THE CEREBELLUM CONVERTS INPUT DATA INTO A HYPER LOW-RESOLUTION GRANULE CELL CODE WITH SPATIAL DIMENSIONS: A HYPOTHESIS

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

We present a theory of the inner layer of the cerebellar cortex, the granular layer, where the main excitatory input to the cerebellum is received. We ask how input signals are converted into an internal code and what form that has. 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. We find that input data are converted to a hyper low-resolution internal code. Information is coded in the joint activity of large cell groups and therefore has minimum spatial dimensions – the dimensions of a code group. The conversion exploits statistical effects of random sampling. Code group dimensions are an effect of topography, cell morphologies and granular layer architecture. The activity of a code group 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 is unexpectedly wasteful – there is a huge sacrifice of resolution – but may be a solution to fundamental problems involved in the biological representation of information.

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. In any event, this is not regarded as 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 will argue, *inter alia*, that these influential ideas may need to be reconsidered.

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 neuroanatomy of the granular layer as a strategy that automates the granular layer computation. The computation, in this contention, is the passive result of anatomy (cell morphologies and neuroanatomical architecture) and linear signalling. 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. Yet despite the potential to be dysfunctionally noisy, 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 [1-7], and that cells are adapted for faithful transmission [8-14]. Firing rates that linearly code behavioural and task metrics have been observed 'at all levels of the cerebellar circuit' [15 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 [12, 13, 16, 17]. 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 short review in a dedicated section before the Discussion 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. Parameter values (usually a range) are usually available from the literature.

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 for support). We work in reverse. We propose first a mechanism that can explain the evidence, then infer function from the mechanism. In our process, high-level ideas come last. We use modelling in a different role, to quantify and test the ideas by simulating the mechanism. In our approach, the hypothesis and simulation are *both* the mechanism/model. The simulation is the hypothesis in quantified form. While there is a computational element, and the ideas are quantified with a computer simulation, it is not primarily a computational paper.

Note that we do not propose to be mathematically interesting. The heart of the proposal is that the physiological computation is in fact not sophisticated, against expectation – at least, there is a plausible argument that this is the case, which we

present as an alternative to current models. For that reason, physiology has a high profile in the paper. The sophistication of the biological design is found in the fact that physiology is adapted to isolate computationally relevant parameters (and computations themselves) from interference by other variables. Biophysical events which accomplish this result are complex. If we are correct, a future challenge will be to add them, so that the model is populated with a smaller scale of detail. Here, however, they are superfluous.

The next section describes our methodology in more detail. We encourage the reader to read the first subsection (2.1 'The relationship of physiology and model parameters') before proceeding because it describes how the computation is related to the physical hardware of the system.

2. Methods

2.1 The relationship of physiology and model parameters

2.1.1 SUMMARY OF METHODOLOGY

We propose a physiologically detailed hypothesis which we recreate *in silico* in the form of a network of unit functions. Functions are attributed to identified anatomical hardware, connected by known anatomy, and the functions themselves are

individually evidenced. In this form, we are able to model large cell networks in high *relevant* definition, without simplifying assumptions. 'Relevant' is italicised to emphasise that parameter values represent data – numbers used to code information – and not biophysics. Hence the claim that the *computation* is not simplified.

We argue that the computation is a passive effect of anatomy. The simulation provides a demonstration that processing by anatomy provides a mechanism that (i) self-regulates the fraction of granule cells that fire at any time, and (ii) converts input signals into an unexpectedly low-resolution internal code with spatial dimensions.

2.1.2 COMPUTATIONAL ARCHITECTURE

The granular layer receives input to the cerebellum and converts it to an internal code.

Input is received from mossy fibres. Mossy fibres contact and excite granule cells and Golgi cells. Golgi cells inhibit granule cells. Excitation of Golgi cells and excitation and inhibition of granule cells all happen in a structure called a glomerulus.

Granular layer anatomy is the physical form of an orderly network of computational units. A unit takes inputs and returns output. Inputs and output represent values (that is, numbers). A unit is the site of biophysical events which relate inputs and output – the unit function.

Physically, the granular layer appears anatomically seamless – a carpet of cells. But they are wired so that units are organised functionally in layers. The outputs of a layer provide the data sampled (i.e., received as inputs) by units in the next layer.

Data are constantly refreshed. Computations in a layer run in parallel, continuously, so unit outputs are analogue. This occurs in all layers at once. The aggregate of computations relates input and output of the system: the (granular layer) network computation.

All layers of unit functions span the network, forming a long, thin strip. All strips are sagittal, and therefore parallel. Information is coded collectively in the concurrent outputs of a layer. The code is a frequency distribution, so that a whole layer, at any given instant in time, codes information that is both indivisible (cannot be subdivided into smaller units of information) and yet coded in any random sample of unit output values. The biophysical form of outputs depends on the layer.

2.1.3 HOW ARE UNITS CONNECTED?

By anatomy, so we know what the connections are, because anatomy has been reported in detail. The number of inputs to a unit is usually anatomically randomly variable in a reported range. In some cases, we simulate that by generating the number with a distributed probability, and in others we use the mean.

It is unnecessary to know which individual cells contact each other because (biologically and in the model) contact is at random inside topographically defined

148 dimensions. This has the asset for modelling that a target population can be
149 represented by randomly sampling data received as input to a location.

150 Biologically, a location is a volume with spatial dimensions. *In silico*, we can represent
151 that as a population of unit functions. We know from anatomy (or can derive) the size
152 of unit populations, population ratios, and convergence and divergence ratios.

153 **2.1.4 WHAT DETERMINES THE SHAPE AND SIZE OF A POPULATION OF UNIT**
154 **FUNCTIONS?**

155 The cerebellar cortex is anatomically seamless — there are no physical boundaries that
156 divide cells into groups, just organisation that reflects topography. Microzones are an
157 example (microzones are defined by their climbing fibre input) [18].

158 Microzones are divisions of the molecular layer. Mapping studies suggest that the
159 granular layer may have similar organisation, but the picture is less clear [19].
160 However, the morphology and termination pattern of mossy fibres suggest a
161 functional division into strips that may mirror microzones.

162 Mossy fibres give rise to sagittally-extending collaterals, the direction of the long axis
163 of microzones (Fig 1). Collaterals branch terminally, and each branch ends in a cluster
164 of terminals, so that a single cell terminates in a spaced-out row of terminal clusters
165 which are always lined up in the same direction (parallel to microzones but in the
166 granular layer) [20-23]. Clusters are separated by a randomly variable distance typically

167 not less than 200 µm [20, 21] measured from edge to edge of clusters (so they are
168 separated by a greater distance measured from centre to centre). Terminals are small
169 compared to the average volume occupied by a cluster, and intimately and randomly
170 intermingled with terminals of other mossy fibres.

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

175 The dimensions which enclose a cluster (cluster size) are variable (because it depends
176 on the number of terminals in a cluster), averaging 200 µm sagittally x 150 µm
177 mediolaterally [21]. Say we divide the granular layer into strips and strips into 'fields'
178 (each perhaps 1/100th part of a strip, but still each containing thousands of granule
179 cells). Fields are average cluster size. Fields are simply nominal, a convenient modelling
180 device. But strips are functional – the minimum dimensions at which the cerebellum
181 makes any effort to tell signals apart, we submit.

182 The shape and size of a population of unit functions are a consequence of this
183 topography – the dimensions of a strip where mossy fibre terminals are mixed up at
184 random. This is the first way that mossy fibre morphology is the key to the cerebellar
185 code. A sagittal row of 75–100 fields has the area (viewed from the cerebellar surface)

of a mid-range microzone [18, 24-26]. We term a row of fields a 'rank' (strip has been used in the literature to refer to other divisions of the cerebellar cortex).

2.1.5 SAMPLING WITH REPLACEMENT

2.1.5.1 Terminal branching mimics independent sampling by a field of input rates to a rank

We propose that terminal branching of mossy fibre collaterals simulates independent random sampling, by fields, of input rates to a rank. That is, inclusion of a rate in a sample does not affect the probability that it is including in any other sample, also called sampling with 'replacement'.

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 would mean any number of fields can receive the same signal, up to all of them. Clearly this is not true – in reality there is an anatomical limit. However, even if there was real replacement, the number of repeat samplings is constrained by a probability distribution – effectively, a cap. 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 rates were in fact sampled with replacement but with a probability (that a given sample contains a given rate) that reflects the average number of terminal branches per cell. 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

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

where N is the number of fields per rank, each receiving innervation by a number of mossy fibres, x, given by the derivation of mossy fibre rates in section 2.2, and assuming m = 5. 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 value of m is consistent with what we know about the biological range of the number of terminal clusters per mossy fibre [21, 27]. However, some different value would still make the same point. In hypothetical conditions in

which it is theoretically possible for a mossy fibre signal to be selected any number of times, the number is restricted statistically to a small range. The number has the same distributed probability as the number of mossy fibre terminal branches per cell. Physiology in this way mimics sampling with replacement.

2.1.5.2 Terminal clustering mimics independent sampling by single cells of input rates to a field

Granule cells and Golgi cells both receive contact from a random sample of mossy fibres which innervate a field. We propose that mossy fibre terminal clustering simulates independent random sampling, by individual cells, of input rates to a field. The principle is the same here as with terminal branching, 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 other cells cannot also receive the same signal. 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 terminal 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 mossy fibre terminal, and we estimate that Golgi cells receive an average of 4 inputs to each of their basal dendrites (see derivation of

parameters in later sections). 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 'Derivation of mossy fibre rates' later in this section).

2.1.5.3 The combined effect of terminal branching and terminal clustering

The combined effect is a biological facsimile of simultaneous random sampling with replacement by individual granule cells and Golgi cells of firing rates received as input to a whole rank, notwithstanding that they each receive fixed contact from a tiny fraction of mossy fibres. Note that it is the population of mossy fibre *rates* that are hypothetically sampled with replacement, and not *signals.* The population of signals includes copies of rates because of branching and clustering.

2.1.6 WHAT IS THE COMPUTATION?

So: the effect of duplicating mossy fibre signals – by terminal branching and clustering – is to mimic sampling with replacement. Granule cells and Golgi cells each individually, effectively (i.e., for granular layer computational purposes), independently randomly sample mossy fibre input rates to a whole rank. This is the second way that mossy fibre morphology is the key to the cerebellar code.

This has the result that unit functions provide the physiological instrument of (continuously) computing (a value linearly related to) the sum of independent random samples of the afferent data layer. Doing this many times (in parallel) generates a distribution with a narrower range and a linearly related mean, by the central limit theorem.

Repeating that layer on layer progressively narrows the data range, while conserving a linear relationship of all the means. The range of granule cell rates is further narrowed because granule cells which fire – a small fraction – are confined (by the glomerular competition – see section 5) to those which receive mossy fibre signals at the higher end of the mossy fibre range.

2.2 Derivation of mossy fibre rates

2.2.1 RATIONALE

We run the computation *in silico*. We represent the granular layer as a network of unit computations organised in layers, as described. Support for the size of unit populations, population ratios, and convergence and divergence ratios is cited in context, in the relevant section of the main text. Support for unit functions is discussed in the section before the Discussion.

The simulation is not an additional step in ideas or detail. It is a quantified form of the hypothesis. The hypothesis describes a computational mechanism; the simulation performs the computations. The purpose of the simulation is to confirm that the mechanism is feasible, with plausible evidence. This section describes the derivation of mossy fibre rates received by a rank and by fields taking account of terminal branching and terminal clustering, the randomly variable number of terminals per cluster and random variation of the number of terminals in a cluster which fall inside and outside a field when a cluster straddles a field boundary.

The frequency distribution of mossy fibre rates received as input to a granular layer rank is unreported. We used a large bandwidth in the physiological range and varied the shape of the distribution to confirm that the operation of the mechanism does not

depend on the input to it. The shape of the mossy fibre distribution is included in all figures.

2.2.2 THE NUMBER OF MOSSY FIBRES THAT INNERVATE A RANK

Each field receives a random sample mossy fibre firing rates received as input to a rank. To simulate that, we need the size of the sampled population – the number of mossy fibre signals received by a rank – and sample size, the number received by a field.

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

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

We work from an estimate of the total number of terminals per field, which is, in turn, derived from convergent estimates of the number of granule cells per field. (Estimate 1) There are an estimated 1.92×10^6 granule cells per µl [28]. The number of fields which fit in that volume is $(1{,}000/150)^2 \times (1{,}000/200) = \sim 222$, assuming a field depth of

312 150 µm [21]. The number of granule cells in a field is, therefore 1,920,000/222 = 8,649.
313 (Estimate 2) An estimated 350,000 parallel fibres pass through a single Purkinje cell's
314 territory [28]. Parallel fibres run about 3 mm in each direction from the granule cell
315 soma [29, 30]. The Purkinje cell dendritic arbour fills the molecular layer vertically and
316 averages ~200 µm wide. Therefore, 6,000/150 = 40 fields fit in a row that provides
317 parallel fibres that pass through a Purkinje cell dendritic territory, giving 350,000/40 =
318 8,750 granule cells per field. Estimates vary of the number of granule cells which
319 extend a dendrite into a single glomerulus. 50 is mid-range. (8,750/50) x 4 = 700
320 glomeruli per field.

321 Granule cells have 3–5 dendrites, averaging 4, and receive input to each of them from
322 a single mossy fibre (so an average granule cell receives 4 inputs). Estimates of
323 divergence of a single mossy fibre terminal onto granule cells vary [1]. It is unclear how
324 much this reflects biological diversity versus differences in opinion and methods. Here
325 we use a mid-range estimate of 1:50 [31, 32]. Assuming a mossy fibre terminal receives
326 a single granule cell dendrite from each of d granule cells, and given an average of 4
327 dendrites per granule cell and z granule cells per field, the total number of mossy fibre
328 terminals in a field is $N = (z \times 4)/d = 700$ given $z = 8{,}750$ and $d = 50$. A convergent
329 estimate is provided by multiplying the per field number of mossy fibres by the
330 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 2 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 x 9)) = 0.159 that a mossy fibre contributes 5 terminals to a field (regardless of cluster size). Similarly, there is a probability of ((6 x 28)/((6 x 28) + 27 + 25 + 22))/7 = 0.099 that a given cell contributes 7 terminals to a field.

Table 2. 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	

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 fraction of terminals provided by a given cluster size is $x_j/(x_4 + x_5 + \cdots + x_{12})$, and this fraction is multiplied by N to give the number of terminals. The number of terminals is divided by the expected mean number of terminals that fall inside field limits (for that cluster size) to give the expected number of mossy fibres that provide them. The number of j-terminal mossy fibres that innervate a field is thus given by

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

360 The results are shown in Table 3. The equal number of mossy fibres (mfs) in each
361 column is what we would expect, because we made it an assumption that all sizes
362 occur with equal probability, and mf values are expected values – i.e., in proportions
363 predicted by their probability. The mf total is the expected number of mossy fibres
364 afferent to a field (which contribute at least one terminal).

365 **Table 3. 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

366

367 R2 ratios – the relative incidence of afferent mossy fibres that contribute 7 terminals
368 to a field, 6 terminals, 5 and so on, down to 1 – and the expected number of mossy
369 fibres that contribute them, are shown in Table 4.

370 **Table 4. 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	

371

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

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

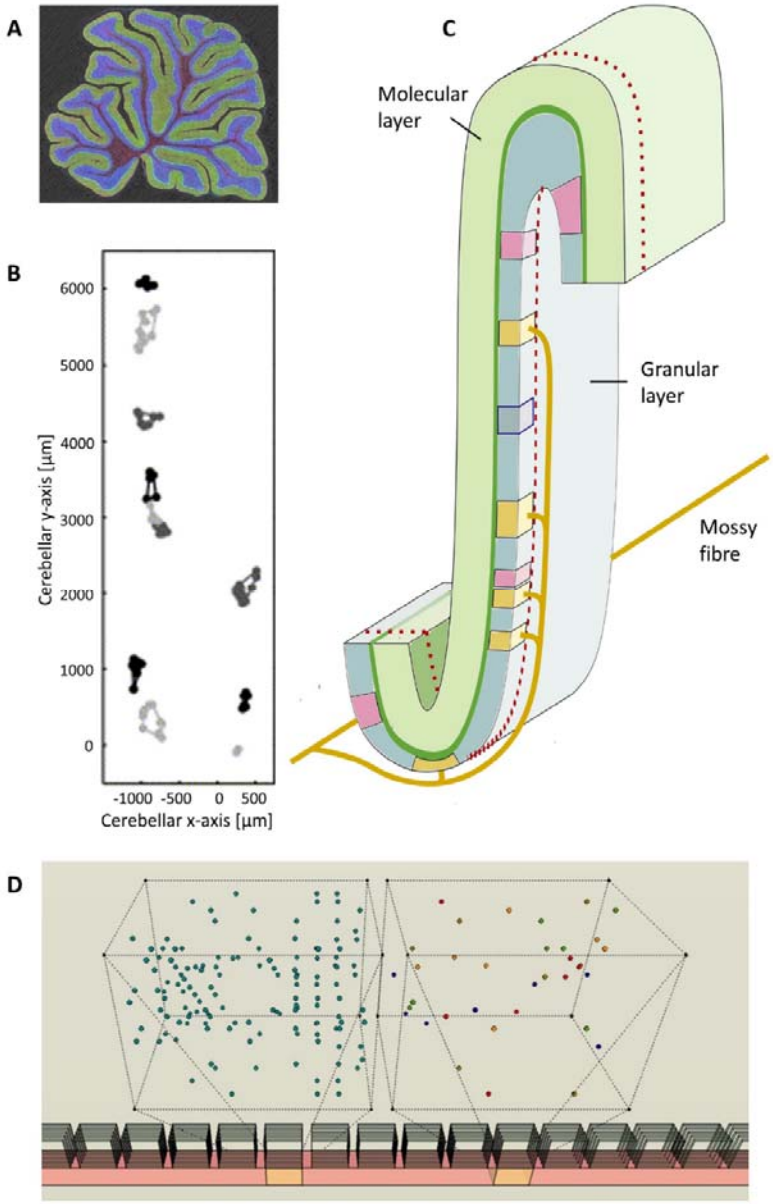
Assumptions. Cluster sizes occur with equal probability. The probabilities of each outcome when a cluster straddles field boundaries (how many terminals are inside and how many outside) are as stated. A single mossy fibre ends in an average of 5 terminal clusters per rank.

2.2.3 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 each number of terminals 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 physiological distance, so that a field cannot contain terminals of more than one cluster per mossy fibre. Fields receive a randomly variable number of copies of most signals in the cerebellum and the simulation. 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.



408

409 **Fig 1. Mossy fibre termination. A** Sagittal section of the rat cerebellum (obtained by
410 multi-photon microscopy, courtesy Tom Deerinck). The outer layer of the cerebellar
411 cortex, the molecular layer, is stained green and the inner layer, the granular layer, is
412 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. Cerebellar x-axis: parallel fibre axis; Cerebellar y-axis: parasagittal axis. **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. The Golgi cell computation

3.1 Hypothesis

Despite forming an anatomically seamless carpet of cells, Golgi cells are organised functionally in groups: ensembles. An ensemble is the population of a sagittal row of three fields. This is the minimum unit of the Golgi cell computation. Golgi cell morphology creates ensemble dimensions. The Golgi cell axon ramifies profusely, giving rise to a dense plexus. The axonal field is sagittally elongated – that is, in the

435 direction of the long axis of a microzone (mean range 650 +/- 179 µm by 180 +/- 40 µm
436 [in mice: 33]) – and is the depth of the granular layer, which it fills vertically. Because
437 of the dimensions and orientation of the Golgi cell plexus, all Golgi cells in a sagittal
438 row of three fields inhibit substantially the whole of the middle field. Conversely, each
439 field defines a functional population of Golgi cells – an ensemble – being those cells
440 with convergent input to that field. A single Golgi cell contacts thousands of granule
441 cells [34, 35].

442 Each mossy fibre terminal is ensheathed by a semi-permeable membrane which
443 restricts neurotransmitter diffusion, a structure termed a glomerulus. Excitation of
444 granule cells and Golgi cells by mossy fibres, and inhibition of granule cells by Golgi
445 cells, all take place here. Fine, beaded axon fibres enter glomeruli, where they inhibit
446 granule cells [34]. 'In the adult rat cerebellum, each granule cell dendrite receives 2.6 ±
447 0.55 synaptic contacts from Golgi axon terminals' [36] citing [32].[1] However, the large
448 majority of inhibition (98%) is by spillover [4], where neurotransmitter released into
449 the synaptic cleft spills out into the glomerular space. Golgi cells release GABA, an
450 inhibitory neurotransmitter. This is detected by high-affinity $GABA_A$ receptors located
451 perisynaptically and extrasynaptically on granule cells [37, 38]. Even synaptically-

[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.

452 received signals have most of their effect (i.e., the large majority of charge transfer is
453 mediated) via spillover.[2]

454 So, an ensemble converts mossy fibre rates received by a three-field row into
455 inhibition of granule cells in the middle field, almost exclusively via spillover. The
456 computation exploits statistical effects called the central limit theorem. Say a
457 population of values (which can have any frequency distribution) is randomly sampled.
458 There are *n* samples, sample size *m*. If we take the mean (or sum) of each sample, and
459 we take enough samples, the sample means have a normal or near-normal
460 distribution. (This works even if *m* is a range.) The new distribution is centred on the
461 same mean as the sampled distribution, but narrower.

462 There are three unit populations, each a step in the ensemble computation. First, Golgi
463 cell basal dendrites each quasi-independently randomly sample mossy fibre rates; unit
464 output is a sustained dendritic state which is, at any given instant in time, a linear
465 function of the sample mean. The unit at the second step is the Golgi cell soma; unit
466 output is a linear function of the mean of dendritic states in the physiological form of
467 Golgi cell firing rates, reflecting a sustained somatic state. Third, middle-field glomeruli
468 each quasi-independently randomly sample firing rates of Golgi cells afferent to the

[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. 36. 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.

field. The effect of the computation is to synchronise inhibition of granule cells in the middle field.

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 randomly sampled from the same mossy fibre distribution. Unit layers also conflate, so that a unit population extends the length of a rank. The result is to synchronise inhibition of granule cells in the middle field: the Golgi cell population of a rank functions as a single unit.

See section 6.1 for derivation of parameters not given here.

3.2 Results

3.2.1 AN ENSEMBLE

Mossy fibre rates sampled by Golgi cells in each field were obtained by sampling from a distribution representing signals received as input to a rank, which were randomly generated subject to constraints on the shape of the distribution, and allowing for mossy fibre terminal branching and clustering, as described in the Methods section (Fig 2 column 1).

This gave the number of mossy fibre signals, firing rates and the number of copies of each rate received by each field, and therefore input to an ensemble territory,

populated by 30 Golgi cells, 10 per field. There are ~700 mossy fibre terminals per field. As far as we know (and we propose), contact by mossy fibres on Golgi cells is at random at field scale, i.e., signals received by a single basal dendrite are provided by a random sample of local terminals. The number of physiological samplings is equal to the number of basal dendrites per field; sample size is the number of mossy fibres which make contact on each dendrite. The effect is sustained dendritic depolarisation under modulation by a linear function of input rates, which we take as the mean. Sources for this paragraph are cited in section 6.1.

