WHAT IF…? COMPUTATIONAL ANATOMY IN THE CEREBELLAR GRANULAR LAYER

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

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

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

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

Keywords: theory, hypothesis, neuroscience, cerebellum, granular, network, layer, code.

ABSTRACT

The cerebellum is made up of repeating, anatomically overlapping networks. Neural modelling has a tradition of learning algorithms that are implemented physiologically by learned changes in synaptic transmission strength. Here we propose a reinterpretation of the evidence with regard to information processing in 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 any input the same way, and it provides an otherwise elusive mechanism to coordinate the firing of microzone-grouped Purkinje cells downstream, which control the output cells of the cerebellar network. Since we argue that anatomy is the chief instrument of the granular layer computation, we can replicate it *in silico*, down to local anatomical noise, so that we are able to use a simulation to test if the ideas are computationally viable.

1. Introduction

It is thought that neuronal firing rates provide the physiological form of numerical values that are used to code information, and that neural networks provide the physiological implementation of computations which take firing rates as inputs. In the traditional cerebellar model, the supervised learning model (Raymond & Medina, 2018, for an excellent review), control of Purkinje cells by granule cell rates is displaced by learning coded as parallel fibre synaptic weight changes (Albus, 1989; Brunel et al, 2004; Dean et al, 2010; Fujita, 1982). After training, when a Purkinje cell receives a remembered pattern, it fires at a learned rate. It is implicit that a granule cell code is redundant because control is surrendered to learning.

However, the evidence is inconclusive. Experimental work has not provided conclusive support for the expected role of long-term depression in motor learning (Schonewille et al, 2011; ten Brinke et al, 2015; Welsh et al, 2005) or for finely-graded modification of synaptic weights predicted by learning models (Ekerot & Jörntell, 2001; Isope & Barbour, 2002). There is accumulating evidence of linear transmission of rate-coded data in the cerebellum (Duguid et al, 2015; Duguid et al, 2012; Jelitai et al, 2016; Rancz et al, 2007; Ritzau-Jost et al, 2014; Walter & Khodakhah, 2006; 2009), and of contact between cells that is adapted for faithful transmission of rate information (Cesana et al, 2010; Delvendahl & Hallermann, 2016; Moore & Blazis, 1989; Saviane & Silver, 2006). There is also evidence that cells fire at a rate that is a linear function of

behavioural and task metrics ‘at all levels of the cerebellar circuit’ (Raymond & Medina, 2018 p.239, citing 30 sources).

The main source of glutamatergic input to the cerebellum is from mossy fibres which terminate on granule cells in the inner layer of the cerebellar cortex, the granular layer. Mossy fibres also contact Golgi cells, which inhibit granule cells. The granule cell axon rises into the outer layer, the molecular layer, where it divides in two to form parallel fibres. Parallel fibres run parallel to each other and the cerebellar surface and contact Purkinje cells. Purkinje cells carry the only output of the cerebellar cortex.

We propose an alternative. We argue that mossy fibre input to the cerebellum, received in the granular layer, is functionally compartmentalised, broadly mirroring microzonal organisation of the molecular layer. Granule cells code information in their collective activity at this scale. High resolution mossy fibre input is converted to a low-resolution granule cell code. Compartment dimensions are the size of a minimum unit of internally coded information, and in that sense the internal code is pixellated. There is no finer resolution. High definition of mossy fibre data is sacrificed and once sacrificed is not recovered downstream, counter to the general expectation that detail increases the power of brain computations. To make that argument, we present a case that there is a computational effect, unaided, of anatomy. Cell morphologies and the scaffold-like architecture of the cerebellar cortex are adapted for this function.

An important consequence of code in this form, spanning a network, is that the entire Purkinje cell population of a microzone downstream receives, at the same time, the same code (notwithstanding that almost all of them receive none of the same signals).

The case in evidence for these proposals is mainly but not exclusively a combination of anatomy and linear transmission. Anatomy is well established and uncontentious (but not all of it is well known). The evidence is not new. Rather we propose a novel reinterpretation of the evidence. The evidence of linear transmission is still relatively sparse and more contentious, but growing. Each step in the argument has a section on evidence. At each stage we present our interpretation of the evidence (the hypothesis) and then quantify and test the ideas by modelling (providing test results). As the chief instrument of computations is anatomy, model parameters and the values they take (usually an estimated range) are available from the literature, down to anatomical variability of single cells.

A key part of the proposals is that the cerebellum exploits statistical effects. We can therefore use simulations to test the proposals – not that they *could* work, since that is obvious, but that cell morphologies, numbers and connectivity are a plausible substrate. Rather than having all the results in one section, we present results and interpretations in each section.

2. