WHAT IF…? COMPUTATIONAL ANATOMY IN THE GRANULAR LAYER OF THE CEREBELLAR CORTEX

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

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

[b]Department of Experimental Medical Science, Lund University, BMC F10, 22184 Lund, Sweden

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

Summary

Neural networks are believed to carry out sophisticated computations. In the influential traditional cerebellar learning model, and many artificial neural networks, computations are implemented by learned synaptic changes that selectively adjust synaptic transmission strength. These ideas have become widely influential. We propose an alternative in the form of a reinterpretation of the evidence relating to the granular layer of the cerebellar cortex, the entry point of the cerebellum. The main proposals are that there is a computational effect, unaided, of the anatomy and spatial organisation of granular layer cells and wiring, with the result that high resolution input data are converted on entry into a (very) low resolution internal code. Internally-coded information has spatial dimensions which mirror the functional organisation of the cerebellar cortex into long thin strips called microzones. Data are coded in the collective activity of granule cells at this scale. There is no finer internal resolution. This feels counter-intuitive because it is counter to the general expectation that high dimensionality increases the power of brain computations. However, it means the cerebellum is able to handle all input the same way, and it provides an otherwise elusive mechanism to orchestrate firing of large cell groups – microzones – downstream, which powerfully modulate firing of the output cells of the cerebellar network. There is no algorithm and the response is unlearned. The model is not abstract and has good evidence.

Keywords: cerebellum, theory, model, computational anatomy, cerebellar cortex, granular layer, network, code.

1. Introduction

The cerebellum typically makes up 20–30 % of the vertebrate brain by mass and volume (about half that in humans). Most of that is made up of the cerebellar cortex, which is divided into two layers. The inner layer, 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. Each mossy fibre terminal is surrounded by a semi-permeable membrane, a structure called a glomerulus. Golgi cells extend their axons into glomeruli, so granule cells receive both excitation and inhibition there. The granule cell axon rises into the outer layer of the cerebellar cortex, the molecular layer, where it divides in two. The two branches, termed parallel fibres, travel in opposite directions making synaptic contact in passing on Purkinje cells, which carry the sole output of the cerebellar cortex.

In the influential supervised learning model (Raymond and Medina, 2018, for an excellent review), control of Purkinje cells by granule cell rates is displaced by learned changes of parallel fibre synaptic weights (Albus, 1989; Brunel et al., 2004; Dean et al., 2010; Fujita, 1982). This obviates a requirement to explain how granule cell firing rates are derived from mossy fibre rates, or encode mossy fibre rate information, because it is implicit that a granule cell rate code would be redundant, because control of Purkinje cells is surrendered to learning.

There is accumulating evidence that rate-coded information is linearly conserved in communication between cerebellar neurons (Duguid et al., 2015; Duguid et al., 2012; Jelitai et al., 2016; Park et al., 2012; Rancz et al., 2007; Ritzau-Jost et al., 2014; Walter and Khodakhah, 2006, 2009), and that cells are adapted for faithful transmission (Cesana et al., 2010; Delvendahl and Hallermann, 2016; Moore and Blazis, 1989; Sargent et al., 2005; Saviane and Silver, 2006; Turecek et al., 2016, 2017). Firing rates that linearly code behavioural and task metrics have been observed 'at all levels of the cerebellar circuit' (Raymond and Medina, 2018 p.239, citing 30 sources). Evidence of linear signalling in the granular layer is rich. To decongest the main text, we include a detailed review in the Methods section, to which we cross-refer from the text.

It is difficult to reconcile linear relationships with network models that use synaptic weight changes because the theorised function of weight changes is to supplant a purely rate-coded response of networked units. Modelling the behaviour of a large cell network can feel like a blunt trauma anyway. For one thing, physiological units are not generic. Cells and synapses even of the same type both vary. We work with 'sampling units' – these can be cells, parts of cells, or some other structure. It forms part of our proposal that the cerebellum exploits statistical effects of random sampling of afferent rate data. Sampling units provide model parameters. We field evidence in the main text to explain the derivation of model parameters. Parameter values are supplied by the literature.

We use sampling units on the following basis in reasons. Cell-to-cell 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. Models vary in how they address that. Broadly, models tend to be divided between network models and biological neuron models. The difference is often partly a scale preference – network models sacrifice biophysics and biological neuron models aren't networks. On the face of it, the task is to identify (or select or abstract) and relate parameters to give the expected outcome. We present the view that this is not a modelling problem but a biological problem. We mean, the cerebellum must solve the problem of disengaging or otherwise compensating for intrinsic non-linear properties of neurons and synapses (Alvina et al., 2008; Sargent et al., 2005; Schonewille et al., 2006; Turecek et al., 2016). The reason is to isolate the effect of variables that code information and control for all other variables. Other variables are noise, including biophysical variables. The significance of evidence of linearity is that despite the seeming clamour, linear relationships are in fact isolated. In this view, the cerebellum is made up of a repeating network of linear units because it is adapted to isolate computational variables by an aggressive strategy of noise reduction. The units are sampling units.

We take it, in general, that output of a sampling unit is a linear function of the mean of input rates, $f(\langle r \rangle)$, which we take as the mean itself. We envisage that in motor circuits during behaviour signalling is analogue, avoiding the noise caused by discrete

signals. So, for example, Golgi cell dendritic signalling is a sustained state under modulation by fluctuating mossy fibre rates. Similarly, GABA spillover into the intraglomerular space, almost exclusively the agent of granule cell inhibition, becomes elevated by a build-up during behaviour, so that this too is at a sustained level, controlled by afferent Golgi cell rates. Thus, at each step, data are coded at any moment in a snapshot of a continuously variable state of change.

Historically, it has been challenging to imagine a useful function of linear units. We propose it has a statistical outcome which emerges at scale. We submit that

(i) there are reliable statistical effects of concurrent unit computations, which relate mossy fibre rates and granule cell rates at collective level; in

(ii) functionally-bounded cell groups, even though the granular layer is an anatomically seamless mesh of neurons; and that this is

(iii) the unaided result of the triad of cell type morphologies, network geometry and linear cerebellar signalling, and that it

(iv) provides a substrate for rate information coded in mossy fibre signals to be converted to control and coordination of Purkinje cells without synaptic weight changes in the historically expected role. This is because the form and spatial dimensions of the granule cell code mean all locations of a microzone receive the same code without receiving the same signals.

As we use a common function, we use a normalised scale. The output of the model is the granule cell code, and a regulated fraction of active granule cells. We include a statement of the relevant parts of the hypothesis at or near the beginning of each section.

2. **Mossy fibres**

2.1 Anatomy, 'ranks' and 'files'

Mossy fibres give rise to sagittally-extending collaterals, the direction of the long axis of microzones (Fig 1). These in turn give rise to multiple distal branches, each ending in a cluster of terminals (Shinoda and Sugihara, 2013; Shinoda et al., 2000; Sultan and Heck, 2003; Wu et al., 1999). Terminal clusters are sagittally aligned and separated by a randomly variable distance typically not less than 200 µm (Sultan and Heck, 2003; Wu et al., 1999) measured from edge to edge of clusters (so they are separated by a greater distance measured from centre to centre). Terminals themselves are intimately and randomly intermingled with terminals of other mossy fibres.

Accordingly, mossy fibres do not terminate at a point but a region with dimensions: the field occupied by a terminal cluster. A single mossy fibre gives rise to several

terminal branches and each of those to several terminals, such that a sagittal strip of the granular layer receives multiple copies of every signal at each of multiple locations.

Cluster size is variable (because it depends on the number of terminals and how tightly clustered they are), averaging 200 µm sagittally x 150 µm mediolaterally (Sultan and Heck, 2003). We term the average dimensions of a cluster a 'field'. The granular layer is anatomically seamless. A field is simply a nominal division that we use for modelling. A sagittal row of 75–100 fields laid end to end has the area (viewed from the cerebellar surface) of a mid-range microzone (Dean et al., 2010; Garwicz et al., 1998; Oscarsson, 1979; Ozden et al., 2009). We term a row of fields a 'rank'. Fields in a rank are not anatomically *or* functionally discrete – they are simply a device.

Figure 1

2.2 Hypothesis

Mossy fibre anatomy is an anatomical solution to a computational problem. The cerebellum exploits statistical effects of random sampling called the central limit theorem. For the effects to occur, samples must be independent, also termed sampling 'with replacement' (in the illustration sometimes used, coloured balls selected at

random from a bag are put back in the bag between samplings). Clearly, physical replacement is impossible because neither a terminal cluster nor a terminal can be in more than one place at the same time. However, anatomy simulates replacement. The distributed probabilities of the number of terminal branches per mossy fibre are the same as the probabilities of the number of repeat samplings if rates received by a rank were in fact randomly sampled with replacement by fields. Similarly, the distributed probabilities of the number of terminals per cluster are the same as the probabilities of number of repeat samplings if rates received by a field were in fact randomly sampled with replacement by individual neurons. Note that it is the population of *rates* that are hypothetically sampled with replacement, and not *signals.* The population of signals includes copies of rates because of branching and clustering. Copies are an anatomical device to approximate the result of sampling rates with replacement.

2.3 Terminal branching simulates 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 should mean any number of fields can receive the same signal, up to all of them. Clearly this is not true – there is still an anatomical limit. But in practice even with real replacement the number of repeat samplings is

constrained by a probability distribution. If, then, the number of terminal branches per mossy fibre has the same probability distribution, branching has the same result as if firing rates received by each field in a rank were in fact a random sample with replacement of firing rates received by the rank as a whole.

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

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

The results are shown in Table 1.

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

k	0	1	2	3	4	5	6	7	8	9	10	11
p	.007	.034	.084	.140	.175	.175	.146	.104	.065	.036	.018	.008

Although data are scarce, the number of repeat selections is consistent with what we know about the biological range of the number of terminal clusters per mossy fibre (Eccles et al., 1967; Sultan and Heck, 2003). Estimated by this method, the mean is 4–5. The anatomical cap on the number of fields that receive a signal (set by the number of terminal branches per cell) has the same distributed probability as the probabilities of the number of repeat samplings if the population of firing rates received by a rank was randomly sampled with replacement 100 times, representing 100 fields. Anatomy simulates replacement.

2.4 Terminal clusters simulate sampling with replacement

Terminal clustering of mossy fibres has the same function as terminal branching but at field scale. As far as we know, granule cells and Golgi cells receive contact from a random sample of mossy fibres that innervate a field, and more exactly a random sample of terminals. If a mossy fibre terminal branch ended in a single terminal, contact on any particular cell would place it beyond the range of most other cells. Mossy fibre termination in a cluster of terminals, and random intermingling of terminals (Fig 1), removes that restriction. Again, to approximate independence, it is only necessary for the number of repeat selections of a signal to have the same probability distribution as if there was actual replacement.

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

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

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

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

So, each field receives, at any given instant in time, a quasi-independent random sample of mossy fibre input rates to a rank, and single neurons individually receive a quasi-independent random sample of input rates to a field. The combined effect is a biological facsimile of simultaneous random sampling with replacement by single neurons of input rates *to a whole rank*.

3. **Golgi cells**

3.1 Derivation of ensemble dimensions and model parameters

4 Purkinje cells fit in a field, based on field size, thickness of the Purkinje cell dendritic plate – 9 µm (Hámori, 1992) – and the distance between them – ~40 µm (Palay and Chan-Palay, 1974). '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' (Lange, 1974). Possibly, estimates are low because smaller Golgi cells at deeper level were not counted (Palay and Chan-Palay, 1974 p.121). We take it that there are 10 Golgi cells per field, a ratio of 1:2.5 with 4 Purkinje cells.

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 direction of the long axis of a microzone (mean range 650 +/- 179 µm by 180 +/- 40 µm (in mice: Barmack and Yakhnitsa, 2008)) – and is the depth of the granular layer, which it fills vertically. Fine, beaded axon fibres enter glomeruli, where they inhibit granule cells (Hámori and Szentágothai, 1966). Because of the dimensions and 3-dimensional orientation of the Golgi cell plexus, all Golgi cells in a sagittal row of three fields inhibit substantially the whole of the middle field. Each field defines a functional population of Golgi cells – an ensemble – being those cells with convergent input to

that field. A single Golgi cell contacts thousands of granule cells (D'Angelo et al., 2013; Hámori and Szentágothai, 1966).