Taking the sample means generates a new set of values (Fig 2 column 2) with a smaller range and centred on the mean, which is linearly related to the mean of the sampled distribution. Basal dendritic charge transfer is passive and little filtered [10, 39]. Integration of dendritic states at the Golgi soma is represented by taking the mean of dendritic charge, for each simulated cell. Golgi cell firing rates have a linear relationship with the amplitude of depolarising current [40], suggesting (and we assume) a linear conversion of somatic membrane potential into the Golgi cell firing rate.

The output is a third data set, Golgi cell firing rates (Fig 2 column 3), on the same normalised scale. Each glomerulus in the middle field receives innervation from a random subset of Golgi cells afferent to the field, and accordingly a random sample of Golgi cell firing rates, mimicking replacement. Golgi cells inhibit granule cells in the

507 glomerulus. A granule cell extends a single dendrite into a glomerulus, typically
508 receiving synaptic contact from two or three Golgi cells [36]. However, inhibition of
509 granule cells by Golgi cells is almost entirely by GABA spillover into the intraglomerular
510 space [4], via extrasynaptic receptors [37, 38], as noted. This effectively increases the
511 Golgi cell to granule convergence ratio.

512 Golgi cells fire [41] at a time-varying rate in the behaving animal, so that a glomerulus
513 receives continuous input. As a result, there is a sustained build-up of glomerular
514 GABA during behaviour [42] at an adjustable concentration controlled by Golgi cell
515 firing rates [43]. Signalling by spillover is sometimes assumed to be ambient and slow.
516 However, the action of glomerular GABA spillover has a fast phasic component – not as
517 fast as synaptic transmission (~1 ms) but with a rise time of only a few milliseconds
518 [43]. Unlike the spiky appearance of synaptically-induced inhibitory postsynaptic
519 currents, spillover [3] generates a sustained outward current.

520 Spillover concentration is under continuous modulation at a level which we take as a
521 linear function of the mean of afferent Golgi cell rates (the sum also works). Sample
522 size is provided by the Golgi cell to glomerulus convergence ratio. The number of Golgi
523 cells afferent to a glomerulus is unreported, as far as we know. We take it as a random
524 number in the range 8–12, generated for each glomerulus. We tested other ratios to
525 confirm that the ratio does not materially affect the ensemble computation
526 (Supplementary Materials).

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

531 **3.2.2 ENSEMBLE SUMMARY**

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

541 **3.2.3 A RANK**

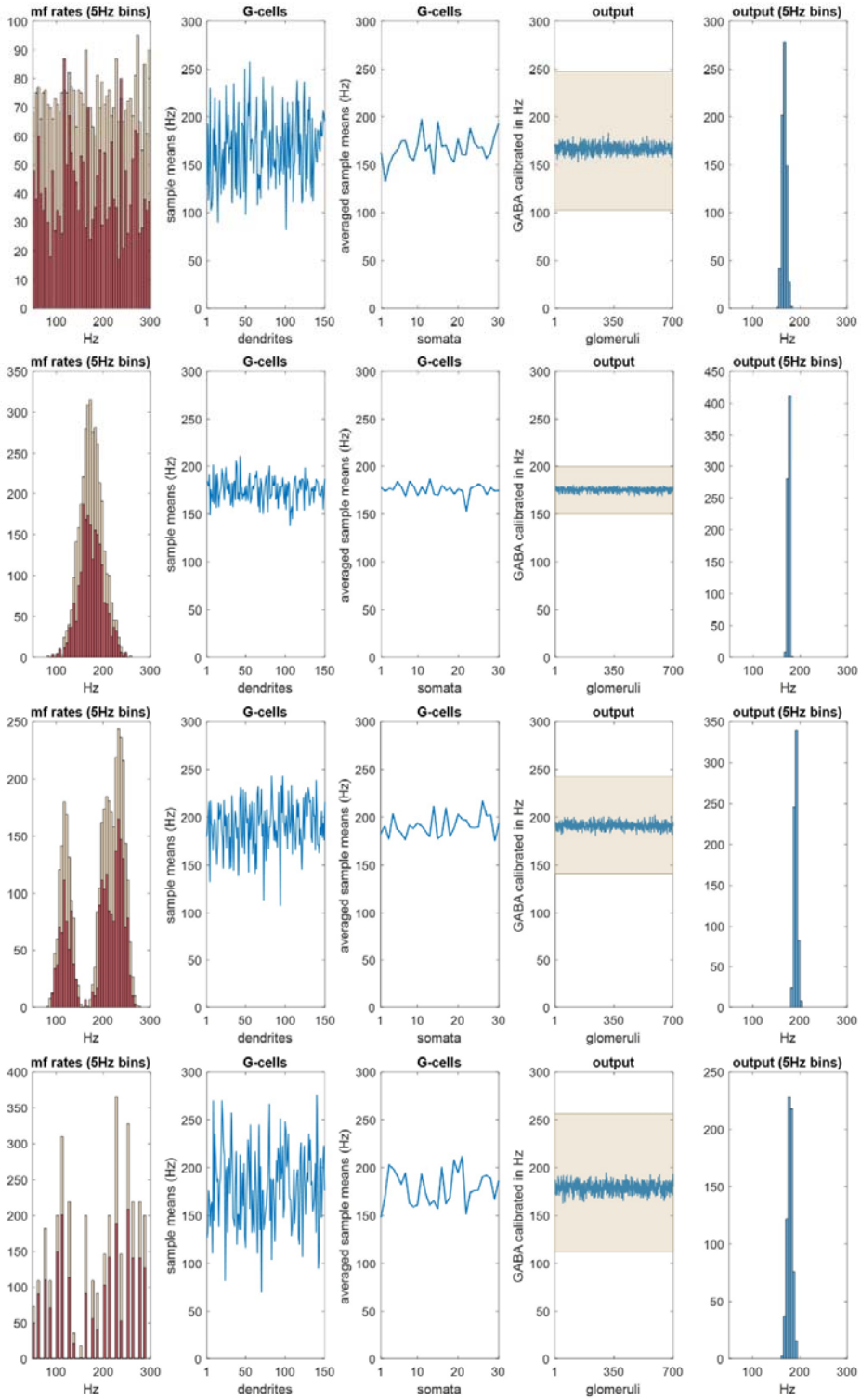
542 We next used the ensemble conversion to derive inhibition of granule cells in a sagittal
543 row of 100 fields, a rank. The results are shown in Figure 3. As ensembles all perform
544 the same computation and effectively randomly sample (with anatomically simulated
545 replacement) the same data (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, 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).

3.2.4 RANK SUMMARY

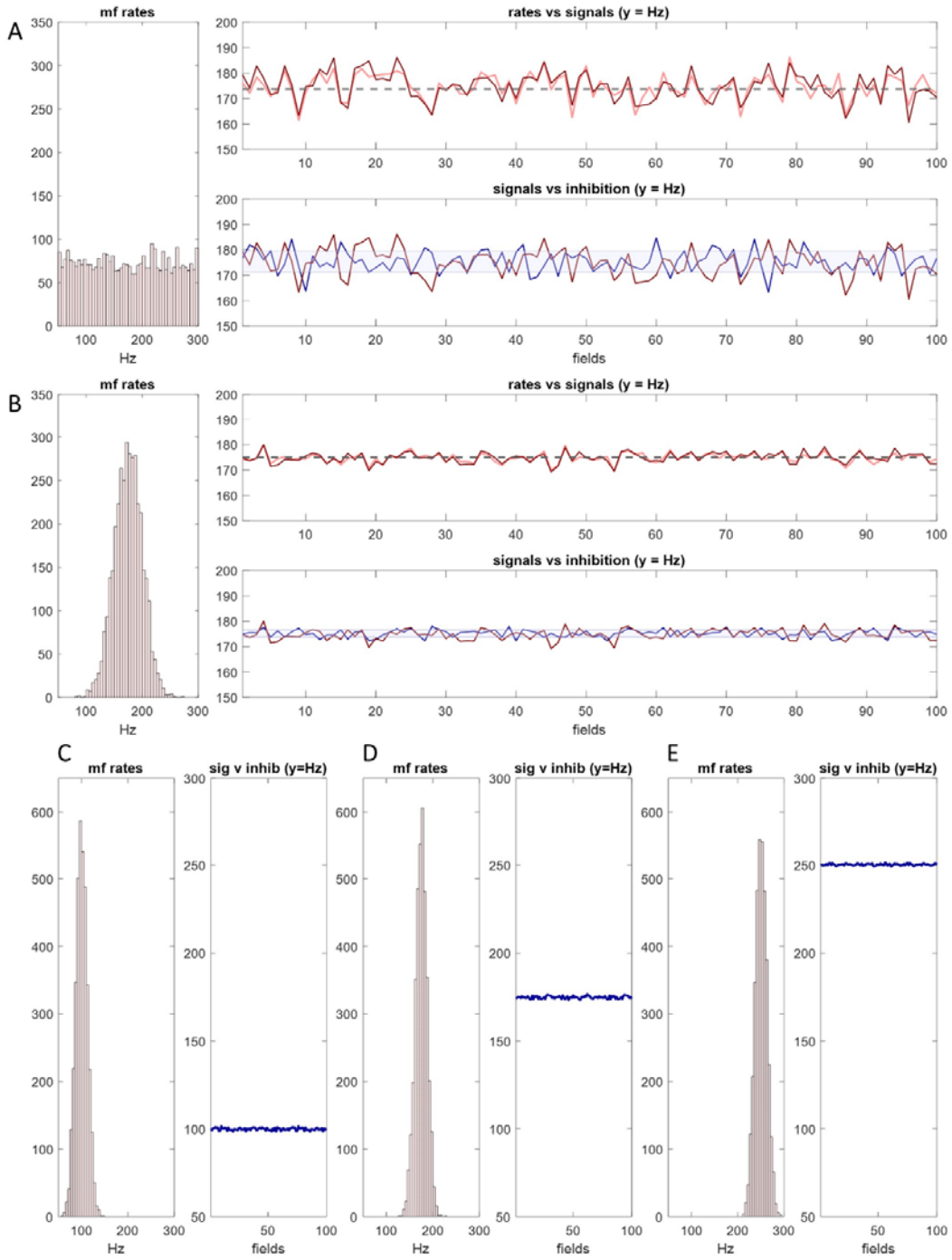
Golgi cells in a rank act as a single functional unit. Ensembles conflate in the sagittal direction functionally as they do anatomically. As a result, the whole granule cell population *of a rank* receive synchronised inhibition which is linearly related to, and varies with, the mean of mossy fibre rates.


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

565

566 **Fig 2. The Golgi cell ensemble conversion. Column 1** We randomly generated (3,645)
567 values representing mossy fibre rates with a uniform, normal, twin-peaked or
568 discontinuous distribution, simulating input to a row of 100 fields. Rate data were
569 randomly sampled to derive input rates to each field, and copies added to reflect the
570 fact that each cell afferent to a field contributes a randomly variable number of
571 terminals. See Methods for derivation of mossy fibre rates. Pale red: input rates to 100
572 fields without copies added. Dark red: input rates to a row of three fields with copies
573 added. **Column 2** The first-column dark red data were randomly sampled 150 times,
574 sample size 4, representing mossy fibre input to a 3-field population of 30 Golgi cells,
575 each with 5 basal dendrites (the mean) which each receive contact from 4 mossy
576 fibres. See section 6.1 for derivation of these numbers. Data are the sample means,
577 representing dendritic depolarisation. **Column 3** We next randomly sampled column 2
578 data *without* replacement 30 times, sample size 5, and took the sample means,
579 representing somatic integration of dendritic signals. **Column 4** Column 3 data were
580 randomly sampled 700 times, sample size 8–12, reflecting convergent input from 3
581 fields of Golgi cells onto 700 glomeruli in the middle field (which receives the output of
582 a three-field Golgi cell ensemble), and took the sample means, representing
583 intraglomerular GABA concentration, the result of spillover. **Column 5** The column 4
584 data are presented as a histogram. **All rows** Column 1 mossy fibre data are converted
585 to narrowly-focussed inhibition of granule cells. The mean of the column 4/5 data is
586 linearly related to the mean of the mossy fibre distribution. This result is independent
587 of the shape of the original distribution.

