgley, S.A. and M. Lidierth, *The discharges of cerebellar Golgi cells during locomotion*
1154 *in the cat.* J Physiol, 1987. **392**: p. 315-32.
1155 72. Galliano, E., P. Mazzarello, and E. D'Angelo, *Discovery and rediscoveries of Golgi cells.* J
1156 Physiol, 2010. **588**(Pt 19): p. 3639-55.
1157 73. Eccles, J.C., R. Llinás, and K. Sasaki, *Golgi cell inhibition in the cerebellar cortex.* Nature,
1158 1964. **204**: p. 1265-6.
1159 74. Holtzman, T., et al., *Different responses of rat cerebellar Purkinje cells and Golgi cells*
1160 *evoked by widespread convergent sensory inputs.* J Physiol, 2006. **574**(Pt 2): p. 491-
1161 507.
1162 75. Vos, B.P., A. Volny-Luraghi, and E. De Schutter, *Cerebellar Golgi cells in the rat:*
1163 *receptive fields and timing of responses to facial stimulation.* Eur J Neurosci, 1999.
1164 **11**(8): p. 2621-34.
1165 76. Hartmann, M.J. and J.M. Bower, *Tactile responses in the granule cell layer of cerebellar*
1166 *folium crus IIa of freely behaving rats.* J Neurosci, 2001. **21**(10): p. 3549-63.
1167 77. Arenz, A., et al., *The contribution of single synapses to sensory representation in vivo.*
1168 Science, 2008. **321**(5891): p. 977-80.

1169 78. van Kan, P.L., A.R. Gibson, and J.C. Houk, *Movement-related inputs to intermediate*
1170 *cerebellum of the monkey.* J Neurophysiol, 1993. **69**(1): p. 74-94.
1171 79. Giovannucci, A., et al., *Cerebellar granule cells acquire a widespread predictive*
1172 *feedback signal during motor learning.* Nat Neurosci, 2017.
1173 80. Pietrajtis, K. and S. Dieudonné, *Golgi Neurons*, in *Handbook of the Cerebellum and*
1174 *Cerebellar Disorders*, M. Manto, et al., Editors. 2013, Springer Netherlands: Dordrecht.
1175 p. 829-852.
1176 81. Eyre, M.D. and Z. Nusser, *Only a Minority of the Inhibitory Inputs to Cerebellar Golgi*
1177 *Cells Originates from Local GABAergic Cells.* eNeuro, 2016. **3**(2).
1178 82. Dumoulin, A., A. Triller, and S. Dieudonné, *IPSC kinetics at identified GABAergic and*
1179 *mixed GABAergic and glycinergic synapses onto cerebellar Golgi cells.* J Neurosci, 2001.
1180 **21**(16): p. 6045-57.
1181 83. Hull, C. and W.G. Regehr, *Identification of an inhibitory circuit that regulates cerebellar*
1182 *Golgi cell activity.* Neuron, 2012. **73**(1): p. 149-58.
1183 84. Vervaeke, K., et al., *Rapid desynchronization of an electrically coupled interneuron*
1184 *network with sparse excitatory synaptic input.* Neuron, 2010. **67**(3): p. 435-51.
1185 85. Ruigrok, T.J., R.A. Hensbroek, and J.I. Simpson, *Spontaneous activity signatures of*
1186 *morphologically identified interneurons in the vestibulocerebellum.* J Neurosci, 2011.
1187 **31**(2): p. 712-24.
1188 86. Rossi, D.J., M. Hamann, and D. Attwell, *Multiple modes of GABAergic inhibition of rat*
1189 *cerebellar granule cells.* J Physiol, 2003. **548**(Pt 1): p. 97-110.
1190 87. Mapelli, L., S. Solinas, and E. D'Angelo, *Integration and regulation of glomerular*
1191 *inhibition in the cerebellar granular layer circuit.* Front Cell Neurosci, 2014. **8**: p. 55.
1192 88. Ros, H., et al., *Neocortical networks entrain neuronal circuits in cerebellar cortex.* J
1193 Neurosci, 2009. **29**(33): p. 10309-20.
1194 89. Rasmussen, A., et al., *Golgi cell activity during eyeblink conditioning in decerebrate*
1195 *ferrets.* Cerebellum, 2014. **13**(1): p. 42-5.
1196 90. DiGregorio, D.A., et al., *Desensitization properties of AMPA receptors at the cerebellar*
1197 *mossy fiber granule cell synapse.* J Neurosci, 2007. **27**(31): p. 8344-57.
1198 91. DiGregorio, D.A., Z. Nusser, and R.A. Silver, *Spillover of glutamate onto synaptic AMPA*
1199 *receptors enhances fast transmission at a cerebellar synapse.* Neuron, 2002. **35**(3): p.
1200 521-33.
1201 92. Chabrol, F.P., et al., *Synaptic diversity enables temporal coding of coincident*
1202 *multisensory inputs in single neurons.* Nat Neurosci, 2015. **18**(5): p. 718-27.
1203 93. Sherman, S.M. and R.W. Guillery, *Distinct functions for direct and transthalamic*
1204 *corticocortical connections.* J Neurophysiol, 2011. **106**(3): p. 1068-77.

1205