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 (Palay and Chan-Palay, 1974). 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.

3.2 Hypothesis

Despite forming an anatomically seamless dense mesh of axonal filaments that fills the granular layer, Golgi cells work in groups: ensembles. An ensemble is the population of a sagittal row of three fields. Groups are not anatomically bounded or even functionally discrete, but simply function as a single unit for the anatomical reason that the range of Golgi cells in 3 fields includes all or most of the middle field. Ensembles perform a computational conversion of mossy fibre rates into inhibition of granule

cells. The computation is (chiefly) a statistical result of random sampling, in two steps. At the first step, Golgi cell basal dendrites receive contact from a random sample of mossy fibres that innervate a field (strictly speaking, they randomly sample terminals); and in the second step, middle-field glomeruli each receive innervation by a random sample of Golgi cells. The effect of the computation is to synchronise inhibition of granule cells in the middle field. The strength of inhibition is linearly related to the mean of mossy fibre firing rates received as input to the entire rank, even though only a modest fraction of rates are received by three fields.

Ensembles conflate sagittally anatomically and functionally. Ensembles in a sagittal strip all return the same output (because they all perform the same computation on a large random sample of the same data). Effectively: the Golgi cell population of a whole rank functions as a single unit.

3.3 The Golgi cell computation

Say a population of values (with any frequency distribution) is randomly sampled. There are n samples, sample size m. If we take the mean (or sum) of each sample, and n is large, the sample means have a normal or near-normal distribution, by the central limit theorem. (This works even if m is a range.) The new distribution is centred on the same mean as the sampled distribution, but narrower. Modest values of n or m give an untidy result. However, if the new distribution is then itself randomly sampled, and the process is repeated, the results get tidier and narrower.

282 We simulated random sampling by Golgi cells of mossy fibre rates, and downstream,
283 random sampling by glomeruli of Golgi cell rates, both with replacement. We randomly
284 generated mossy fibre rates received by each field in a rank, taking into account mossy
285 fibre terminal branching and clustering, random variation of the number of terminals
286 per cluster and clusters per cell, and the fact that most clusters fall partly inside and
287 partly outside a field (see Methods, Derivation of mossy fibre rates). Rates were
288 obtained by sampling from a distribution representing signals received as input to the
289 rank as a whole, which were randomly generated subject to constraints on the shape
290 of the distribution.

291 This gave the number of mossy fibre signals, firing rates and the number of copies of
292 each rate received by each field, and therefore input to an ensemble territory,
293 populated by 30 Golgi cells, 10 per field. There are ~700 mossy fibre terminals per
294 field. As far as we know (and we propose), contact by mossy fibres on Golgi cells is at
295 random at field scale, i.e., inputs to a single basal dendrite are provided by a random
296 sample of local terminals. The number of physiological samplings is equal to the
297 number of basal dendrites per field; sample size is the number of mossy fibres which
298 make contact on each dendrite. The effect is sustained dendritic depolarisation under
299 modulation by a linear function of input rates (see Methods, Linear transmission of
300 mossy fibres to Golgi cells), which we take as the mean.

301 Taking the sample means generates a new set of values (Fig 2 column 2) with a smaller
302 range and centred on the mean, which is linearly related to the mean of the sampled
303 distribution. Basal dendritic charge transfer is passive and little filtered (Cesana et al.,
304 2010; Kanichay and Silver, 2008). Integration of dendritic states at the Golgi soma is
305 represented by taking the mean of dendritic charge, for each simulated cell. Golgi cell
306 firing rates have a linear relationship with the amplitude of depolarising current
307 (Dieudonné, 1998), suggesting (and we assume) a linear conversion of somatic
308 membrane potential into the Golgi cell firing rate.

309 The output is a third data set, Golgi cell firing rates (Fig 2 column 3), on the same
310 normalised scale. Each glomerulus in the middle field receives innervation from a
311 random subset of Golgi cells afferent to the field, and accordingly independently
312 randomly samples Golgi cell firing rates. Golgi cells inhibit granule cells in the
313 glomerulus. A granule cell extends a single dendrite into a glomerulus, typically
314 receiving synaptic contact from two or three Golgi cells (Rossi and Hamann, 1998).
315 However, inhibition of granule cells by Golgi cells is almost entirely by GABA spillover
316 into the intraglomerular space (Duguid et al., 2012), via extrasynaptic receptors
317 (Brickley et al., 1996; Nusser et al., 1998) (Methods, Linear transmission of Golgi cells
318 to granule cells). This effectively increases the Golgi cell to granule convergence ratio,
319 since the number of Golgi cells that are functionally afferent to a granule cell is not the
320 number that make direct contact, but the number afferent to a glomerulus. The
321 membrane around a glomerulus restricts GABA diffusion so that signalling is by

modulation of the intraglomerular concentration of GABA, which in turn reflects the rate it is released. The rate it is released by a single cell is proportional to the firing rate of that cell. The sum of rates is proportional to the mean. The convergence ratio is sample size; the number of samples is the number of glomeruli per field (~700, the same as terminals). We took the sample means to represent intraglomerular GABA concentration, which physiologically is under continuous modulation at a sustained level which is a linear function of the mean of afferent Golgi cell rates. The number of Golgi cells afferent to a glomerulus is unreported, as far as we know. We take it as a random number in the range 8–12, generated for each glomerulus, to reflect likely physiological variation. We tested other ratios to confirm that the ratio does not affect the ensemble computation (Supplementary Materials).

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

3.4 Results

3.4.1 An ensemble

There is a progressive narrowing of the data range at each step. Inputs to the system are converted to narrowly-grouped and normally-distributed outputs. Input values

represent mossy fibre signals received by a row of 3 fields; output values represent inhibition of granule cells in the middle field by GABA spillover in glomeruli. The narrow output range indicates that inhibition is locally synchronised, i.e., granule cells in a field receive co-modulated inhibition. Inhibition is statistically constrained to a linear relationship with mossy fibre rates, acting through the ensemble conversion. These results are independent of the shape of mossy fibre rates.

Figure 2

3.4.2 A rank

We next used the ensemble conversion to derive inhibition of granule cells in a sagittal row of 100 fields, a rank. The results are shown in Figure 3. As ensembles all perform the same computation and take as input (at any moment) the same frequency distribution of input rates, or nearly the same distribution, they all return the same output, or nearly the same output. There are 700 outputs per field. The mean of outputs, taken for each field, has a narrow range centred on the mean of mossy fibre input rates to the whole rank. Field-to-field variation of output appears to be inherent in the mechanism, i.e., caused by uncertainty of the outcome of random sampling, so it has a predictable range which the system can be built to tolerate.

So: 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 the mean of mossy fibre rates.

Figure 3

4. Granule cells

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

4.1.1 Hypothesis

Anatomy simulates independent random sampling by a single granule cell of the population of mossy fibre input rates to the whole rank (section 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. During behaviour, this takes place ceaselessly. The strength of inhibition is synchronised between glomeruli that populate a rank (section 3), so that all mossy fibre signals compete, at any

377 moment, against the same inhibitory threshold. The outcome in each glomerulus is
378 independent of the outcome at other terminals, including other terminals that
379 innervate the same granule cell. To fire, a granule cell must receive a minimum
380 number of mossy fibre inputs that prevail in the glomerular competition. The
381 probability that enough prevail depends, at any moment, on the strength of inhibition.
382 Since inhibition is synchronised, the probability is the same for all granule cells at any
383 time. The number that fire is very near the expected proportion (the number predicted
384 by probability) because granule cells are numerous. Because the linear relationship of
385 mossy fibre firing rates and inhibition is always conserved by the Golgi cell
386 computation, the probability is invariant, and therefore so is the proportion of active
387 granule cells.

388 Note: the idea of an all-or-nothing result of the intraglomerular 'balance' of excitation
389 and inhibition has appeared previously (Gilbert and Miall, 2021a).

390 **4.1.2 Anatomy**

391 Granule cells receive input to each dendrite from a single mossy fibre. Granule cells
392 have 3–5 dendrites, averaging 4 (the simulation assumes 4, but the result is the same
393 with 3–5). Each dendrite of a single granule cell extends into a different glomerulus.
394 Estimates vary of the number of granule cells that receive contact from a single mossy
395 fibre terminal. 50 is mid-range (Gao et al., 2016; Jakab and Hámori, 1988), but
396 estimates vary significantly (Ritzau-Jost et al., 2014). 50 is consistent with our

397 estimates of the number of terminals and granule cells in a field. The granule cell
398 sampling unit is the whole cell since it receives only a single input rate to each
399 dendrite. Sample size is therefore 4, and the number of samples per field is the
400 number of granule cells, which we estimate at 8,750 (see Methods).

401 **4.1.3 Feed-forward inhibition regulates the probability that granule cells fire**

402 Excitation and inhibition compete for influence in a glomerulus. It is unclear if there is
403 always an emphatic winner, but a swing in the dominant direction may be amplified by
404 positive feedback, acting through presynaptic receptors. Activation of presynaptic
405 $GABA_B$ receptors on mossy fibres inhibits glutamate release (Mitchell and Silver,
406 2000a), while activation of presynaptic mGluR2 receptors on Golgi cells inhibits GABA
407 release (Mitchell and Silver, 2000b). We postulate that only mossy fibre signals which
408 prevail in the glomerular competition contribute significantly to granule cell
409 depolarisation. To fire, a granule cell must receive a minimum number of competitive
410 mossy fibre signals – thought to be three (Billings et al., 2014; Jörntell and Ekerot,
411 2006). Certainly, one is too few (Bengtsson and Jörntell, 2009; Chadderton et al., 2004;
412 Duguid et al., 2012).

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

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

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 is synchronised and scale 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.

4.1.4 Does physiology have that outcome?

Probability is an abstraction with no direct physiological counterpart. The simulation does not include a binomial function but derives output from input by comparing excitatory to inhibitory input to each glomerulus, the way we claim it is derived physiologically. The equation is simply shorthand – biology exploits chance and large numbers to generate the same result. To test that, we incorporated the Golgi cell ensemble conversion into a model that recodes mossy fibre signals into granule cell signals. 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 per section 3 and Methods, so taking into account,

434 *inter alia*, that a field receives a randomly variable number of copies of most signals,
435 and most terminal clusters lie only partly inside a field. Inhibition – outputs of the
436 ensemble conversion, 700 per field – is provided by the Golgi cell model. Excitation and
437 inhibition received by each glomerulus are randomly paired. Pairing is used to identify
438 the subset of granule cells that fire (cells that receive at least 3 mossy fibre signals
439 which exceed the inhibitory threshold). Output of the model is the number of granule
440 cells that fire in each field, the mean of mossy fibre rates they each receive (the sample
441 means), and the mean of the sample means in each field ('field means'). The results
442 are shown in Figure 4.

443 **4.1.5 Results**

444 Feed-forward inhibition – mossy fibre → Golgi cell → granule cell – may be the
445 primary regulator of the number of granule cells that fire. Field-to-field random
446 variability of the number is inversely proportional to the fraction that are active,
447 suggesting that, if primary regulation is in this form, the fraction is low. We anticipate
448 nonetheless a smaller range relative to the mean, indicating the physiological presence
449 of a supplementary mechanism that we have not represented in the model.

450 A local feedback mechanism may be provided by granule cell ascending axon contact
451 on Golgi cells. The ascending portion of the granule cell axon makes synaptic contact
452 on local Golgi cells, on basal dendrites (Cesana et al., 2010). A Golgi cell receives ~400
453 functional synaptic contacts from this source, around 25% of the granule cells that a

454 Golgi cell receives contact from in total. During sustained mossy fibre discharge,
455 synapses from local granule cells contribute equivalent charge to Golgi cells as mossy
456 fibre synapses and 'display similarly fast kinetics' (Cesana et al., 2013). Short-latency
457 inhibitory feedback may add homeostatic fine adjustment to regulation at local level,
458 at field scale.

459 It is a long-standing idea (Albus, 1971) that granule cell activity may be regulated by
460 feedback via parallel fibres, which contact Golgi cell apical dendrites. We have
461 previously proposed that gap junctions between apical dendrites may provide the
462 means for Golgi cells to detect the density of parallel fibre activity (Gilbert and
463 Rasmussen, 2024). However, feedback by this pathway is not sensitive to local
464 fluctuation of granule cell active numbers.

465 The results are the same with any (tested) distribution of mossy fibre rates and also
466 whether mossy fibres are all active or some are active and others are not. Active cells
467 can be in any random permutation, subject to there being a minimum number that are
468 active at any time.

469

470 **Figure 4**

471

472 ## 4.2 Granule cells encode mossy fibre rate information at a collective
473 level that mirrors organisation of the molecular layer into microzones

474 ### 4.2.1 Hypothesis

475 Anatomy simulates independent random sampling by a single granule cell of the
476 population of mossy fibre input rates to the whole rank (section 2). The granule cell
477 firing rate is a linear function of depolarising somatic charge, in turn a linear function
478 of the mean (or sum) of dendritic depolarisations. Cells that fire receive rate-
479 proportional dendritic depolarisation that is undiluted by inhibition because they
480 receive dominant mossy fibre signals. Derived in this way, granule cell rates are
481 collectively well predicted by mossy fibre rates, by the central limit theorem. This
482 provides the substrate for granule cells to collectively encode mossy fibre rate
483 information. Code groups are functionally defined by the topography of mossy fibre
484 input to the cerebellum, as Golgi cell networks are. The output of a rank is the smallest
485 unit of granule cell encoded information. The permutation of active mossy fibres, and
486 of granule cells, and how firing rates are distributed among active cells, all code
487 nothing.

488 ### 4.2.2 Do granule cell firing rates – the output of a rank – have the hypothesised
489 features?

We ran the same simulation as section 4.1 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. We take granule cell firing rates to be a linear function of the mean of input rates to each active cell, which we take as the mean itself. Field means are therefore the mean of granule cell firing rates in each field. The results are shown in Figures 5 and 6. We present evidence relating to mossy fibre → granule cell linear transmission in the Methods section (Methods, Linear transmission of mossy fibres to granule cells).

The narrowing and predicted shape of the granule cell frequency distribution is evident in all conditions. The field 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 very close to a straight line. The field means follow the mean of mossy fibre rates at a precisely fixed distance. When mossy fibre rates have a random uniform distribution, the field means remain almost a straight line, and the distance from the mossy fibre mean varies with the size of the mossy fibre range.

A straight line indicates that fields all return the same output – not the same firing rates, but the same narrowed range, clustered around the mean, and attracted to a normal distribution (more evident with narrowly and/or normally distributed mossy fibre rates). This holds even though a single granule cell receives a tiny fraction of

510 mossy fibre rates, and each field receives only a small random sample of mossy fibre
511 signals received as input to the rank as a whole, and it is independent of the shape of
512 the distribution of mossy fibre rates.

513 This is not in itself necessarily in fatal conflict with the traditional cerebellar learning
514 model, but has the alternative interpretation that granule cell firing conserves mossy
515 fibre rate information, coded in the output of a whole rank. This is therefore the
516 minimum resolution of granule cell information but with the great flexibility that the
517 same information is coded in any random sample of granule cell signals, at any
518 location. It does not matter which cells they are, so there is no need for patterns to be
519 repeated, or received at any particular set of synapses downstream. It is immaterial
520 which cells fire at what rates, or indeed what rates they are, provided any random
521 sample of rates is as if drawn from the same frequency distribution, with replacement.
522 The output of a rank can be likened by analogy to a pixel: it is only divisible into smaller
523 pieces of the same information.

524

525 **Figures 5 and 6**

526

527

5. **Discussion**

Neural networks are believed to carry out sophisticated computations. In the influential traditional cerebellar learning model, and many artificial neural networks, computations are implemented by learned synaptic changes that selectively adjust synaptic transmission strength. We propose a rethink, taking the granular layer of the cerebellar cortex, the input layer of the cerebellar network, as a physiologically-detailed example. The proposals are quantified and tested in silico and empirically testable. Some features are:

i) Information has geometry. Large cell groups talk to each other in a statistical rate code. At any instant, a whole group codes the same information. Accordingly, information has spatial dimensions, the dimensions of a code group.

ii) Granule cell code groups form long, thin strips, mirroring microzonal organisation of the molecular layer, so internal code has a striped pattern. A strip/code group functions as a single unit and spans the cerebellar network sagittally.

iii) Recoding exploits the central limit theorem. The mean of any random sample of granule cell firing rates has a linear relationship with the mean of mossy fibre input rates to the whole strip (and granule cell rates have a reduced range). The pattern of active cells, and which cells fire at what rates, codes nothing.

iv) Accordingly, while firing of a whole code group is the minimum unit of information, it is not the minimum unit of code, because inside those dimensions, the same information is coded in any random sample of firing rates.

v) There is a computational effect, unaided, of the triad of cell morphologies, network architecture and linear transmission. The cerebellum is highly adapted for this result. No learning algorithm or learning is needed.

vi) There is a huge sacrifice of data resolution on entry to the cerebellum, counter to the general expectation that more data should inform more powerful brain computations.

vii) This is functional. Coded in this form, granule cell data can drive coherent firing of large cell groups downstream.

viii) A statistical code may be a physiological strategy to mitigate variable performance of single neurons and synapses, a long-known impediment to precise and replicable biological information capture. It is a long-standing sometimes implicit notion that a repeat of conditions that provide information to the brain generates a repeat of the same brain activity. In fact it is at least not clear there is a faithful correlation, or indeed that conditions are ever exactly repeated. Code that does not depend on high fidelity may be a solution to working with unreliable information.

564 Microzones lie at right angles to parallel fibres, presenting a large target. As a result, a
565 Purkinje cell receives none of the same signals as almost all other Purkinje cells in the
566 same microzone. However, the form of the granule cell code means the same
567 information is received by all locations of sagittal plane, so that microzone-grouped
568 Purkinje cells receive the same information without receiving the same signals. An
569 estimated 350,000 parallel fibres pass through a Purkinje cell dendritic territory
570 (Harvey and Napper, 1988, 1991; Napper and Harvey, 1988a, b). If even only a small
571 fraction of parallel fibres is active (0.5%), and allowing for the fact that only one in two
572 make contact, and that a majority of synapses are severely depressed or silent (Ekerot
573 and Jörntell, 2001; Isope and Barbour, 2002), a single Purkinje cell still receives around
574 200 signals in any snapshot of parallel fibre activity. This provides a substrate to
575 control firing of Purkinje cells. Control is by a chain of rate codes. Purkinje cells linearly
576 code the strength of granule cell inputs (Park et al., 2012; Walter and Khodakhah,
577 2006, 2009) and behavioural metrics (Chen et al., 2016; Dugué et al., 2017; Herzfeld et
578 al., 2015; Hong et al., 2016; Jelitai et al., 2016). Learned changes of synaptic weights in
579 the form traditionally envisaged may not be needed. Also, it can explain how Purkinje
580 cell firing rates in a microzone are modulated in union.

581 A note on methodology. With our approach, the function of the computation is not
582 inherent in the mechanism. This breaks with a long tradition network modelling.
583 Generally, neural network models solve a problem it is proposed the network has. The
584 proposal is an assumption of the model; the solution is a computational mechanism;

585 the mechanism is the model. We don't propose a function, only a mechanism. We
586 work towards a function.

587 Computational anatomy provides a tool to habilitate linear signalling, historically not
588 popular with modellers because it makes unit computations less flexible.

589

590 **METHODS**

591 **DERIVATION OF MOSSY FIBRE RATES**

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

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

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

600 The number of mossy fibres afferent to a field has been estimated at 100 (Sultan and Heck,
601 2003). However, that number is a field-sized fraction of the general population. The actual
602 number is larger because mossy fibres end in a dispersed group (cluster) of terminals, and
603 most groups straddle field limits.

604 We work from an estimate of the total number of terminals per field, which is, in turn, derived
605 from convergent estimates of the number of granule cells per field. (Estimate 1) There are an
606 estimated 1.92×10^6 granule cells per µl (Harvey and Napper, 1988). The number of fields
607 which fit in that volume is $(1{,}000/150)^2 \times (1{,}000/200) = \sim 222$, assuming a field depth of 150 µm
608 (Sultan and Heck, 2003). The number of granule cells in a field is, therefore 1,920,000/222 =
609 8,649. (Estimate 2) An estimated 350,000 parallel fibres pass through a single Purkinje cell's
610 territory (Harvey and Napper, 1988). Parallel fibres run about 3 mm in each direction from the
611 granule cell soma (Brand et al., 1976; Mugnaini, 1983). The Purkinje cell dendritic arbour fills
612 the molecular layer vertically and averages ~200 µm wide. Therefore, 6,000/150 = 40 fields fit
613 in a row that provides parallel fibres that pass through a Purkinje cell dendritic territory, giving
614 350,000/40 = 8,750 granule cells per field. Estimates vary of the number of granule cells which
615 extend a dendrite into a single glomerulus. 50 is mid-range. (8,750/50) x 4 = 700 glomeruli per
616 field.

617 Granule cells have 3–5 dendrites, averaging 4, and receive input to each of them from a single
618 mossy fibre (so an average granule cell receives 4 inputs). Estimates of divergence of a single
619 mossy fibre terminal onto granule cells vary (Ritzau-Jost et al., 2014). It is unclear how much
620 this reflects biological diversity versus differences in opinion and methods. Here we use a mid-
621 range estimate of 1:50 (Gao et al., 2016; Jakab and Hámori, 1988). Assuming a mossy fibre
622 terminal receives a single granule cell dendrite from each of *d* granule cells, and given an
623 average of 4 dendrites per granule cell and *z* granule cells per field, the total number of mossy
624 fibre terminals in a field is $N = (z \times 4)/d = 700$ given $z = 8{,}750$ and $d = 50$. A convergent
625 estimate is provided by multiplying the per field number of mossy fibres by the average
626 number of terminals per cluster.

627 Cluster size varies. We take the range as 4–12 terminals per cluster (Sultan and Heck, 2003).
628 We assume the number varies at random and cluster sizes occur with equal probability. A
629 maximum of 7 fits in a field (because a field shares the dimensions of the area that encloses a
630 mean cluster). Fields are not anatomically bounded, so many clusters straddle field limits. To
631 reflect that, we assume there is a 1/7 probability that a 7-terminal cluster (or larger) has 7
632 terminals inside a field and also the same probability of any other number. A 6-terminal cluster
633 has a 2/7 chance that all 6 terminals are inside and a 1/7 chance of any other number, and so
634 on.

635 Table 1 shows the derivation of the expected (i.e., predicted by probability) ratios of terminals
636 contributed to a field by each cluster size (R1), not counting terminals that lie outside when a
637 cluster straddles field boundaries; and the ratios of mossy fibres contributing $n = 1$–7 terminals
638 (R2), regardless of cluster size. R1 ratios are given by the sum of numbers in each row. So, for
639 example, the ratio of cells that end in 8 terminals to cells that end in 5 is 28:25. R2 ratios are
640 given by the number of appearances of each number in the grid – so the expected ratio of cells
641 which contribute 7 terminals to the number which contribute 4 is 6:12. The ratios effectively
642 give the distributed probability of possible states. For example, there is a probability of $10/(6 +$
643 $8 + 10 + 12 + (3 \times 9)) = 0.159$ that a mossy fibre contributes 5 terminals to a field (regardless of
644 cluster size). Similarly, there is a probability of $((6 \times 28)/((6 \times 28) + 27 + 25 + 22))/7 = 0.099$ that
645 a given cell contributes 7 terminals to a field.