588


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

589

590

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

609

610

611

612

613

4. Regulation of granule cells

4.1 Hypothesis

Anatomy simulates independent random sampling by a single granule cell of the population of mossy fibre input rates to a rank (section 2). There is a competition for influence between mossy fibre excitation and Golgi cell inhibition of granule cells. The competition takes place in the glomerulus. Granule cells receive input to each dendrite from a single mossy fibre. Granule cells have 3–5 dendrites, averaging 4 (we assume 4, but the result is the same with 3–5). Each dendrite of a single granule cell extends into a different glomerulus. Estimates vary of the number of granule cells that receive contact from a single mossy fibre terminal. 50 is mid-range [31, 32] and consistent with our estimates of the number of terminals and granule cells in a field. Excitation and inhibition compete for influence in a glomerulus. It is unclear if there is always an emphatic winner, but a swing in the dominant direction may be amplified by positive feedback, acting through presynaptic receptors. Activation of presynaptic $GABA_B$ receptors on mossy fibres inhibits glutamate release [44], while activation of presynaptic mGluR2 receptors on Golgi cells inhibits GABA release [45]. To fire, a granule cell must receive at least 3 mossy fibre signals which exceed the inhibitory threshold. In experimental conditions, using stimulation that excites very high mossy fibre rates, granule cells can be induced to fire in response to only one input [46].

However, other accounts suggest one is too few [4, 47, 48], and three are needed [49, 50]. We assume 3. Only mossy fibre signals which prevail in the glomerular competition contribute significantly to granule cell depolarisation. During behaviour, the competition is ceaseless. The strength of inhibition is synchronised between glomeruli that populate a rank (section 3), so that all mossy fibre signals compete, at any moment, against the same inhibitory threshold. The outcome in each glomerulus is quasi-independent of the outcome at other terminals, including other terminals that innervate the same granule cell. It is not in fact independent because it is the same at all terminals that arise from the same mossy fibre. However, a single granule cell receives a quasi-independent random sample of outputs of the glomerular competition for the same reason it receives a quasi-independent random sample of mossy fibre rates (section 2). 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 (section 3), the probability is the same, at any time, for all granule cells. The probability that a given granule cell receives three competitive signals and therefore fires is a function of the percentage of mossy fibre signals that are competitive. Expressed as an equation, the expected number of granule cells that fire in a field is the result of

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

where T is the total number of granule cells per field, n is the number of dendrites per granule cell, and j is the minimum number of dendrites that must receive dominant excitation, $e > i$. The proportion of granule cells that fire is near the expected number by the law of large numbers. The percentage of mossy fibre signals that are competitive is always the same because inhibition scales with the mean of mossy fibre rates (because Golgi cells translate mossy fibre firing rates into proportional inhibition). As the percentage is unvarying, so is the probability. This relationship is always conserved regardless of the original distribution (the frequency distribution of mossy fibre rates received as input to the system). So, the probability does not vary either between granule cells or with time. The number of granule cells that fire is very near the expected proportion (the number predicted by probability) because granule cells are numerous. Since the probability is invariant, therefore so is the proportion of active granule cells.

Note: the idea of an all-or-nothing result of the intraglomerular 'balance' of excitation and inhibition has appeared previously [51].

4.2 Results

The model compares mossy fibre signal frequency with inhibition at each of 4 dendrites of each of 8,750 granule cells per field, in each of 100 fields. Mossy fibre input rates to each field are derived, and duplicate signals added, in the way described

671 earlier. Inhibition – outputs of the ensemble conversion, 700 per field – is provided by
672 the Golgi cell model (which takes the same mossy fibre distribution as input). Mossy
673 fibre signals are randomly sampled. Output of the model is the number of granule cells
674 that fire in each field and the mean of mossy fibre rates they each receive (the sample
675 means). The number that fire in each field is shown in Figure 4.

676 Feed-forward inhibition – mossy fibre → Golgi cell → granule cell – reliably and closely
677 regulates the fraction of granule cells that are active when the number is small.
678 Variability of the number is inversely proportional to the fraction that are active,
679 suggesting that, if primary regulation is provided by feed-forward inhibition (and we
680 are correct about the mechanism), the physiological number is low – termed sparse
681 code. We would nonetheless expect to see greater precision, suggesting the
682 physiological presence of a supplementary mechanism(s) that we have not
683 represented in the model.

684 A local feedback mechanism may be provided by granule cell ascending axon contact
685 on Golgi cells. The ascending portion of the granule cell axon makes synaptic contact
686 on local Golgi cells, on basal dendrites [10]. A Golgi cell receives ~400 functional
687 synaptic contacts from this source, around 25% of the granule cells that a Golgi cell
688 receives contact from in total. During sustained mossy fibre discharge, synapses from
689 local granule cells contribute equivalent charge to Golgi cells as mossy fibre synapses

690 and ‘display similarly fast kinetics’ [52]. Short-latency inhibitory feedback may add
691 homeostatic fine adjustment to feed-forward regulation at local level.

692 It is a long-standing idea [53] that granule cell activity may be regulated by feedback
693 via parallel fibres, which contact Golgi cell apical dendrites. We have previously
694 proposed that gap junctions between apical dendrites may provide the mechanism for
695 Golgi cells to detect the density of parallel fibre activity [54]. Historically, the feedback
696 pathway – mossy fibre → granule cell → Golgi cell → granule cell – has been regarded
697 as a more intuitive candidate for a regulatory role. It was at one time thought that
698 feedforward inhibition might be involved, instead, in control or adjustment of spike
699 timing. We do not incorporate the gap junction model for these reasons. 1) It has
700 appeared previously. 2) By our calculation, as presented here, the feedforward
701 pathway is the dominant regulator, suggesting the feedback pathway has a supporting,
702 fine-tuning role, so that omitting it has a modest effect on the results. We might
703 expect that feedforward inhibition is the dominant partner because it operates at local
704 (field) level. Strong local regulation is necessary to ensure that active granule cells are
705 uniformly distributed in the general population. The feedback pathway is not sensitive
706 to local fluctuation of granule cell active numbers. 3) Our scope is a single rank. Parallel
707 fibres have a range of around 3mm in each direction from the granule cell soma [29,
708 30], so that it would be necessary to model a larger region (about 40 times the size),
709 which extends into the molecular layer, to generate parallel fibre input to a strip of
710 Golgi cells. 4) It is a different (and independent, although related) mechanism. 5) They

711 have (related but) different functions. The feed-forward pathway is a local regulator of
712 the fraction of active granule cells (with an effect that flows through to parallel fibres),
713 whereas the feedback pathway is a regulator of the fraction of active parallel fibres
714 only. While the feedback mechanism works by modulating inhibition of granule cells,
715 modulation is very weakly related to local conditions.

716 Uniform density of active granule cells, and therefore strong local regulation, is
717 important for mossy fibre information to be faithfully coded by granule cells. The
718 granule cell code is the subject of the next section.

719 The results are independent of (i) the frequency distribution of mossy fibre rates; (ii)
720 whether mossy fibres are all active or some are active and some are not; and (iii) the
721 permutation they are active in (which ones are active and which are not).

722

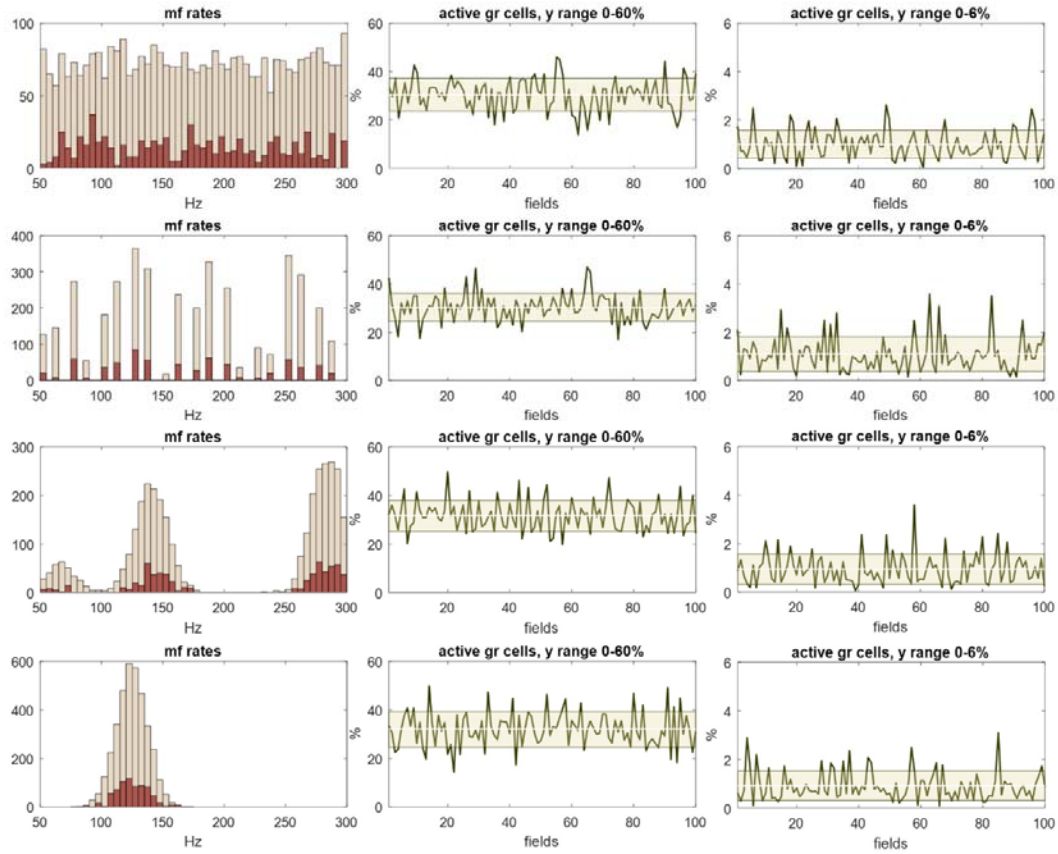


Fig 4. Regulation of granule cell activity by the feed-forward pathway (mossy fibres→Golgi cells→granule cells) We incorporated the Golgi cell ensemble conversion into a simulation of the conversion of mossy fibre signals into granule cell signals (recoding). One of the outputs of recoding is regulation 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, and whether there is an effect of the distribution 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. Regulation is poor: there is wide variation from field to field. **Column 3** The same as column 2 with an adjustment (a co-efficient) to sparsen granule cell activity. y axes have 1/10 the range of column 2. Regulation is strict: the percentage of active cells 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. The granule cell code

5.1 Hypothesis

Granule cells code information in their collective activity, in functionally defined groups. Granule cell code groups are functionally defined by the topography of mossy fibre input to the cerebellum, as Golgi cell networks are. The output of a rank is the smallest unit of granule cell encoded information. Granule cell encoded information, accordingly, has spatial dimensions, the dimensions of a code group. Information is coded in the frequency distribution of granule cell firing rates. Therefore, the same information is coded, at any moment, in firing of any random sample of co-active cells,

757 provided the sample is large enough. Because granule cells are very numerous, the
758 minimum dimensions of code are much smaller than a code group, because they are
759 the size of dimensions that are large enough to contain a sufficient number of co-
760 active cells at the regulated level of granule cell activity. The permutation of active
761 mossy fibres, and of granule cells, and how firing rates are distributed among active
762 cells, all code nothing.

763 Anatomy simulates independent random sampling by a single granule cell of the
764 population of mossy fibre input rates to a rank (section 2). The sampling unit is the
765 whole cell since it receives only a single input rate to each dendrite. Sample size is
766 therefore 4 (on average), and the number of samples per field is the number of
767 granule cells, which we estimate at 8,750 (see Methods). The unit function is only
768 operative in the case of a cell that receives at least the minimum number of
769 competitive mossy fibre signals. Then, somatic depolarisation is a linear function of the
770 mean (or sum) of competitive input rates (and is in turn straightforwardly related to
771 the granule cell firing rate – see section 6 for a discussion and references). We
772 postulate that the linear input-output relationship is undiluted by inhibition because
773 dominant signals displace a significant effect of inhibition. Derived in this way, there is
774 a reliable statistical relationship of granule cell rates and mossy fibre rates, again by
775 the central limit theorem. The range of granule cell rates is narrower than the range of
776 mossy fibre rates, centred on the mean, and attracted to a normal distribution. The
777 mean of the new distribution (unit outputs) is linearly related to the mean of mossy

fibre rates. This result is repeated in all fields. As all fields randomly sample, at any time, the same mossy fibre distribution, and perform the same computation, they all return the same granule cell code. This provides the substrate for granule cells to collectively encode mossy fibre rate information.

5.2 Results

We ran the same simulation as section 4 to test whether, and how, mossy fibre rates are related to granule cell rates, and if fields in a rank all have the same output. As before, the model compares mossy fibre rates with inhibition at each of 4 dendrites of each of 8,750 granule cells per field, in each of 100 fields. Mossy fibre signals received as input to each field are derived in the same way, described in section 2, and inhibition is provided by the Golgi cell model. The results are shown in Figures 5 and 6. We stated in section 4 that the outputs of the model are the number of granule cells that fire in each field and the mean of input rates they each receive – the sample means.

The narrowing and predicted shape of the granule cell frequency distribution is evident in all conditions. The 'field means' (for each field, the mean of the sample means) approach a straight line, the test that a rank operates as a single functional unit. When mossy fibre rates are normally distributed (SD 12.5 Hz), the granule cell distribution is very narrow (and normal if rates are counted in narrower bins) and the field means are

797 very close to a straight line. The field means follow the mean of mossy fibre rates at a
798 fixed distance. In all tested conditions, granule cell rates have a narrowed range, which
799 is the same in all fields, clustered around the mean, and attracted to a normal
800 distribution (more evident with narrowly and/or normally distributed mossy fibre
801 rates). This holds even though a single granule cell receives a tiny fraction of mossy
802 fibre rates, and each field receives only a small fraction of mossy fibre signals received
803 by a rank, and also regardless of the shape of the mossy fibre distribution.

804 What may we infer from that? The purpose of the demonstration is to show that a
805 network of linear units, in the form described, has the predicted output. We know (or
806 can derive) from reported anatomy the size of unit populations, population ratios, and
807 convergence and divergence ratios. Evidence is cited in context. The evidence for
808 linear units is discussed in section 6. Random sampling is a functional feature, a
809 proposal to account for anatomy.

810 Predictions confirmed: input rates to the system in the physiological range, with any
811 distribution and incorporating duplication as a result of terminal branching and
812 clustering, in anatomical detail including variation of single cells, is reliably converted
813 to a granule cell distribution with the predicted features. A key feature is that
814 inhibition of granule cells is synchronised both in a field and in a rank. A second key
815 feature is that mossy fibre rates are converted to a granule cell distribution with the

same shape and mean in all sectors of a rank, so that any random sample of granule cell rates also has that distribution.

The principal outstanding questions, as we see it, are whether:

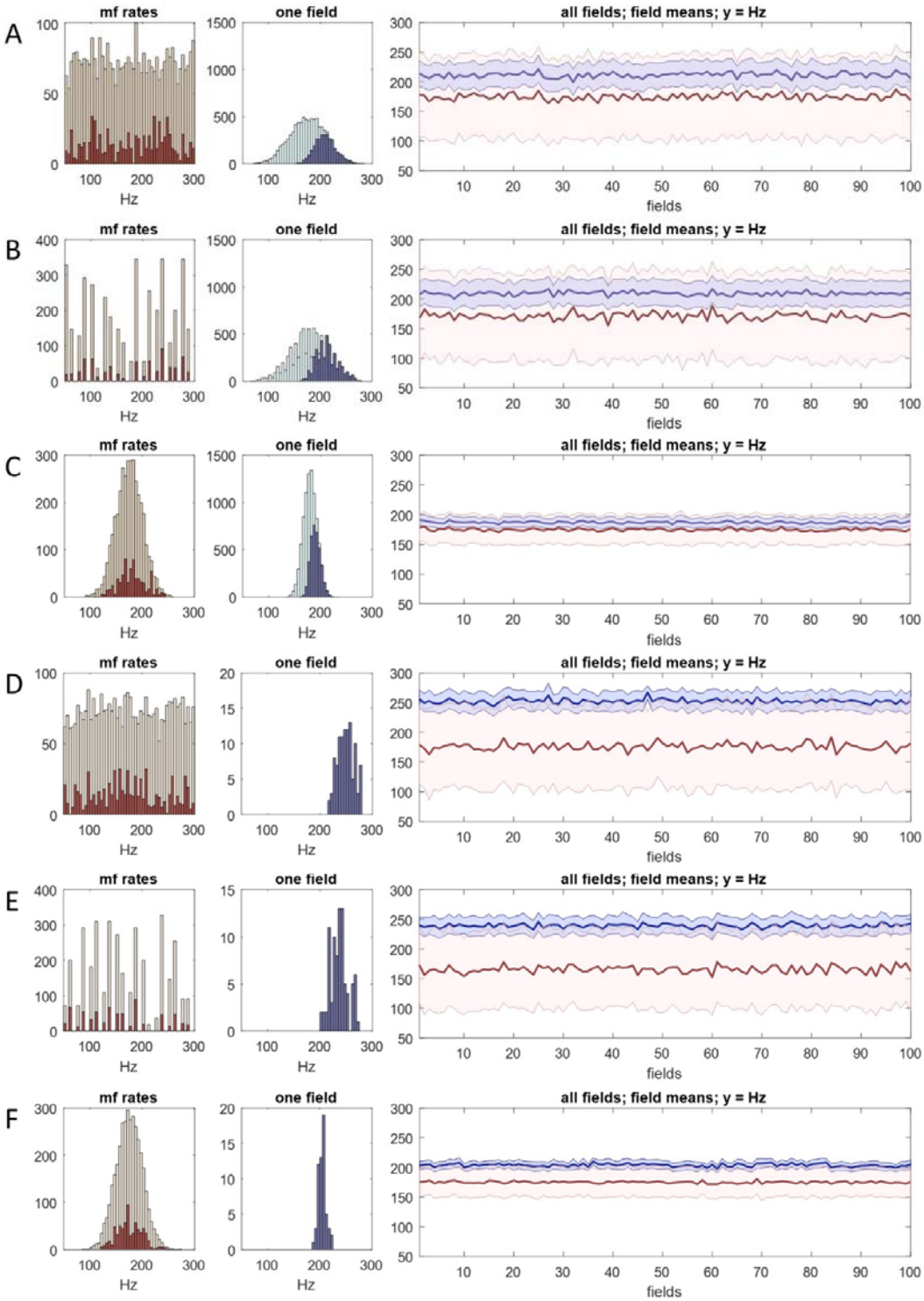
1) Computational units are linear (see next section).

2) Contact between cells in the granular layer is at random below a scale threshold set by range and morphology.

3) Sampling mimics replacement, an inference from anatomy.

4) The frequency distribution of granular layer firing rates is code.

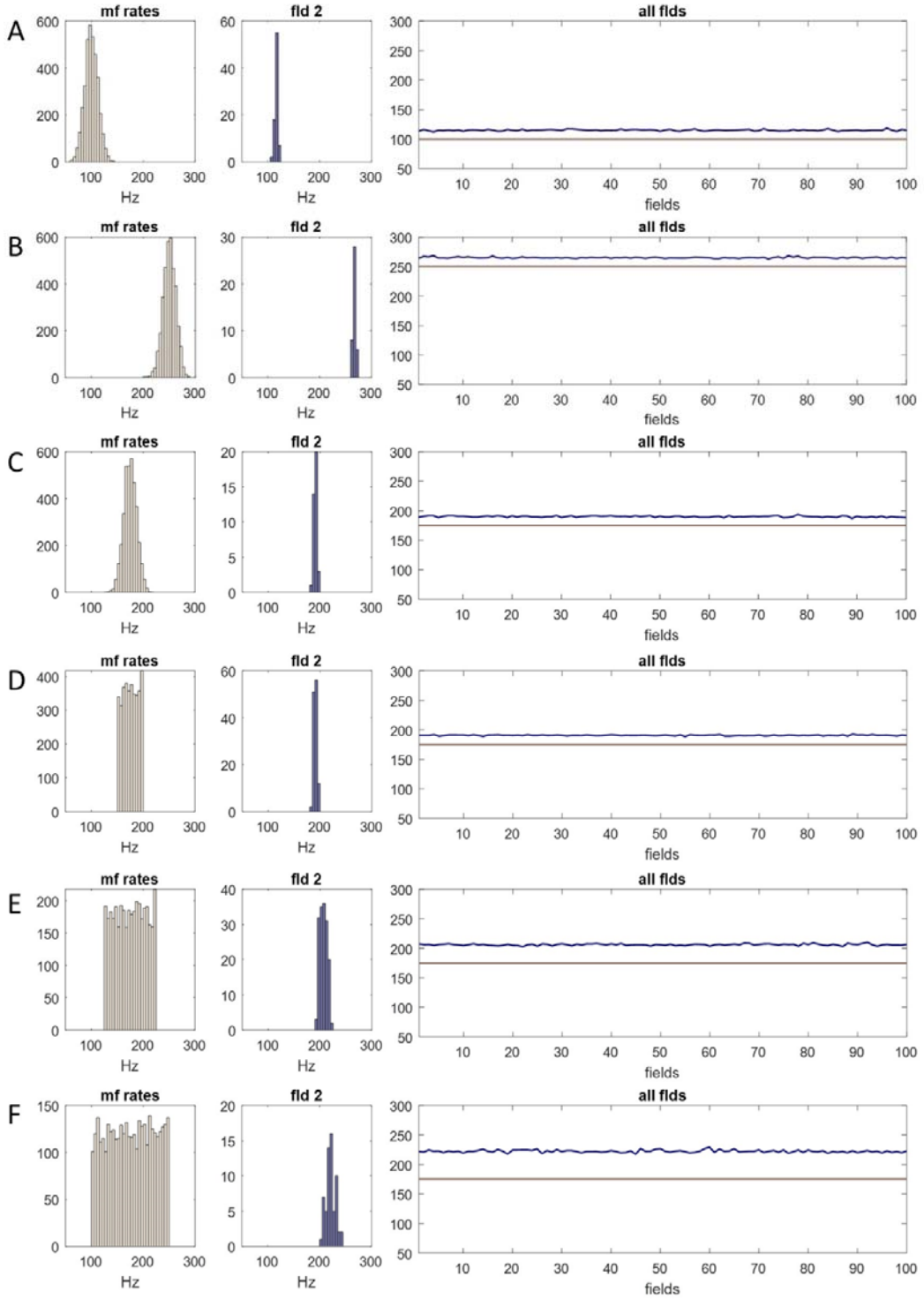
If so, we can go further. Granule cell firing conserves mossy fibre rate information, coded at the collective scale of the output of a rank. This is therefore the minimum resolution of granule cell information. However, it is not the minimum unit of code. The same information is coded in any random sample of granule cell signals, over a minimum number. It is immaterial (i.e. codes nothing) which particular cells are active or which cells fire at what rates, or even what rates those are.