Table 1. Derivation of expected ratios

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

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 therefore given by

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

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

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

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

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

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

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

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

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

Mossy fibre input rates to a field

675 Terminal clustering means a field receives multiple copies of most signals. The proportions of
676 mossy fibres afferent to a simulated field that contribute each possible number of terminals
677 (1–7) are generated in simulated steps. In the first step, the simulation randomly generates,
678 for each field, the cluster size ratio. This means, the relative proportions of afferent mossy
679 fibres (which contribute at least 1 terminal to a field) that end in 4 terminals, 5, 6 and so on up
680 to 12, assuming they occur with equal probability, disregarding how many terminals fall inside
681 and outside field boundaries. In step two, the ratio of cluster sizes is weighted to reflect the
682 greater probability that large clusters contribute more terminals to a field. In step 3, ratios are
683 converted to the fraction of terminals contributed by each cluster size. When multiplied by the
684 total number of terminals, this gives the number of terminals contributed by each cluster size.
685 In step four, the values obtained in step three are divided up between cells that contribute a
686 single terminal, 2 terminals, 3 and so on, to reflect straddling of field limits. The proportions
687 give the number of mossy fibres afferent to a simulated field that contribute a single terminal,
688 two, three and so on up to seven, the maximum.

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

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

697

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

699 It is part of our proposal that Golgi cell basal dendritic signals linearly reflect mossy fibre rates,
700 so that mossy fibre rate information is conserved proportionally in Golgi cell firing rates. Golgi
701 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 (Dieudonné, 1998 p.845). 'Golgi cells can follow peripheral signals in a continuous fashion, modulating their frequency with the intensity of the stimulus [citing (Edgley and Lidierth, 1987; Miles et al., 1980)]' (Galliano et al., 2010). As a result, during movement, the 'output of Golgi cells varies smoothly as a function of input.…They are capable of modulating their output to reflect their input with little noise or signal distortion' (Moore and Blazis, 1989). 'Sensory-evoked Golgi-cell inhibition scales proportionally with the level of mossy fiber excitatory synaptic input', such that inhibition reliably conveys mossy fibre rate information (Duguid et al., 2015). On this evidence: mossy fibre rate information is conserved proportionally in Golgi cell firing rates.

Golgi cells extend their basal dendrites into glomeruli (Eccles et al., 1964; Hámori and Szentágothai, 1966). Contact on them by mossy fibres is multi-synaptic (Palay and Chan-Palay, 1974), contributing to a reliable, rapid (submillisecond) response. The fast rise time and weak distance dependence of the amplitude and timing of Golgi cell EPSCs evoked by mossy fibre signals, suggests that basal dendrites introduce relatively little filtering (Cesana et al., 2010; Kanichay and Silver, 2008), as expected for large-diameter dendrites. We note that the pause in Golgi cell firing following a discrete stimulus under anaesthesia (Duguid et al., 2015; Holtzman et al., 2006; Vos et al., 1999) disappears from Golgi cell recordings during locomotion (Edgley and Lidierth, 1987).

During behaviour in freely-moving animals, mossy fibre activity is dense: a high fraction of mossy fibres are active (Hartmann and Bower, 2001). Both mossy fibres (Arenz et al., 2008; Giovannucci et al., 2017; van Kan et al., 1993) and Golgi cells (Edgley and Lidierth, 1987) have been reported to fire with a sustained, time-varying signature. In the behaving animal, Golgi cell basal dendritic membrane potential is a continuous variable under modulation by sustained inputs, we submit. As charge transfer to the soma is passive, 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 (van Kan et al., 1993). In the simulations, we use the physiological rates, within stated constraints on the shape of the distribution.

731

732 **Other control of Golgi cells?** Do Golgi cells receive a significant influence of input from other
733 sources? The evidence is incomplete but currently suggests other influence is absent or weak.
734 Contrary to early reports, neither Purkinje cells nor climbing fibres contact Golgi cells (Pietrajtis
735 and Dieudonné, 2013, who give references), and only a modest minority of inhibitory inputs to
736 Golgi cells (if any) are from molecular layer interneurons (Eyre and Nusser, 2016), which
737 generate weak synaptic currents (Dumoulin et al., 2001), consistent with either extremely
738 weak or wholly absent innervation (Hull and Regehr, 2012). There is conflicting evidence
739 whether Golgi cells make inhibitory synaptic contact on each other (Hull and Regehr, 2012;
740 Vervaeke et al., 2010). Our simulation does not include an effect on Golgi cell firing by other
741 sources of input.

742

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

744 'In the adult rat cerebellum, each granule cell dendrite receives 2.6 ± 0.55 synaptic contacts
745 from Golgi axon terminals' (Rossi and Hamann, 1998) citing (Jakab and Hámori, 1988).[1]
746 However, the large majority of inhibition (98%) is by spillover (Duguid et al., 2012), where
747 neurotransmitter released into the synaptic cleft spills out into the glomerular space. Golgi
748 cells release GABA, an inhibitory neurotransmitter. This is detected by high-affinity $GABA_A$
749 receptors located perisynaptically and extrasynaptically on granule cells (Brickley et al., 1996;
750 Nusser et al., 1998). Even synaptically-received signals have most of their effect (i.e., the large
751 majority of charge transfer is mediated) via spillover.[2]

[1] Data are sparse. The original data are an estimate derived from visual inspection of slices thirty years ago so far without later corroboration, except from Rossi and Hamann 1998, who derive a lower estimate but propose the disparity is because their estimate is too low.
[2] Because of the long duration of IPSCs generated by spillover (recorded in slices), the total charge carried is three times that of IPSCs generated by directly connected terminals. Rossi, D.J., and Hamann, M. (1998). Spillover-mediated transmission at inhibitory synapses promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry. Neuron 20, 783-795.

752 Golgi cells fire spontaneously (Ruigrok et al., 2011) at a time-varying rate in the behaving
753 animal, so that downstream, a glomerulus receives continuous convergent input. As a result,
754 there is a sustained build-up of glomerular GABA during behaviour (Rossi et al., 2003) at an
755 adjustable concentration controlled by the mean of afferent Golgi cell firing rates (Mapelli et
756 al., 2014). A glomerulus randomly samples Golgi cells afferent to a field. Golgi cells provide the
757 only inhibitory input to granule cells.

758 Signalling by spillover is sometimes assumed to be ambient and slow. However, the action of
759 glomerular GABA spillover has a fast phasic component – not as fast as synaptic transmission
760 (~1 ms) but with a rise time of only a few milliseconds (Mapelli et al., 2014). Unlike the spiky
761 appearance of synaptically-induced inhibitory postsynaptic currents, spillover (Duguid et al.,
762 2015) generates a sustained outward current. The response time to multiple inputs is likely
763 faster than the experimentally isolated response to a discrete signal (because the combined
764 rate of GABA release is higher).

765 **Golgi cell oscillations.** Golgi cells fire autonomously. Under anaesthesia, firing falls into a slow
766 oscillating pattern (Ros et al., 2009). Under excitatory synaptic input, however, this disappears
767 (Vervaeke et al., 2010). Anaesthesia 'has a strong influence on spontaneous activity of Golgi
768 cells' (Pietrajtis and Dieudonné, 2013). Discharge at low rates with irregular timing under
769 anaesthesia (Ruigrok et al., 2011) is replaced by higher rates with more regular timing without
770 anaesthesia (Edgley and Lidierth, 1987; Miles et al., 1980; Rasmussen et al., 2014). (This
771 paragraph is a cautionary note about evidence obtained under anaesthesia, which has been
772 used to argue that oscillations provide a form of signalling.)

773

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

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

777 The short and equal length of granule cell dendrites would suggest light and equal filtering.
778 The mossy fibre-granule cell connection has a range of adaptations which support high-fidelity

779 transmission of high frequency signals across a wide bandwidth (Delvendahl and Hallermann,
780 2016). Faithful transmission of rate information is reported (Rancz et al., 2007; Ritzau-Jost et
781 al., 2014). Vesicle release and replenishment are fast (Sargent et al., 2005; Saviane and Silver,
782 2006) and postsynaptic AMPA receptors operate in their linear range (Sargent et al., 2005),
783 where they are resistant to desensitisation (DiGregorio et al., 2007). Multiple contacts are
784 made by a mossy fibre with each granule cell (Jakab and Hámori, 1988). Conversion of
785 depolarising somatic charge into granule cell firing is a simple function, such that granule cells
786 'have a relatively linear and uncomplicated conversion of depolarisation level to spike rate'
787 (Bengtsson and Jörntell, 2009 p.2393, citing Jorntell and Ekerot 2006 and D'Angelo et al 1998).

788 Glutamate spillover enhances precision and reliability of transmission (DiGregorio et al., 2002;
789 Sargent et al., 2005). During behaviour, at sustained physiological mossy fibre rates, there may
790 be a build-up of intraglomerular glutamate, increasing the relative influence of spillover in a
791 balance with synaptic transmission, assisted by short-term synaptic depression. In these
792 conditions, spillover may dominate, as it dominates inhibitory transmission (Duguid et al.,
793 2012; Mapelli et al., 2014). Spillover from multiple release sites (200-400 at a single mossy
794 fibre terminal) is plausibly sufficient to mitigate variability of vesicular size, release probability,
795 and the number of release points at synapses made on any individual cell, increasing fidelity
796 and equality of excitation of granule cells.

797 Against that, there are reported to be heterogeneous synaptic weights. Different strength and
798 short-term dynamics of the mossy fibre to granule cell connection have been reported in
799 vestibular slices, with $GABA_A$ receptors blocked. The amplitude of the response, measured at
800 the soma, is reported to be statistically correlated to the source of the mossy fibre that is
801 stimulated (Chabrol et al., 2015), suggesting that the strength of the connection depends on
802 the source.

803 However, the presence of weights does not necessarily mean that the function is to modulate
804 firing rates. Rather, it is suggested that the function may temporally code the source of inputs
805 because different combinations have different response onset times. An alternative, which
806 they also mention, is that some inputs may be in a supporting role, reminiscent of 'driver' and
807 'modulatory' signals in the thalamus and cortex (Sherman and Guillery, 2011). Feasibly, signals

808 in a supporting role must be present for the postsynaptic granule cell to fire, but do not affect
809 the granule cell firing rate.

Grants

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

Acknowledgement

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

Author contributions

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

Declaration of interests

The authors declare no competing interests.

825 **REFERENCES**

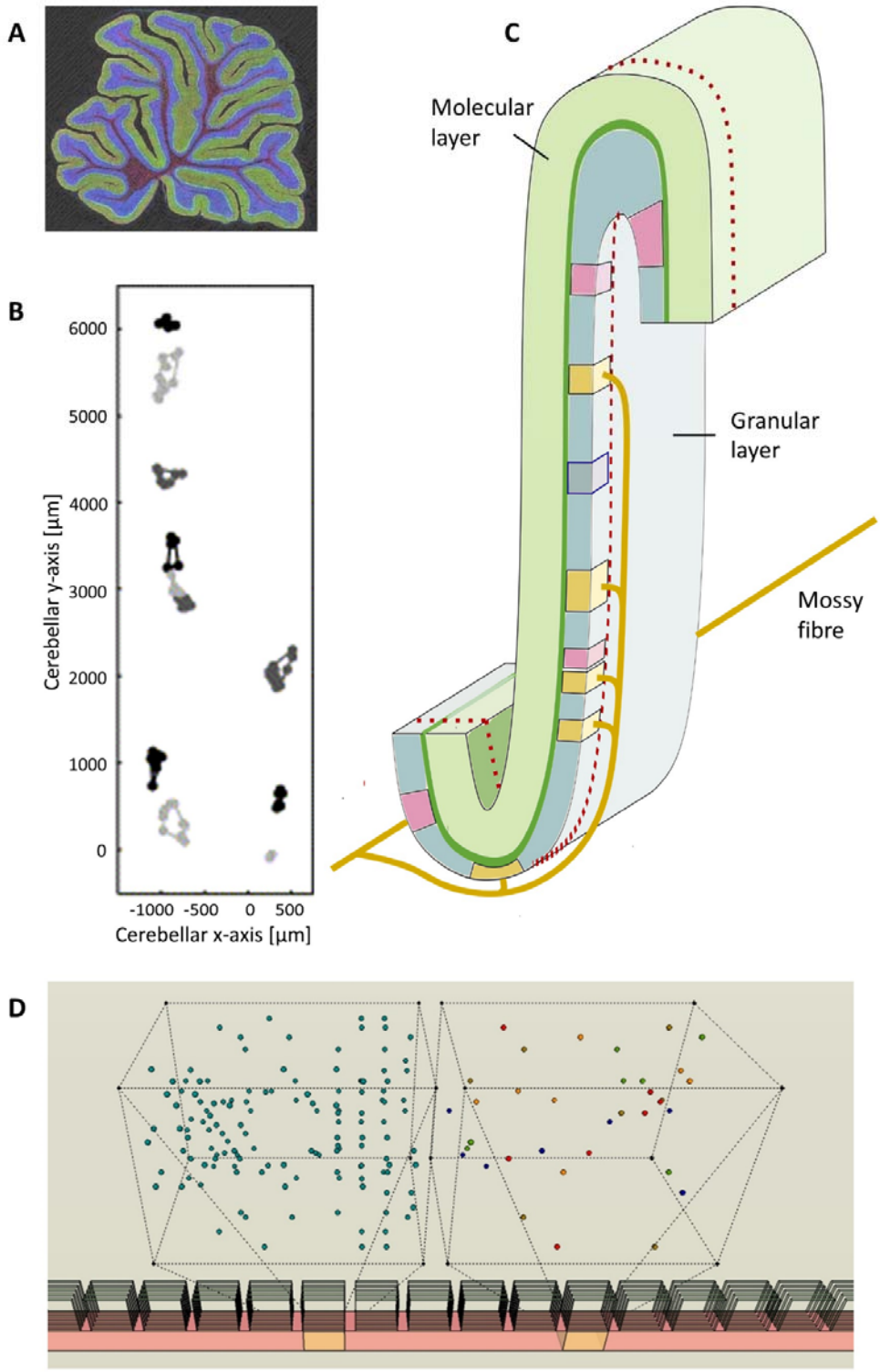
826 Albus, J.S. (1971). A Theory of Cerebellar Function. Mathematical Biosciences 10, 25-61.
827 Albus, J.S. (1989). The Marr and Albus theories of the cerebellum-two early models of
828 associative memory. In Digest of Papers COMPCON Spring 89 Thirty-Fourth IEEE Computer
829 Society International Conference: Intellectual Leverage, pp. 577-582.
830 Alvina, K., Walter, J.T., Kohn, A., Ellis-Davies, G., and Khodakhah, K. (2008). Questioning the
831 role of rebound firing in the cerebellum. Nature neuroscience 11, 1256-1258.
832 Arenz, A., Silver, R.A., Schaefer, A.T., and Margrie, T.W. (2008). The contribution of single
833 synapses to sensory representation in vivo. Science (New York, NY) 321, 977-980.
834 Barmack, N.H., and Yakhnitsa, V. (2008). Functions of interneurons in mouse cerebellum. The
835 Journal of neuroscience : the official journal of the Society for Neuroscience 28, 1140-1152.
836 Bengtsson, F., and Jörntell, H. (2009). Sensory transmission in cerebellar granule cells relies on
837 similarly coded mossy fiber inputs. Proceedings of the National Academy of Sciences of the
838 United States of America 106, 2389-2394.
839 Billings, G., Piasini, E., Lorincz, A., Nusser, Z., and Silver, R.A. (2014). Network structure within
840 the cerebellar input layer enables lossless sparse encoding. Neuron 83, 960-974.
841 Brand, S., Dahl, A.L., and Mugnaini, E. (1976). The length of parallel fibers in the cat cerebellar
842 cortex. An experimental light and electron microscopic study. Experimental brain research 26,
843 39-58.
844 Brickley, S.G., Cull-Candy, S.G., and Farrant, M. (1996). Development of a tonic form of synaptic
845 inhibition in rat cerebellar granule cells resulting from persistent activation of GABAA
846 receptors. The Journal of physiology 497 (Pt 3), 753-759.
847 Brunel, N., Hakim, V., Isope, P., Nadal, J.P., and Barbour, B. (2004). Optimal information
848 storage and the distribution of synaptic weights: perceptron versus Purkinje cell. Neuron 43,
849 745-757.
850 Cesana, E., Dieudonné, S., Isope, P., Bidoret, C., D'Angelo, E., and Forti, L. (2010). Novel granule
851 cell-Golgi cell excitatory input in the cerebellar granular layer. In FENS Abstr.
852 Cesana, E., Pietrajtis, K., Bidoret, C., Isope, P., D'Angelo, E., Dieudonne, S., and Forti, L. (2013).
853 Granule cell ascending axon excitatory synapses onto Golgi cells implement a potent feedback
854 circuit in the cerebellar granular layer. The Journal of neuroscience : the official journal of the
855 Society for Neuroscience 33, 12430-12446.
856 Chabrol, F.P., Arenz, A., Wiechert, M.T., Margrie, T.W., and DiGregorio, D.A. (2015). Synaptic
857 diversity enables temporal coding of coincident multisensory inputs in single neurons. Nature
858 neuroscience 18, 718-727.
859 Chadderton, P., Margrie, T.W., and Hausser, M. (2004). Integration of quanta in cerebellar
860 granule cells during sensory processing. Nature 428, 856-860.
861 Chen, S., Augustine, G.J., and Chadderton, P. (2016). The cerebellum linearly encodes whisker
862 position during voluntary movement. eLife 5, e10509-e10509.
863 D'Angelo, E., Solinas, S., Mapelli, J., Gandolfi, D., Mapelli, L., and Prestori, F. (2013). The
864 cerebellar Golgi cell and spatiotemporal organization of granular layer activity. Frontiers in
865 neural circuits 7, 93.
866 Dean, P., Porrill, J., Ekerot, C.F., and Jörntell, H. (2010). The cerebellar microcircuit as an
867 adaptive filter: experimental and computational evidence. Nature reviews Neuroscience 11,
868 30-43.

869 Delvendahl, I., and Hallermann, S. (2016). The Cerebellar Mossy Fiber Synapse as a Model for
870 High-Frequency Transmission in the Mammalian CNS. Trends Neurosci 39, 722-737.
871 Dieudonné, S. (1998). Submillisecond kinetics and low efficacy of parallel fibre-Golgi cell
872 synaptic currents in the rat cerebellum. The Journal of physiology 510 (Pt 3), 845-866.
873 DiGregorio, D.A., Nusser, Z., and Silver, R.A. (2002). Spillover of glutamate onto synaptic AMPA
874 receptors enhances fast transmission at a cerebellar synapse. Neuron 35, 521-533.
875 DiGregorio, D.A., Rothman, J.S., Nielsen, T.A., and Silver, R.A. (2007). Desensitization properties
876 of AMPA receptors at the cerebellar mossy fiber granule cell synapse. The Journal of
877 neuroscience : the official journal of the Society for Neuroscience 27, 8344-8357.
878 Dugué, G.P., Tihy, M., Gourévitch, B., and Léna, C. (2017). Cerebellar re-encoding of self-
879 generated head movements. eLife 6.
880 Duguid, I., Branco, T., Chadderton, P., Arlt, C., Powell, K., and Hausser, M. (2015). Control of
881 cerebellar granule cell output by sensory-evoked Golgi cell inhibition. Proceedings of the
882 National Academy of Sciences of the United States of America 112, 13099-13104.
883 Duguid, I., Branco, T., London, M., Chadderton, P., and Hausser, M. (2012). Tonic inhibition
884 enhances fidelity of sensory information transmission in the cerebellar cortex. The Journal of
885 neuroscience : the official journal of the Society for Neuroscience 32, 11132-11143.
886 Dumoulin, A., Triller, A., and Dieudonné, S. (2001). IPSC kinetics at identified GABAergic and
887 mixed GABAergic and glycinergic synapses onto cerebellar Golgi cells. The Journal of
888 neuroscience : the official journal of the Society for Neuroscience 21, 6045-6057.
889 Eccles, J.C., Ito, M., and Szentágothai, J. (1967). The cerebellum as a neuronal machine (Berlin,
890 New York etc.: Springer-Verlag).
891 Eccles, J.C., Llinás, R., and Sasaki, K. (1964). Golgi cell inhibition in the cerebellar cortex. Nature
892 204, 1265-1266.
893 Edgley, S.A., and Lidierth, M. (1987). The discharges of cerebellar Golgi cells during locomotion
894 in the cat. The Journal of physiology 392, 315-332.
895 Ekerot, C.F., and Jörntell, H. (2001). Parallel fibre receptive fields of Purkinje cells and
896 interneurons are climbing fibre-specific. The European journal of neuroscience 13, 1303-1310.
897 Eyre, M.D., and Nusser, Z. (2016). Only a Minority of the Inhibitory Inputs to Cerebellar Golgi
898 Cells Originates from Local GABAergic Cells. eNeuro 3.
899 Fujita, M. (1982). Adaptive filter model of the cerebellum. Biological cybernetics 45, 195-206.
900 Galliano, E., Mazzarello, P., and D'Angelo, E. (2010). Discovery and rediscoveries of Golgi cells.
901 The Journal of physiology 588, 3639-3655.
902 Gao, Z., Proietti-Onori, M., Lin, Z., Ten Brinke, M.M., Boele, H.J., Potters, J.W., Ruigrok, T.J.,
903 Hoebeek, F.E., and De Zeeuw, C.I. (2016). Excitatory Cerebellar Nucleocortical Circuit Provides
904 Internal Amplification during Associative Conditioning. Neuron 89, 645-657.
905 Garwicz, M., Ekerot, C.F., and Jorntell, H. (1998). Organizational principles of cerebellar
906 neuronal circuitry. News in Physiological Sciences 13, 28-34.
907 Gilbert, M., and Miall, R.C. (2021a). How and Why the Cerebellum Recodes Input Signals: An
908 Alternative to Machine Learning. Neuroscientist DOI: 10.1177/1073858420986795.
909 Gilbert, M., and Rasmussen, A. (2024). Gap Junctions May Have A Computational Function In
910 The Cerebellum: A Hypothesis. The Cerebellum.
911 Giovannucci, A., Badura, A., Deverett, B., Najafi, F., Pereira, T.D., Gao, Z., Ozden, I., Kloth, A.D.,
912 Pnevmatikakis, E., Paninski, L., *et al.* (2017). Cerebellar granule cells acquire a widespread
913 predictive feedback signal during motor learning. Nature neuroscience.