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

831

832 **Fig 5. A rank functions as a single unit.** The same simulation as Figure 4 but data are
833 granule cell firing rates or derived from granule cell firing rates. **Column 1** Mossy fibre
834 rates were generated in the same way as earlier figures (the number of signals is the
835 same in all rows). **Column 2** Pale green: sample means – the mean of input rates to
836 each granule cell in a single field selected at random; blue: sample means of the subset
837 of granule cells which meet the conditions to fire. In rows A–C, ~30% fire. In rows D–F,
838 ~1% fire, y axes are rescaled and green data are omitted. **Column 3** Red: mean of
839 mossy fibre input rates to each field in a rank, including duplicate signals; pink: SD of
840 mossy fibre input rates to each field; blue: field means (i.e., the mean of granule cell
841 firing rates in each field); pale blue: SD of granule cell firing rates in each field.
842 Regardless of the shape and large range of the mossy fibre distribution, the field
843 means always converge towards a straight line which is a fixed distance from the
844 overall mean of mossy fibre rates. The amount and direction of field-to-field random
845 fluctuation of the blue data is independent of local fluctuation of the red data, and in a
846 smaller range. Uncertainty inherent in random sampling – first by single fields of input
847 rates to a rank, and then by Golgi cells and granule cells of input rates to a field – is not
848 compounded.

849



850

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 of the sampled distribution.

6. Support for linear transmission

There is a growing body of evidence of linear communication in the granular layer, indicating that unit functions are linear. This section, except the first subsection, is a short review.

6.1 Derivation of some Golgi cell ensemble parameters

4 Purkinje cells fit in a field, based on field size, thickness of the Purkinje cell dendritic plate – 9 µm [55] – and the distance between them – ~40 µm [56]. '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' [57]. Possibly, estimates are low because smaller Golgi cells at deeper level were not counted [56 p.121]. We take it that there are 10 Golgi cells per field, a ratio of 1:2.5 with 4 Purkinje cells.

Contact by mossy fibres on a Golgi cell is on basal dendrites. 4–6 basal dendrites arise from the Golgi cell soma. We assume 5, so a population of ensemble-grouped Golgi cells has a combined total of 30 x 5 = 150 basal dendrites. The number of mossy fibres which contact a Golgi cell is unreported. Dendrites radiate from the cell body. A single dendrite has a range of 50–100 µm [56]. The longest dimension of a field – which contains ~700 mossy fibre terminals – is 200 µm. We estimate the mean number of mossy fibres that contact a single dendrite is 4, so 20 per cell, or ~3% of the population of terminals (or 0.6% per dendrite).

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

6.2 Linear transmission of mossy fibres to Golgi cells

6.2.1 MAIN DISCUSSION

Golgi cells are large interneurons whose cell bodies and basal dendrites lie in the granular layer. Mossy fibres directly contact Golgi cell basal dendrites. Golgi cell firing frequency increases linearly with the amplitude of depolarising current [40 p.845]. 'Golgi cells can follow peripheral signals in a continuous fashion, modulating their

890 frequency with the intensity of the stimulus [citing [58, 59]]' [60]. As a result, during
891 movement, the 'output of Golgi cells varies smoothly as a function of input.…They are
892 capable of modulating their output to reflect their input with little noise or signal
893 distortion' [11]. 'Sensory-evoked Golgi-cell inhibition scales proportionally with the
894 level of mossy fiber excitatory synaptic input', such that inhibition reliably conveys
895 mossy fibre rate information [3]. On this evidence: mossy fibre rate information is
896 conserved proportionally in Golgi cell firing rates.

897 Golgi cells extend their basal dendrites into glomeruli [34, 61]. Contact on them by
898 mossy fibres is multi-synaptic [56], contributing to a reliable, rapid (submillisecond)
899 response. The fast rise time and weak distance dependence of the amplitude and
900 timing of Golgi cell EPSCs evoked by mossy fibre signals, suggests that basal dendrites
901 introduce relatively little filtering [10, 39], as expected for large-diameter dendrites.
902 We note that the pause in Golgi cell firing following a discrete stimulus under
903 anaesthesia [3, 62, 63] disappears from Golgi cell recordings during locomotion [59].

904 During behaviour in freely-moving animals, mossy fibre activity is dense: a high
905 fraction of mossy fibres are active [64]. Both mossy fibres [65-67] and Golgi cells [59]
906 have been reported to fire with a sustained, time-varying signature. In the behaving
907 animal, Golgi cell basal dendritic membrane potential is a continuous variable under
908 modulation by sustained inputs, we submit. As charge transfer to the soma is passive,
909 polarisation of the soma is likewise sustained, under modulation by dendritic states.

It is worth noting that very high mossy fibre rates can be generated in experimental conditions. However, the typical physiological range is 50–300 Hz [66]. In the simulations, we use the physiological rates, within stated constraints on the shape of the distribution.

6.2.2 IS THERE OTHER CONTROL OF GOLGI CELLS?

Do Golgi cells receive a significant influence of input from other sources? The evidence is incomplete but currently suggests other influence is absent or weak. Contrary to early reports, neither Purkinje cells nor climbing fibres contact Golgi cells [68, who give references], and only a modest minority of inhibitory inputs to Golgi cells (if any) are from molecular layer interneurons [69], which generate weak synaptic currents [70], consistent with either extremely weak or wholly absent innervation [71]. There is conflicting evidence whether Golgi cells make inhibitory synaptic contact on each other [71, 72]. Our simulation does not include an effect on Golgi cell firing by other sources of input.

6.2.3 GOLGI CELL OSCILLATIONS

Golgi cells fire autonomously. Under anaesthesia, firing falls into a slow oscillating pattern [73]. Under excitatory synaptic input, however, this disappears [72]. Anaesthesia 'has a strong influence on spontaneous activity of Golgi cells' [68]. Discharge at low rates with irregular timing under anaesthesia [41] is replaced by

929 higher rates with more regular timing without anaesthesia [58, 59, 74]. This paragraph
930 is a cautionary note about evidence obtained under anaesthesia, which has been used
931 to argue that oscillations provide a form of signalling.

932 6.3 Linear transmission of mossy fibres to granule cells

933 There is significant evidence that granule cell firing rates are linearly related to input
934 rates, but also evidence that has been given a conflicting interpretation. We take them
935 in turn.

936 The short and equal length of granule cell dendrites would suggest light and equal
937 filtering. The mossy fibre-granule cell connection has a range of adaptations which
938 support high-fidelity transmission of high frequency signals across a wide bandwidth
939 [8]. Faithful transmission of rate information is reported [1, 46]. Vesicle release and
940 replenishment are fast [9, 12] and postsynaptic AMPA receptors operate in their linear
941 range [12], where they are resistant to desensitisation [75]. Multiple contacts are
942 made by a mossy fibre with each granule cell [32]. Conversion of depolarising somatic
943 charge into granule cell firing is a simple function, such that granule cells 'have a
944 relatively linear and uncomplicated conversion of depolarisation level to spike rate' [48
945 p.2393, citing Jorntell and Ekerot 2006 and D'Angelo et al 1998].

946 Glutamate spillover enhances precision and reliability of transmission [12, 76]. During
947 behaviour, at sustained physiological mossy fibre rates, there may be a build-up of

948 intraglomerular glutamate, increasing the relative influence of spillover in a balance
949 with synaptic transmission, assisted by short-term synaptic depression. In these
950 conditions, spillover may dominate, as it dominates inhibitory transmission [4, 43].
951 Spillover from multiple release sites (200–400 at a single mossy fibre terminal) is
952 plausibly sufficient to mitigate variability of vesicular size, release probability, and the
953 number of release points at synapses made on any individual cell, increasing fidelity
954 and equality of excitation of granule cells.

955 Against that, there are reported to be heterogeneous mossy fibre-granule cell synaptic
956 weights. Different strength and short-term dynamics of the mossy fibre to granule cell
957 connection have been reported in vestibular slices, with $GABA_A$ receptors blocked. The
958 amplitude of the response, measured at the soma, is reported to be statistically
959 correlated to the source of the mossy fibre that is stimulated [77], suggesting that the
960 strength of the connection depends on the source.

961 However, the presence of weights does not necessarily mean that the function is to
962 modulate firing rates. Rather, the authors themselves suggest that the function may
963 temporally code the source of inputs because different combinations have different
964 response onset times. An alternative, which they also mention, is that some inputs
965 may be in a supporting role, reminiscent of 'driver' and 'modulatory' signals in the
966 thalamus and cortex [78]. Feasibly, signals in a supporting role must be present for the
967 postsynaptic granule cell to fire, but do not affect the granule cell firing rate.

968 Heterogenous granule cell subtypes have been inferred from recordings of granule
969 cells in rat slices, from the finding that when stimulation is extended beyond 500 ms
970 firing of most cells changes, either slowing down or (in fewer cases) speeding up.
971 However, the significance is not clear. Granule cells typically fire in short bursts –
972 shorter than the duration of the stimulation tested, so perhaps not long enough for
973 the effects observed. Short granule cell burst duration may be because a granule cell
974 normally does not meet the conditions to fire for long, because few fire at any time,
975 and the permutation of active and inactive inputs to a single granule cell, and the firing
976 rates of active inputs, are both in a state of restless change. Also, excitation of granule
977 cells receives a concurrent regulatory influence of a network of Golgi cells which
978 populate a region which receives 100s of time-varying mossy fibre signals at any time
979 (so, not the conditions tested). Confirmation of functional significance of adaptive
980 firing in experimental conditions is outstanding.

981 6.4 Linear transmission of Golgi cells to granule cells

982 The hypothesis includes the proposal that inhibition of granule cells is, at any time,
983 linearly related to the mean firing rate of Golgi cells afferent to each glomerulus. In
984 support, we cite evidence that GABA spillover into the intraglomerular space is
985 proportional to afferent Golgi cell rates and that inhibition of granule cells is almost
986 exclusively mediated by spillover, and therefore linearly reflects the intraglomerular

concentration of GABA. Citations appear in the Golgi cell section but for convenience they are repeated here.

Fine, beaded axon fibres enter glomeruli, where they inhibit granule cells [34]. 'In the adult rat cerebellum, each granule cell dendrite receives 2.6 ± 0.55 synaptic contacts from Golgi axon terminals' [36] citing [32].[3] However, the large majority of inhibition (98%) is by spillover [4], where neurotransmitter released into the synaptic cleft spills out into the glomerular space. Golgi cells release GABA, an inhibitory neurotransmitter. This is detected by high-affinity $GABA_A$ receptors located perisynaptically and extrasynaptically on granule cells [37, 38]. Even synaptically-received signals have most of their effect (i.e., the large majority of charge transfer is mediated) via spillover.[4]

Golgi cells fire [41] at a time-varying rate in the behaving animal, so that a glomerulus receives continuous input. As a result, there is a sustained build-up of glomerular GABA during behaviour [42] at an adjustable concentration controlled by Golgi cell firing rates [43]. Signalling by spillover is sometimes assumed to be ambient and slow. However, the action of glomerular GABA spillover has a fast phasic component – not as

[3] 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.

[4] 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. 36. 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.

fast as synaptic transmission (~1 ms) but with a rise time of only a few milliseconds [43]. Unlike the spiky appearance of synaptically-induced inhibitory postsynaptic currents, spillover [3] generates a sustained outward current.

7. Discussion

7.1 Summary

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

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.

2. Physiology is the physical form of a network of linear units organised in layers. Units randomly sample afferent activity. Code groups communicate by a facsimile of independent random sampling. Anatomy mimics random sampling with replacement by a single cell of the entire population of firing rates in the afferent cell layer.

1020 3. Data processing exploits statistical effects of random sampling, such that i) inhibition
1021 of granule cells in a rank is synchronised and tracks mossy fibre rates; and ii) the
1022 frequency distribution of granule cell rates is reliably (because statistically) well-
1023 predicted by mossy fibre rates.

1024 4. High-dimensional mossy fibre information recodes as a very low-resolution granule
1025 cell code. Large granule cell groups function as a single unit, coding information in their
1026 collective activity.

1027 5. At any instant, a whole group codes the same information. Accordingly, granule cell-
1028 coded information has spatial dimensions, the dimensions of code groups. As code
1029 groups are long and thin, information has a striped pattern viewed from the cerebellar
1030 surface.