914 Hámori, J. (1992). Anatomy and neurochemical anatomy of the cerebellum. In Cerebellar
915 degenerations : clinical neurobiology, A. Plaitakis, ed. (Kluwer Academic).
916 Hámori, J., and Szentágothai, J. (1966). Participation of Golgi neuron processes in the
917 cerebellar glomeruli: an electron microscope study. Experimental brain research 2, 35-48.
918 Hartmann, M.J., and Bower, J.M. (2001). Tactile responses in the granule cell layer of
919 cerebellar folium crus IIa of freely behaving rats. The Journal of neuroscience : the official
920 journal of the Society for Neuroscience 21, 3549-3563.
921 Harvey, R.J., and Napper, R.M. (1988). Quantitative study of granule and Purkinje cells in the
922 cerebellar cortex of the rat. The Journal of comparative neurology 274, 151-157.
923 Harvey, R.J., and Napper, R.M. (1991). Quantitative studies on the mammalian cerebellum.
924 Progress in neurobiology 36, 437-463.
925 Herzfeld, D.J., Kojima, Y., Soetedjo, R., and Shadmehr, R. (2015). Encoding of action by the
926 Purkinje cells of the cerebellum. Nature 526, 439-442.
927 Holtzman, T., Rajapaksa, T., Mostofi, A., and Edgley, S.A. (2006). Different responses of rat
928 cerebellar Purkinje cells and Golgi cells evoked by widespread convergent sensory inputs. The
929 Journal of physiology 574, 491-507.
930 Hong, S., Negrello, M., Junker, M., Smilgin, A., Thier, P., and De Schutter, E. (2016). Multiplexed
931 coding by cerebellar Purkinje neurons. eLife 5.
932 Hull, C., and Regehr, W.G. (2012). Identification of an inhibitory circuit that regulates cerebellar
933 Golgi cell activity. Neuron 73, 149-158.
934 Isope, P., and Barbour, B. (2002). Properties of unitary granule cell-->Purkinje cell synapses in
935 adult rat cerebellar slices. The Journal of neuroscience : the official journal of the Society for
936 Neuroscience 22, 9668-9678.
937 Jakab, R.L., and Hámori, J. (1988). Quantitative morphology and synaptology of cerebellar
938 glomeruli in the rat. Anatomy and embryology 179, 81-88.
939 Jelitai, M., Puggioni, P., Ishikawa, T., Rinaldi, A., and Duguid, I. (2016). Dendritic excitation-
940 inhibition balance shapes cerebellar output during motor behaviour. Nature communications
941 7, 13722.
942 Jörntell, H., and Ekerot, C.F. (2006). Properties of somatosensory synaptic integration in
943 cerebellar granule cells in vivo. The Journal of neuroscience : the official journal of the Society
944 for Neuroscience 26, 11786-11797.
945 Kanichay, R.T., and Silver, R.A. (2008). Synaptic and cellular properties of the feedforward
946 inhibitory circuit within the input layer of the cerebellar cortex. The Journal of neuroscience :
947 the official journal of the Society for Neuroscience 28, 8955-8967.
948 Lange, W. (1974). Regional differences in the distribution of golgi cells in the cerebellar cortex
949 of man and some other mammals. Cell and tissue research 153, 219-226.
950 Mapelli, L., Solinas, S., and D'Angelo, E. (2014). Integration and regulation of glomerular
951 inhibition in the cerebellar granular layer circuit. Frontiers in cellular neuroscience 8, 55.
952 Miles, F.A., Fuller, J.H., Braitman, D.J., and Dow, B.M. (1980). Long-term adaptive changes in
953 primate vestibuloocular reflex. III. Electrophysiological observations in flocculus of normal
954 monkeys. Journal of neurophysiology 43, 1437-1476.
955 Mitchell, S.J., and Silver, R.A. (2000a). GABA spillover from single inhibitory axons suppresses
956 low-frequency excitatory transmission at the cerebellar glomerulus. The Journal of
957 neuroscience : the official journal of the Society for Neuroscience 20, 8651-8658.

958 Mitchell, S.J., and Silver, R.A. (2000b). Glutamate spillover suppresses inhibition by activating
959 presynaptic mGluRs. Nature 404, 498-502.
960 Moore, J.W., and Blazis, D.E.J. (1989). 11 - Simulation of a Classically Conditioned Response: A
961 Cerebellar Neural Network Implementation of the Sutton–Barto–Desmond Model. In Neural
962 Models of Plasticity, J.H. Byrne, and W.O. Berry, eds. (Academic Press), pp. 187-207.
963 Mugnaini, E. (1983). The length of cerebellar parallel fibers in chicken and rhesus monkey. The
964 Journal of comparative neurology 220, 7-15.
965 Napper, R.M., and Harvey, R.J. (1988a). Number of parallel fiber synapses on an individual
966 Purkinje cell in the cerebellum of the rat. The Journal of comparative neurology 274, 168-177.
967 Napper, R.M., and Harvey, R.J. (1988b). Quantitative study of the Purkinje cell dendritic spines
968 in the rat cerebellum. The Journal of comparative neurology 274, 158-167.
969 Nusser, Z., Sieghart, W., and Somogyi, P. (1998). Segregation of different GABAA receptors to
970 synaptic and extrasynaptic membranes of cerebellar granule cells. The Journal of neuroscience
971 : the official journal of the Society for Neuroscience 18, 1693-1703.
972 Oscarsson, O. (1979). Functional units of the cerebellum - sagittal zones and microzones.
973 Trends in Neurosciences 2, 143-145.
974 Ozden, I., Sullivan, M.R., Lee, H.M., and Wang, S.S. (2009). Reliable coding emerges from
975 coactivation of climbing fibers in microbands of cerebellar Purkinje neurons. The Journal of
976 neuroscience : the official journal of the Society for Neuroscience 29, 10463-10473.
977 Palay, S.L., and Chan-Palay, V. (1974). Cerebellar cortex : cytology and organization (Berlin:
978 Springer).
979 Park, S.M., Tara, E., and Khodakhah, K. (2012). Efficient generation of reciprocal signals by
980 inhibition. Journal of neurophysiology 107, 2453-2462.
981 Pietrajtis, K., and Dieudonné, S. (2013). Golgi Neurons. In Handbook of the Cerebellum and
982 Cerebellar Disorders, M. Manto, J.D. Schmahmann, F. Rossi, D.L. Gruol, and N. Koibuchi, eds.
983 (Dordrecht: Springer Netherlands), pp. 829-852.
984 Rancz, E.A., Ishikawa, T., Duguid, I., Chadderton, P., Mahon, S., and Hausser, M. (2007). High-
985 fidelity transmission of sensory information by single cerebellar mossy fibre boutons. Nature
986 450, 1245-1248.
987 Rasmussen, A., Zucca, R., Jirenhed, D.A., Johansson, F., Ortenblad, C., Svensson, P., and
988 Hesslow, G. (2014). Golgi cell activity during eyeblink conditioning in decerebrate ferrets.
989 Cerebellum (London, England) 13, 42-45.
990 Raymond, J.L., and Medina, J.F. (2018). Computational Principles of Supervised Learning in the
991 Cerebellum. Annual review of neuroscience 41, 233-253.
992 Ritzau-Jost, A., Delvendahl, I., Rings, A., Byczkowicz, N., Harada, H., Shigemoto, R., Hirrlinger, J.,
993 Eilers, J., and Hallermann, S. (2014). Ultrafast action potentials mediate kilohertz signaling at a
994 central synapse. Neuron 84, 152-163.
995 Ros, H., Sachdev, R.N., Yu, Y., Sestan, N., and McCormick, D.A. (2009). Neocortical networks
996 entrain neuronal circuits in cerebellar cortex. The Journal of neuroscience : the official journal
997 of the Society for Neuroscience 29, 10309-10320.
998 Rossi, D.J., and Hamann, M. (1998). Spillover-mediated transmission at inhibitory synapses
999 promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry. Neuron
1000 20, 783-795.
1001 Rossi, D.J., Hamann, M., and Attwell, D. (2003). Multiple modes of GABAergic inhibition of rat
1002 cerebellar granule cells. The Journal of physiology 548, 97-110.

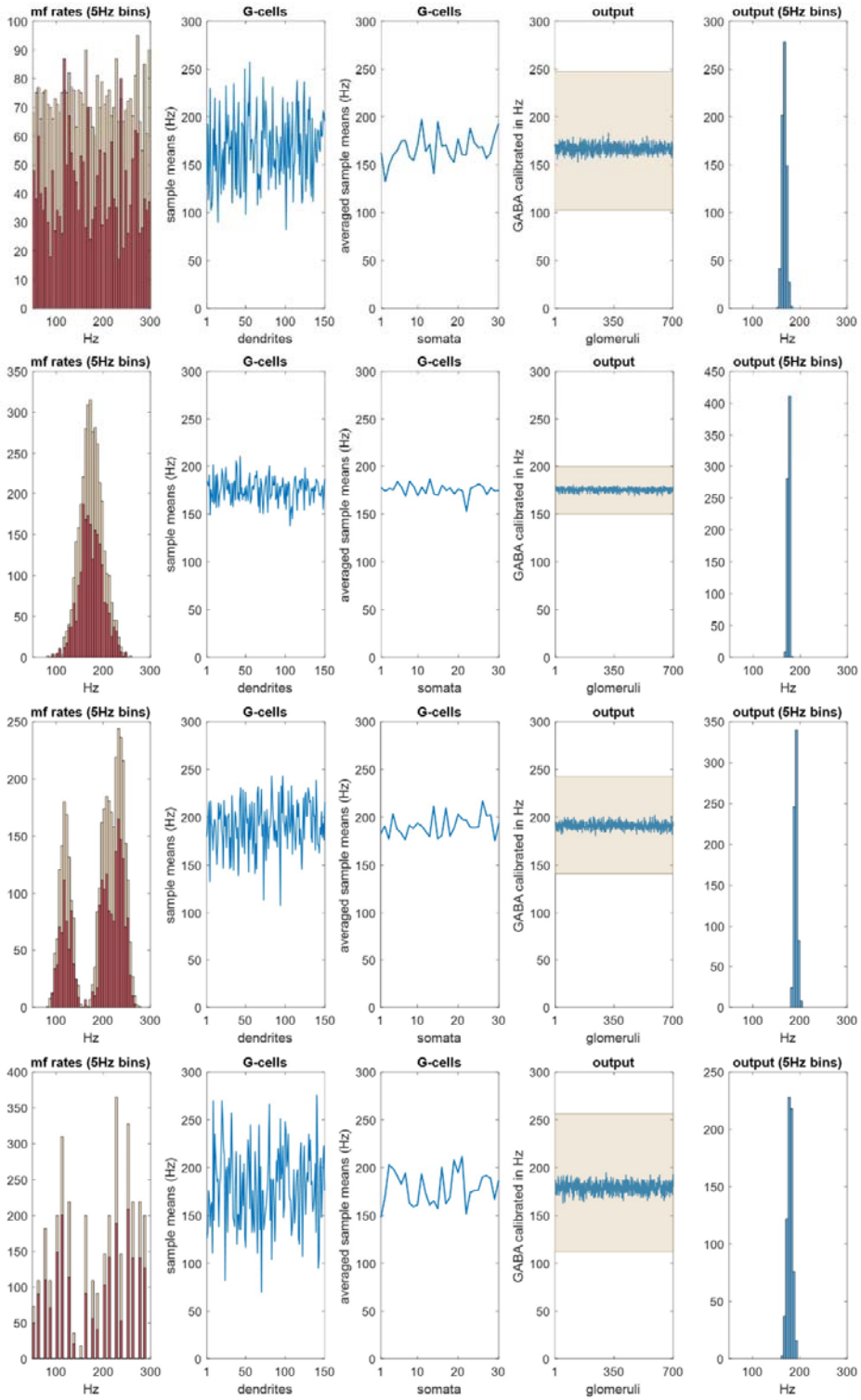
1003 Ruigrok, T.J., Hensbroek, R.A., and Simpson, J.I. (2011). Spontaneous activity signatures of
1004 morphologically identified interneurons in the vestibulocerebellum. The Journal of
1005 neuroscience : the official journal of the Society for Neuroscience 31, 712-724.
1006 Sargent, P.B., Saviane, C., Nielsen, T.A., DiGregorio, D.A., and Silver, R.A. (2005). Rapid vesicular
1007 release, quantal variability, and spillover contribute to the precision and reliability of
1008 transmission at a glomerular synapse. The Journal of neuroscience : the official journal of the
1009 Society for Neuroscience 25, 8173-8187.
1010 Saviane, C., and Silver, R.A. (2006). Fast vesicle reloading and a large pool sustain high
1011 bandwidth transmission at a central synapse. Nature 439, 983-987.
1012 Schonewille, M., Khosrovani, S., Winkelman, B.H., Hoebeek, F.E., De Jeu, M.T., Larsen, I.M.,
1013 Van der Burg, J., Schmolesky, M.T., Frens, M.A., and De Zeeuw, C.I. (2006). Purkinje cells in
1014 awake behaving animals operate at the upstate membrane potential. Nature neuroscience 9,
1015 459-461; author reply 461.
1016 Sherman, S.M., and Guillery, R.W. (2011). Distinct functions for direct and transthalamic
1017 corticocortical connections. Journal of neurophysiology 106, 1068-1077.
1018 Shinoda, Y., and Sugihara, I. (2013). Axonal Trajectories of Single Climbing and Mossy Fiber
1019 Neurons in the Cerebellar Cortex and Nucleus. In Handbook of the Cerebellum and Cerebellar
1020 Disorders, S.J.D. Manto M., Rossi F., Gruol D.L., Koibuchi N., ed. (Springer, Dordrecht).
1021 Shinoda, Y., Sugihara, I., Wu, H.S., and Sugiuchi, Y. (2000). The entire trajectory of single
1022 climbing and mossy fibers in the cerebellar nuclei and cortex. Progress in brain research 124,
1023 173-186.
1024 Sultan, F., and Heck, D. (2003). Detection of sequences in the cerebellar cortex: numerical
1025 estimate of the possible number of tidal-wave inducing sequences represented. Journal of
1026 physiology, Paris 97, 591-600.
1027 Turecek, J., Jackman, S.L., and Regehr, W.G. (2016). Synaptic Specializations Support
1028 Frequency-Independent Purkinje Cell Output from the Cerebellar Cortex. Cell reports 17, 12.
1029 Turecek, J., Jackman, S.L., and Regehr, W.G. (2017). Synaptotagmin 7 confers frequency
1030 invariance onto specialized depressing synapses. Nature 551, 503.
1031 van Kan, P.L., Gibson, A.R., and Houk, J.C. (1993). Movement-related inputs to intermediate
1032 cerebellum of the monkey. Journal of neurophysiology 69, 74-94.
1033 Vervaeke, K., Lorincz, A., Gleeson, P., Farinella, M., Nusser, Z., and Silver, R.A. (2010). Rapid
1034 desynchronization of an electrically coupled interneuron network with sparse excitatory
1035 synaptic input. Neuron 67, 435-451.
1036 Vos, B.P., Volny-Luraghi, A., and De Schutter, E. (1999). Cerebellar Golgi cells in the rat:
1037 receptive fields and timing of responses to facial stimulation. The European journal of
1038 neuroscience 11, 2621-2634.
1039 Walter, J.T., and Khodakhah, K. (2006). The linear computational algorithm of cerebellar
1040 Purkinje cells. The Journal of neuroscience : the official journal of the Society for Neuroscience
1041 26, 12861-12872.
1042 Walter, J.T., and Khodakhah, K. (2009). The advantages of linear information processing for
1043 cerebellar computation. Proceedings of the National Academy of Sciences of the United States
1044 of America 106, 4471-4476.
1045 Wu, H.S., Sugihara, I., and Shinoda, Y. (1999). Projection patterns of single mossy fibers
1046 originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei. The Journal
1047 of comparative neurology 411, 97-118.


A
B
C
D
Molecular layer
Granular layer
Mossy fibre
Cerebellar y-axis [µm]
Cerebellar x-axis [µm]
6000
5000
4000
3000
2000
1000
0
-1000
-500
0
500

1048

1049 **Fig 1. Mossy fibre termination. A** Sagittal section of the rat cerebellum (obtained by
1050 multi-photon microscopy, courtesy Tom Deerinck). The outer layer of the cerebellar
1051 cortex, the molecular layer, is stained green and the inner layer, the granular layer, is
1052 blue. **B** Output of a procedure used to group mossy fibre terminals into clusters (Sultan
1053 and Heck, 2003 fig 6), reproduced with permission. All terminals belong to one mossy
1054 fibre, which terminates in two strips. Clusters are shown in different shades of grey. **C**
1055 Schematic of an area of the folded cerebellar cortex. The strip inside the red dotted
1056 line is the width and depth of a microzone but only part of its length. Termination of
1057 mossy fibre signals is not at a point but at several randomly spaced, sagittally-aligned
1058 locations. Yellow zones are terminal locations of a single mossy fibre. Size of locations
1059 varies depending partly on the number of terminals in each cluster. Pink zones are
1060 terminal locations of a second mossy fibre. In reality, zones overlap extensively and
1061 terminals are intermingled. A mossy fibre mediolateral axon may give rise to several
1062 sagittal branches, of which here only one is shown. The zone enclosed by a blue line
1063 has the dimensions of a 'field'. A field is a nominal division of the granular layer with
1064 the average dimensions of a mossy fibre terminal cluster – 200 µm x 150 µm – which
1065 we use to quantify granular layer computations. A strip of the granular layer with the
1066 dimensions of a sagittal row of 75–100 fields is a 'rank'. Viewed from the cerebellar
1067 surface, a rank has the dimensions of a mid-range estimate of a microzone, but its long
1068 sides are less cleanly defined. **D** Section of a microzone viewed obliquely from the
1069 cerebellar surface, to illustrate the size and random intermingling of mossy fibre
1070 terminals. The molecular layer is represented by squares of about the dimensions of a
1071 Purkinje cell territory. The wire frame boxes are an exploded view of the orange fields.
1072 The right box shows randomly distributed terminals of 5 mossy fibres (each a different
1073 colour). We estimate that ~180 mossy fibres contribute terminals to a field, and a field
1074 contains a total of around 700 terminals. Proportions of terminal size to field size are
1075 preserved. The left box shows terminals of 20 mossy fibres.

1076


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

1077

1078 **Fig 2. Golgi cell ensemble conversion. Column 1** We randomly generated (3,645)
1079 values representing mossy fibre rates with a uniform, normal, twin-peaked or
1080 discontinuous distribution, simulating input to a row of 100 fields. Rate data were
1081 randomly sampled to derive input rates to each field, and copies added to reflect the
1082 fact that each afferent cell contributes a randomly variable number of terminals. Red
1083 data represent copies-included signals received by a row of three fields. See Methods
1084 for derivation of rates. **Column 2** The first-column red data were randomly sampled
1085 150 times, sample size 4, representing mossy fibre input to a 3-field population of 30
1086 Golgi cells, each with 5 basal dendrites (the mean) which each receive contact from 4
1087 mossy fibres. Data are the sample means, representing dendritic depolarisation.
1088 **Column 3** We next randomly sampled column 2 data *without* replacement 30 times,
1089 sample size 5, and took the sample means, representing somatic integration of
1090 dendritic signals. **Column 4** Column 3 data were randomly sampled 700 times, sample
1091 size 8–12, reflecting convergent input from 3 fields of Golgi cells onto 700 glomeruli in
1092 the middle field (which receives the output of a three-field ensemble), and took the
1093 sample means, representing GABA spillover that is a linear function of the mean of
1094 afferent rates. Inhibition of granule cells is almost exclusively mediated by spillover.
1095 **Column 5** The column 4 data are presented as a histogram. **All rows** Column 1 mossy
1096 fibre data are converted to narrowly-focussed inhibition of granule cells. The mean of
1097 the column 4/5 data is linearly related to the mean of the mossy fibre distribution. This
1098 result is independent of the shape of the original distribution.

1099

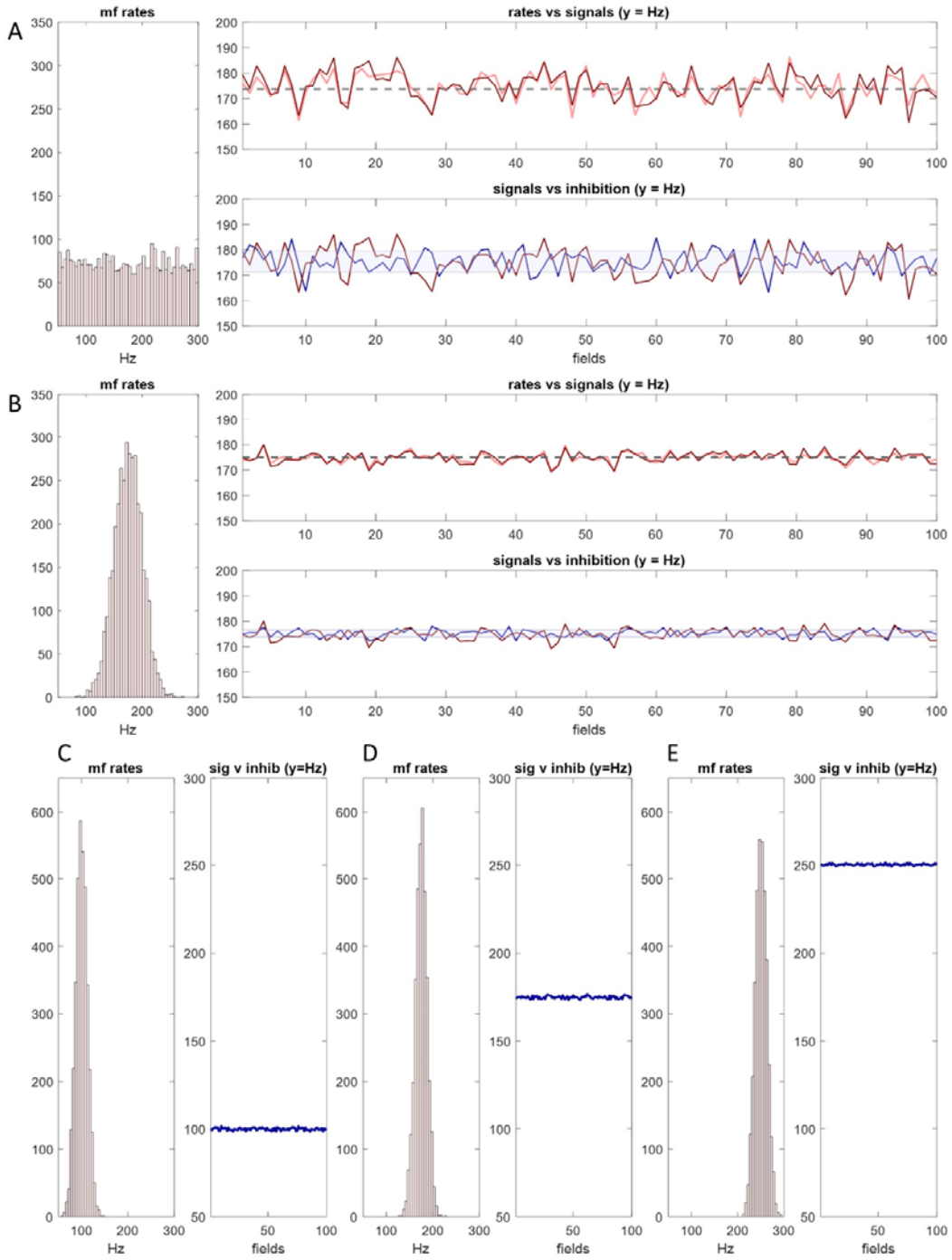


Fig 3. The Golgi cell population of a rank functions as a single unit. A Simulation of Golgi cells in a sagittal row of 100 fields, taking the mossy fibre distribution on the left as input (see Fig 2 and Methods for derivation of mossy fibre rates). **Top right**: pink:

the mean of mossy fibre input rates to each field; red: the mean with duplicate signals added; dashed grey line: the mean of input rates to the whole rank. **Bottom right**: red: same as top right; blue: mean, for each field, of glomerulus data derived in the same way as Figure 2 columns 4 and 5 (the 'inhibitory field mean'); pale blue: SD of glomerulus data. **B** Same as (A) but sampling a different mossy fibre distribution.

A and B

Red data closely track the pink data in top right panels, indicating that duplicate signals (because the mossy fibre axon ends in a cluster of terminals) do not skew the mean of firing rates received by a field.

Blue data across all fields are in a narrowly circumscribed range. 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.

Inhibitory field means are tightly grouped: inhibition of granule cells is synchronised in a rank.

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, unpredictability inherent in the mechanism (because the outcome of random sampling is inherently uncertain) 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).

C The same simulation. Left: the sampled distribution (mossy fibre input rates to a rank); right: the inhibitory field means (we use the means to compare fields). The y axis of the right panel has the same range as the x axis of the left panel. Panels are tall to improve visual comparison to other rows, where the y axis has a smaller range. **D** and **E** The same, moving the sampled distribution to the right. Inhibition of granule cells is precisely and linearly coupled to the mean of mossy fibre input rates to the whole rank.