1031 6. Code groups span the cerebellar network in the sagittal direction, from side to side,
1032 mirroring microzonal organisation of the molecular layer. Like microzones, they are
1033 defined topographically, by their input from outside the cerebellum, rather than
1034 having physical boundaries (but they have fuzzier borders).

1035 7. While firing of a whole code group is the minimum unit of information, it is not the
1036 minimum unit of code. The same information is coded in any random sample of code-
1037 grouped granule cell firing rates. In that sense, the code is homogenised below a scale
1038 threshold.

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

7.2 Commentary

7.2.1 GENERAL COMMENTARY

Speaking generally, there is a conflict between evidence relating to learning and evidence of linear transmission. There is a conflict partly because there is an implicit inference that i) synaptic changes are functional, and ii) synaptic weights are graded. Historically, there has been strong emphasis on researching learning [79, 80]. The traditional learning model, which centrally features synaptic memory, has not received confirmation by experimental work. However, there is still a mainstream feeling that the computation will turn out to involve synaptic weights. In this view, core assumptions are retained, and the aim of further work is to understand how plasticity at multiple sites is related, and related to function.

Such has been the dominance of learning, it is regarded as naive to have a different opinion. However, it does not address a range of evidence relating to linear transmission (discussed in context, in the main text, and reviewed in section 6), and of linear relationships between firing rates and behavioural metrics, suggesting linearity is not merely an experimental artefact (for which Raymond and Medina provide 30 references [15]). We present an alternative approach where the focus is to explain cell

1058 morphologies/general anatomy and linear transmission. We propose there is a
1059 computational effect of anatomy, give linear signalling. Our focus is the granular layer
1060 of the cerebellar cortex. We deal with the linear/non-linear question by dedicating a
1061 section (section 6, as noted) to a review of the evidence for linear transmission, which
1062 is rich for the granular layer.

1063 We know of course that this is contentious. However, it has merits.
1064 1. It has in fact good explanatory value – cell morphologies, granular layer network
1065 geometry, cell ratios, spatial organisation, the shape and size of dendritic and axonal
1066 territories, convergence and divergence ratios, random contact that is on some and
1067 not other cells that lie in axonal range, the duplication of mossy fibre signals as a result
1068 of terminal branching and also locally because branches each end in a cluster of
1069 terminals, the structure and physiology of the glomerulus, and linear signalling
1070 (including evidence that synaptic connections are adapted for high-fidelity
1071 transmission) – are all accounted for.

1072 2. We propose – as artificial neural network models do – that signals traffic flows
1073 through computational units organised in layers, but with the difference that

1074 i. Our units are matched with physiological hardware;
1075 ii. Units form functional populations whose size and spatial dimensions are an
1076 effect of known physiology; and

1077 iii. Unit functions are not conjecture, they are linear, with the basis in evidence of
1078 linear signalling. We note in this connection that non-linear synaptic properties
1079 predicted by experimentally observed synaptic changes have the same problem
1080 with evidence of linearity, but reversed, as linear unit functions do with
1081 synaptic evidence of non-linearity.

1082 3. A linear network, as we describe, has functional advantages. We discuss the
1083 functional significance of the proposed form of the code, and of the computation, in
1084 the remainder of the Discussion. We make three main arguments. In all three, the
1085 form of the computation solves a problem. The first problem is how an anatomically
1086 seamless tangled mat of cells is organised functionally with a view to convert high-
1087 dimensional external signals into coherent firing of microzone-grouped Purkinje cells.
1088 The second is that – against expectation – the pattern in which parallel fibres are active
1089 may not in fact be reliably or perhaps ever replicated. Theory has previously addressed
1090 this problem by making assumptions that have been widely adopted but which are not
1091 safe. The first paragraph of the paper comments on this. The third problem is how
1092 insentient cerebellar circuits are able to process input data usefully with, as far as we
1093 know, little or no information about where it is from or what it represents.

1094 **7.2.2 CONTROL OF MICROZONE-GROUPED PURKINJE CELLS**

1095 There is a huge sacrifice of data resolution on entry to the cerebellum, counter to the
1096 general expectation that data processing is sophisticated. This is functional. Coded in

1097 this form, granule cell data can drive coherent firing of large cell groups downstream.
1098 Microzones lie at right angles to parallel fibres and parallel to code groups. The
1099 statistical form of the granule cell code means the same information is received by all
1100 locations of sagittal plane, so that microzone-grouped Purkinje cells can receive the
1101 same information without receiving any of the same signals. An estimated 350,000
1102 parallel fibres pass through a Purkinje cell dendritic territory [28, 81-83]. If even only a
1103 small fraction of parallel fibres is active (say, 0.5%), and allowing both for the fact that
1104 only one in two make contact and that a majority of synapses are severely depressed
1105 or silent [84, 85], a single Purkinje cell still receives around 200 parallel fibre signals at
1106 working synapses at any time. A granule cell code in this form provides a means to
1107 modulate firing of microzone-grouped Purkinje cells in unison, ultimately under control
1108 of mossy fibre rates. There is evidence that Purkinje cells linearly code the strength of
1109 granule cell inputs [5-7] and linearly code behavioural metrics [2, 86-89].

1110 **7.2.3 REPLICABILITY**

1111 A statistical code may be a physiological strategy to mitigate variable performance of
1112 single neurons and synapses, long known to be a practical problem for replicable
1113 biological information capture. The traditional cerebellar learning model [53, 90, 91]
1114 relies on (and assumes) replicability because pattern replication is necessary for signals
1115 to match up with the correct training-modified parallel fibre synaptic weights. Learned

changes of synaptic weights are also a core feature of many neural network learning models, not just cerebellar models.

However, this may not be safe. First, it is not clear that input to the cerebellum is ever replicated. Around 4,000 mossy fibres innervate the region that supplies parallel fibre input to a single Purkinje cell. It may not be a safe assumption that conditions upstream are routinely repeated with precision sufficient to consistently (or perhaps ever) generate a repeat of the same inputs. Second, single neurons and synapses are 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* so that small differences of input are converted by decorrelation to wholesale changes to output (the pattern of active parallel fibres). 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.

Homogenised granule cell/parallel fibre code removes the problem because replication of the pattern of active granule cells is unnecessary to precisely code mossy fibre information. This is because:

1. It is coded in granule cell firing rates and not the pattern they are active in.

2. It is coded in population activity, in rate statistics.

In the traditional model, a granule cell rate code was effectively redundant. That is not to say granule cell rates had no effect on Purkinje cell firing. But the function of synaptic memory was to *displace* control by rates. They might start out as *any* rates, because learning would generate the 'desired' firing of Purkinje cells following training. In the alternative model, output of the system is an unlearned response to input rates and the pattern of active cells is irrelevant. The colossal sacrifice of resolution is a functional feature.

7.2.4 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.

We propose instead that the solution is to sacrifice the resolution of high-dimensional input data. Source and modality data are disregarded below a topographically-defined scale threshold. It is therefore unnecessary to discriminate below that scale. Mapping studies show that mossy fibre signals (evoked by cutaneous stimulation) share the topographical organisation of climbing fibres [19, 92, 93]. At sub-topographical dimensions, we submit, signal source and type are simply discounted. Therefore,

1155 networks do not need to be adapted to the source or type of input they receive –
1156 anatomically, networks are interchangeable, so network design can be modular. This
1157 feature allowed basic cerebellar wiring to be highly conserved through dramatic
1158 upheavals of the vertebrate phenotype [94, 95].

1159

Glossary

field A nominal division of the granular layer measuring 200 µm (sagittal) x 150 µm (mediolateral), the average dimensions of the area enclosing a mossy fibre terminal cluster.

field mean For each field, the average of the mean of mossy fibre input rates to each granule cell which meets the conditions to fire.

rank A sagittal row of 100 fields.

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. Ritzau-Jost, A., et al., *Ultrafast action potentials mediate kilohertz signaling at a central synapse.* Neuron, 2014. **84**(1): p. 152-163.
2. Jelitai, M., et al., *Dendritic excitation-inhibition balance shapes cerebellar output during motor behaviour.* Nat Commun, 2016. **7**: p. 13722.
3. 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.
4. Duguid, I., et al., *Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex.* J Neurosci, 2012. **32**(32): p. 11132-43.
5. Walter, J.T. and K. Khodakhah, *The linear computational algorithm of cerebellar Purkinje cells.* J Neurosci, 2006. **26**(50): p. 12861-72.
6. 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.
7. Park, S.M., E. Tara, and K. Khodakhah, *Efficient generation of reciprocal signals by inhibition.* J Neurophysiol, 2012. **107**(9): p. 2453-62.
8. Delvendahl, I. and S. Hallermann, *The Cerebellar Mossy Fiber Synapse as a Model for High-Frequency Transmission in the Mammalian CNS.* Trends Neurosci, 2016. **39**(11): p. 722-737.
9. Saviane, C. and R.A. Silver, *Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse.* Nature, 2006. **439**(7079): p. 983-7.
10. Cesana, E., et al. *Novel granule cell-Golgi cell excitatory input in the cerebellar granular layer*. in *FENS Abstr*. 2010.
11. 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.
12. 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.
13. 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.
14. Turecek, J., S.L. Jackman, and W.G. Regehr, *Synaptotagmin 7 confers frequency invariance onto specialized depressing synapses.* Nature, 2017. **551**: p. 503.
15. Raymond, J.L. and J.F. Medina, *Computational Principles of Supervised Learning in the Cerebellum.* Annu Rev Neurosci, 2018. **41**: p. 233-253.
16. Alvina, K., et al., *Questioning the role of rebound firing in the cerebellum.* Nat Neurosci, 2008. **11**(11): p. 1256-8.
17. 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.
18. Oscarsson, O., *Functional units of the cerebellum - sagittal zones and microzones.* Trends in Neurosciences, 1979. **2**: p. 143-145.

1224 19. Apps, R. and R. Hawkes, *Cerebellar cortical organization: a one-map hypothesis.* Nat
1225 Rev Neurosci, 2009. **10**(9): p. 670-81.
1226 20. Wu, H.S., I. Sugihara, and Y. Shinoda, *Projection patterns of single mossy fibers*
1227 *originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei.* J
1228 Comp Neurol, 1999. **411**(1): p. 97-118.
1229 21. Sultan, F. and D. Heck, *Detection of sequences in the cerebellar cortex: numerical*
1230 *estimate of the possible number of tidal-wave inducing sequences represented.* J
1231 Physiol Paris, 2003. **97**(4-6): p. 591-600.
1232 22. Shinoda, Y. and I. Sugihara, *Axonal Trajectories of Single Climbing and Mossy Fiber*
1233 *Neurons in the Cerebellar Cortex and Nucleus*, in *Handbook of the Cerebellum and*
1234 *Cerebellar Disorders*, S.J.D. Manto M., Rossi F., Gruol D.L., Koibuchi N., Editor. 2013,
1235 Springer, Dordrecht.
1236 23. Shinoda, Y., et al., *The entire trajectory of single climbing and mossy fibers in the*
1237 *cerebellar nuclei and cortex.* Prog Brain Res, 2000. **124**: p. 173-86.
1238 24. Garwicz, M., C.F. Ekerot, and H. Jorntell, *Organizational principles of cerebellar*
1239 *neuronal circuitry.* News in Physiological Sciences, 1998. **13**: p. 28-34.
1240 25. Ozden, I., et al., *Reliable coding emerges from coactivation of climbing fibers in*
1241 *microbands of cerebellar Purkinje neurons.* J Neurosci, 2009. **29**(34): p. 10463-73.
1242 26. Dean, P., et al., *The cerebellar microcircuit as an adaptive filter: experimental and*
1243 *computational evidence.* Nat Rev Neurosci, 2010. **11**(1): p. 30-43.
1244 27. Eccles, J.C., M. Ito, and J. Szentágothai, *The cerebellum as a neuronal machine*. 1967,
1245 Berlin, New York etc.: Springer-Verlag. 335 p.
1246 28. Harvey, R.J. and R.M. Napper, *Quantitative study of granule and Purkinje cells in the*
1247 *cerebellar cortex of the rat.* J Comp Neurol, 1988. **274**(2): p. 151-7.
1248 29. Mugnaini, E., *The length of cerebellar parallel fibers in chicken and rhesus monkey.* J
1249 Comp Neurol, 1983. **220**(1): p. 7-15.
1250 30. Brand, S., A.L. Dahl, and E. Mugnaini, *The length of parallel fibers in the cat cerebellar*
1251 *cortex. An experimental light and electron microscopic study.* Exp Brain Res, 1976.
1252 **26**(1): p. 39-58.
1253 31. Gao, Z., et al., *Excitatory Cerebellar Nucleocortical Circuit Provides Internal*
1254 *Amplification during Associative Conditioning.* Neuron, 2016. **89**(3): p. 645-57.
1255 32. Jakab, R.L. and J. Hámori, *Quantitative morphology and synaptology of cerebellar*
1256 *glomeruli in the rat.* Anat Embryol (Berl), 1988. **179**(1): p. 81-8.
1257 33. Barmack, N.H. and V. Yakhnitsa, *Functions of interneurons in mouse cerebellum.* J
1258 Neurosci, 2008. **28**(5): p. 1140-52.
1259 34. Hámori, J. and J. Szentágothai, *Participation of Golgi neuron processes in the cerebellar*
1260 *glomeruli: an electron microscope study.* Exp Brain Res, 1966. **2**(1): p. 35-48.
1261 35. D'Angelo, E., et al., *The cerebellar Golgi cell and spatiotemporal organization of*
1262 *granular layer activity.* Front Neural Circuits, 2013. **7**: p. 93.
1263 36. Rossi, D.J. and M. Hamann, *Spillover-mediated transmission at inhibitory synapses*
1264 *promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry.*
1265 Neuron, 1998. **20**(4): p. 783-95.
1266 37. Nusser, Z., W. Sieghart, and P. Somogyi, *Segregation of different GABAA receptors to*
1267 *synaptic and extrasynaptic membranes of cerebellar granule cells.* J Neurosci, 1998.
1268 **18**(5): p. 1693-703.

38. Brickley, S.G., S.G. Cull-Candy, and M. Farrant, *Development of a tonic form of synaptic inhibition in rat cerebellar granule cells resulting from persistent activation of GABAA receptors.* J Physiol, 1996. **497 (Pt 3)**: p. 753-9.
39. Kanichay, R.T. and R.A. Silver, *Synaptic and cellular properties of the feedforward inhibitory circuit within the input layer of the cerebellar cortex.* J Neurosci, 2008. **28**(36): p. 8955-67.
40. Dieudonné, S., *Submillisecond kinetics and low efficacy of parallel fibre-Golgi cell synaptic currents in the rat cerebellum.* J Physiol, 1998. **510 (Pt 3)**: p. 845-66.
41. Ruigrok, T.J., R.A. Hensbroek, and J.I. Simpson, *Spontaneous activity signatures of morphologically identified interneurons in the vestibulocerebellum.* J Neurosci, 2011. **31**(2): p. 712-24.
42. Rossi, D.J., M. Hamann, and D. Attwell, *Multiple modes of GABAergic inhibition of rat cerebellar granule cells.* J Physiol, 2003. **548**(Pt 1): p. 97-110.
43. Mapelli, L., S. Solinas, and E. D'Angelo, *Integration and regulation of glomerular inhibition in the cerebellar granular layer circuit.* Front Cell Neurosci, 2014. **8**: p. 55.
44. 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.
45. Mitchell, S.J. and R.A. Silver, *Glutamate spillover suppresses inhibition by activating presynaptic mGluRs.* Nature, 2000b. **404**(6777): p. 498-502.
46. Rancz, E.A., et al., *High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons.* Nature, 2007. **450**(7173): p. 1245-8.
47. 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.
48. 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.
49. Billings, G., et al., *Network structure within the cerebellar input layer enables lossless sparse encoding.* Neuron, 2014. **83**(4): p. 960-74.
50. 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.
51. Gilbert, M. and R.C. Miall, *How and Why the Cerebellum Recodes Input Signals: An Alternative to Machine Learning.* Neuroscientist, 2021a. **DOI: 10.1177/1073858420986795**.
52. 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.
53. Albus, J.S., *A Theory of Cerebellar Function.* Mathematical Biosciences, 1971. **10**(1-2): p. 25-61.
54. Gilbert, M. and A. Rasmussen, *Gap Junctions May Have A Computational Function In The Cerebellum: A Hypothesis.* The Cerebellum, 2024.
55. Hámori, J., *Anatomy and neurochemical anatomy of the cerebellum*, in *Cerebellar degenerations : clinical neurobiology*, A. Plaitakis, Editor. 1992, Kluwer Academic.
56. Palay, S.L. and V. Chan-Palay, *Cerebellar cortex : cytology and organization*. 1974, Berlin: Springer.

1313 57. Lange, W., *Regional differences in the distribution of golgi cells in the cerebellar cortex*
1314 *of man and some other mammals.* Cell Tissue Res, 1974. **153**(2): p. 219-26.
1315 58. Miles, F.A., et al., *Long-term adaptive changes in primate vestibuloocular reflex. III.*
1316 *Electrophysiological observations in flocculus of normal monkeys.* J Neurophysiol, 1980.
1317 **43**(5): p. 1437-76.
1318 59. Edgley, S.A. and M. Lidierth, *The discharges of cerebellar Golgi cells during locomotion*
1319 *in the cat.* J Physiol, 1987. **392**: p. 315-32.
1320 60. Galliano, E., P. Mazzarello, and E. D'Angelo, *Discovery and rediscoveries of Golgi cells.* J
1321 Physiol, 2010. **588**(Pt 19): p. 3639-55.
1322 61. Eccles, J.C., R. Llinás, and K. Sasaki, *Golgi cell inhibition in the cerebellar cortex.* Nature,
1323 1964. **204**: p. 1265-6.
1324 62. Holtzman, T., et al., *Different responses of rat cerebellar Purkinje cells and Golgi cells*
1325 *evoked by widespread convergent sensory inputs.* J Physiol, 2006. **574**(Pt 2): p. 491-
1326 507.
1327 63. Vos, B.P., A. Volny-Luraghi, and E. De Schutter, *Cerebellar Golgi cells in the rat:*
1328 *receptive fields and timing of responses to facial stimulation.* Eur J Neurosci, 1999.
1329 **11**(8): p. 2621-34.
1330 64. Hartmann, M.J. and J.M. Bower, *Tactile responses in the granule cell layer of cerebellar*
1331 *folium crus IIa of freely behaving rats.* J Neurosci, 2001. **21**(10): p. 3549-63.
1332 65. Arenz, A., et al., *The contribution of single synapses to sensory representation in vivo.*
1333 Science, 2008. **321**(5891): p. 977-80.
1334 66. van Kan, P.L., A.R. Gibson, and J.C. Houk, *Movement-related inputs to intermediate*
1335 *cerebellum of the monkey.* J Neurophysiol, 1993. **69**(1): p. 74-94.
1336 67. Giovannucci, A., et al., *Cerebellar granule cells acquire a widespread predictive*
1337 *feedback signal during motor learning.* Nat Neurosci, 2017.
1338 68. Pietrajtis, K. and S. Dieudonné, *Golgi Neurons*, in *Handbook of the Cerebellum and*
1339 *Cerebellar Disorders*, M. Manto, et al., Editors. 2013, Springer Netherlands: Dordrecht.
1340 p. 829-852.
1341 69. Eyre, M.D. and Z. Nusser, *Only a Minority of the Inhibitory Inputs to Cerebellar Golgi*
1342 *Cells Originates from Local GABAergic Cells.* eNeuro, 2016. **3**(2).
1343 70. Dumoulin, A., A. Triller, and S. Dieudonné, *IPSC kinetics at identified GABAergic and*
1344 *mixed GABAergic and glycinergic synapses onto cerebellar Golgi cells.* J Neurosci, 2001.
1345 **21**(16): p. 6045-57.
1346 71. Hull, C. and W.G. Regehr, *Identification of an inhibitory circuit that regulates cerebellar*
1347 *Golgi cell activity.* Neuron, 2012. **73**(1): p. 149-58.
1348 72. Vervaeke, K., et al., *Rapid desynchronization of an electrically coupled interneuron*
1349 *network with sparse excitatory synaptic input.* Neuron, 2010. **67**(3): p. 435-51.
1350 73. Ros, H., et al., *Neocortical networks entrain neuronal circuits in cerebellar cortex.* J
1351 Neurosci, 2009. **29**(33): p. 10309-20.
1352 74. Rasmussen, A., et al., *Golgi cell activity during eyeblink conditioning in decerebrate*
1353 *ferrets.* Cerebellum, 2014. **13**(1): p. 42-5.
1354 75. DiGregorio, D.A., et al., *Desensitization properties of AMPA receptors at the cerebellar*
1355 *mossy fiber granule cell synapse.* J Neurosci, 2007. **27**(31): p. 8344-57.

76. DiGregorio, D.A., Z. Nusser, and R.A. Silver, *Spillover of glutamate onto synaptic AMPA receptors enhances fast transmission at a cerebellar synapse.* Neuron, 2002. **35**(3): p. 521-33.
77. Chabrol, F.P., et al., *Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons.* Nat Neurosci, 2015. **18**(5): p. 718-27.
78. Sherman, S.M. and R.W. Guillery, *Distinct functions for direct and transthalamic corticocortical connections.* J Neurophysiol, 2011. **106**(3): p. 1068-77.
79. Hansel, C., D.J. Linden, and E. D'Angelo, *Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum.* Nat Neurosci, 2001. **4**(5): p. 467-75.
80. Gao, Z., B.J. van Beugen, and C.I. De Zeeuw, *Distributed synergistic plasticity and cerebellar learning.* Nature Reviews Neuroscience, 2012. **13**(9): p. 619-635.
81. Harvey, R.J. and R.M. Napper, *Quantitative studies on the mammalian cerebellum.* Prog Neurobiol, 1991. **36**(6): p. 437-63.
82. 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.
83. 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.
84. 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.
85. 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.
86. Dugué, G.P., et al., *Cerebellar re-encoding of self-generated head movements.* Elife, 2017. **6**.
87. Chen, S., G.J. Augustine, and P. Chadderton, *The cerebellum linearly encodes whisker position during voluntary movement.* eLife, 2016. **5**: p. e10509-e10509.
88. Herzfeld, D.J., et al., *Encoding of action by the Purkinje cells of the cerebellum.* Nature, 2015. **526**(7573): p. 439-42.
89. Hong, S., et al., *Multiplexed coding by cerebellar Purkinje neurons.* Elife, 2016. **5**.
90. Fujita, M., *Adaptive filter model of the cerebellum.* Biol Cybern, 1982. **45**(3): p. 195-206.
91. Marr, D., *A theory of cerebellar cortex.* J Physiol, 1969. **202**(2): p. 437-70.
92. Apps, R. and M. Garwicz, *Anatomical and physiological foundations of cerebellar information processing.* Nat Rev Neurosci, 2005. **6**(4): p. 297-311.
93. Ji, Z. and R. Hawkes, *Topography of purkinje cell compartments and mossy fiber terminal fields in lobules ii and iii of the rat cerebellar cortex: Spinocerebellar and cuneocerebellar projections.* Neuroscience, 1994. **61**(4): p. 935-954.
94. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from Myxinoids through Birds*. 1967: The University of Minnesota Press.
95. Larsell, O., *The Comparative Anatomy and Histology of the Cerebellum from Monotremes through Apes*. 1970: The University of Minnesota Press, Minneapolis.