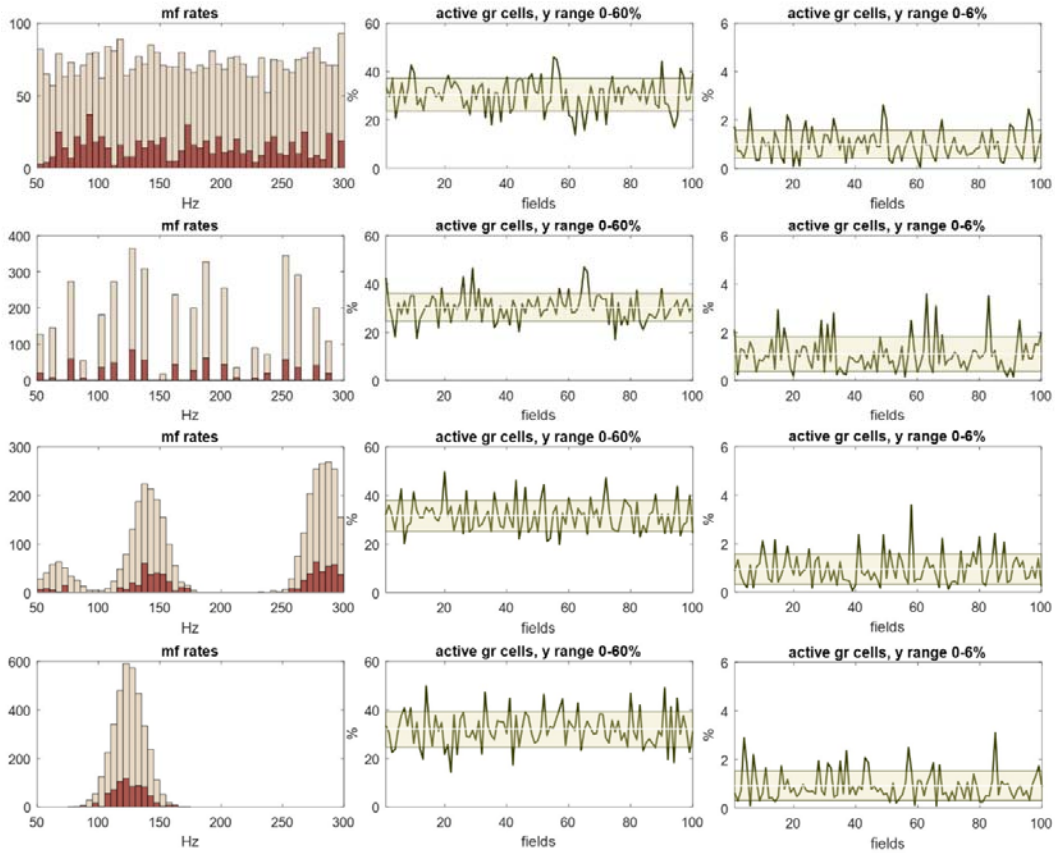
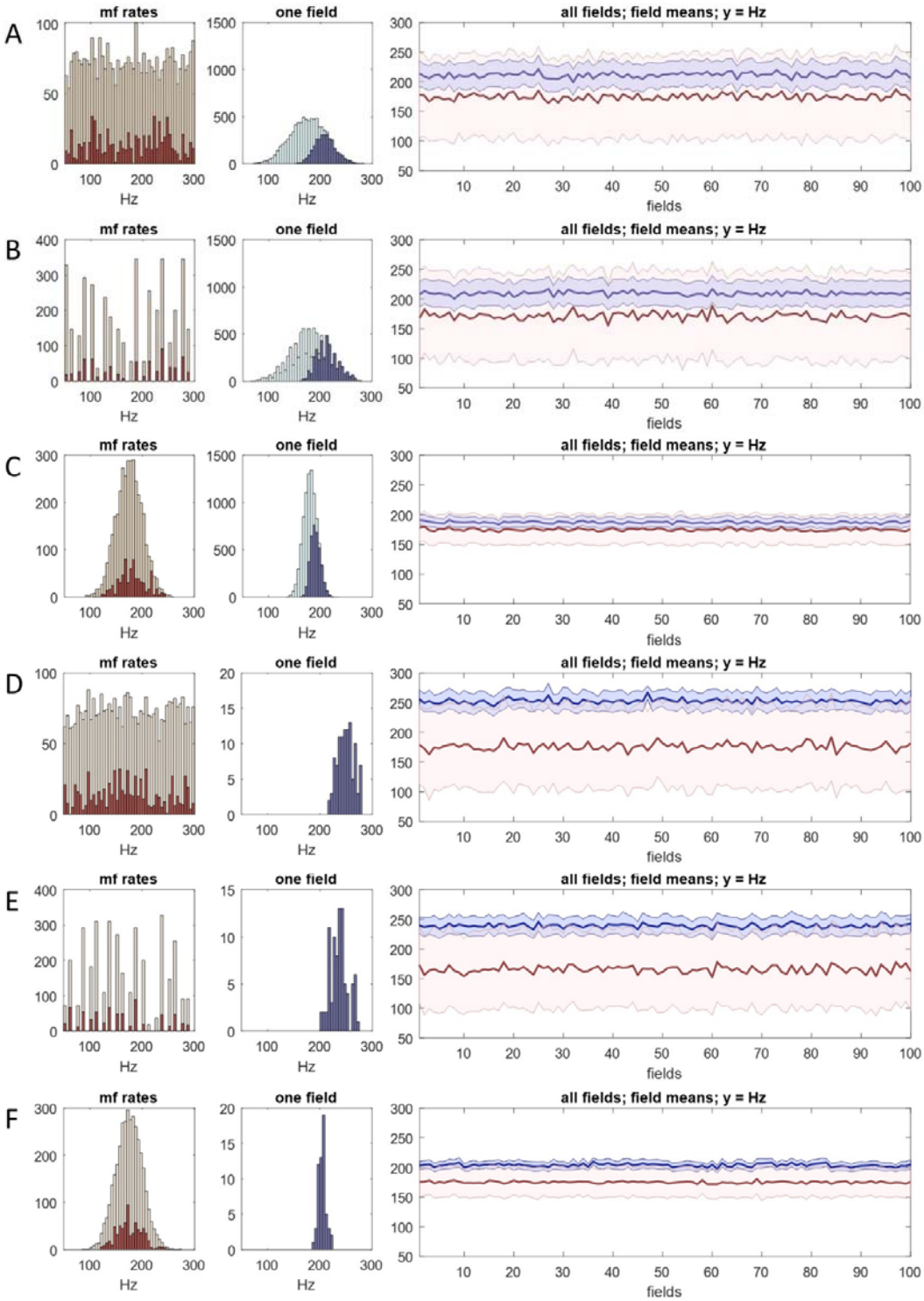


Fig 4. Regulation of granule cell activity by the feed-forward pathway (mossy fibres→Golgi cells→granule cells) We incorporated the Golgi cell ensemble conversion into a simulation of the conversion of mossy fibre signals into granule cell signals (recoding). One of the outputs of recoding is regulation of the percentage of active granule cells. We tested if the recoding computation causes the percentage to be stable, i.e., in the same narrow range in all fields regardless of 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. **Column 3** The same, with an adjustment (a co-efficient) to sparsen granule cell activity. y axes have 1/10 the range of column 2. The percentage is confined to a narrow range. This result is independent of the sampled (i.e., mossy fibre) distribution, suggesting that regulation is

independent of fluctuation of mossy fibre rates. Nonetheless, field-to-field variation relative to the mean is significant. Homeostatic local fine tuning may be provided physiologically by feedback via granule cell ascending axon contact on Golgi cells, not represented in the simulation.


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

1156

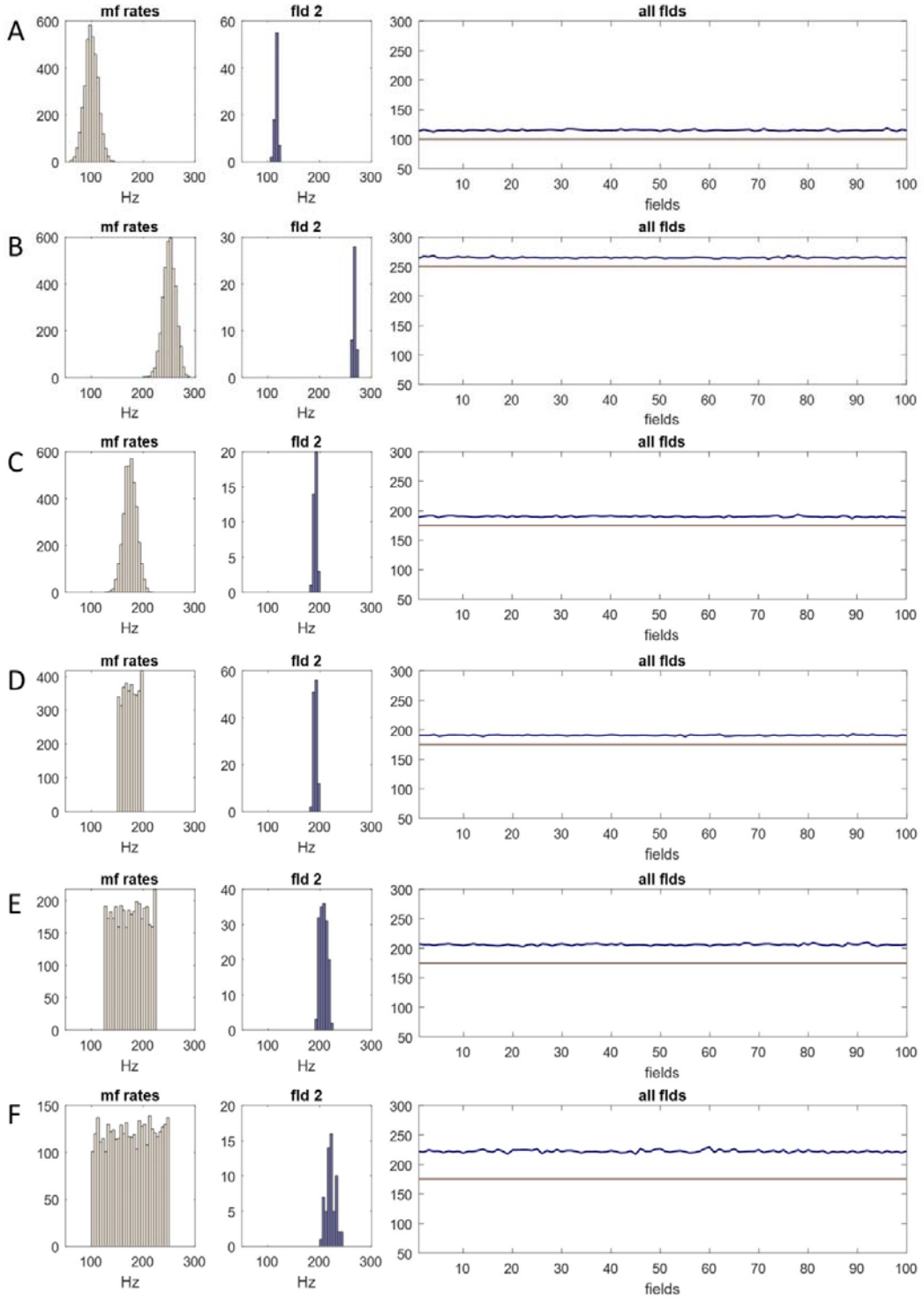
1157 **Fig 5. A rank is a single recoding unit.** The same simulation as Figure 4 but data are
1158 granule cell firing rates or derived from granule cell firing rates. **Column 1** We tested
1159 the effect of changing the shape of the mossy fibre distribution (without changing the
1160 number of signals). **Column 2** Pale green: sample means – the mean of input rates to
1161 each granule cell in a single field selected at random; blue: sample means of the subset
1162 of granule cells which meet the conditions to fire. In rows A–C, ~30% fire. In rows D–F,
1163 ~1% fire, y axes are rescaled and green data are omitted. **Column 3** Red: mean of
1164 mossy fibre input rates to each field in a rank, including duplicate signals; pink: SD of
1165 mossy fibre input rates to each field; blue: field means (i.e., the mean of granule cell
1166 firing rates in each field); pale blue: SD of granule cell firing rates in each field.

1167 **Observations**

1168 Regardless of the shape and large range of the mossy fibre distribution, the field
1169 means always converge towards a straight line which is a fixed distance from the
1170 overall mean of mossy fibre rates.

1171 The amount and direction of field-to-field random fluctuation of the blue data is
1172 independent of local fluctuation of the red data, and in a smaller range. Uncertainty
1173 inherent in random sampling – first by single fields of input rates to a rank, and then by
1174 Golgi cells and granule cells of input rates to a field – is not compounded.

1175



1176

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.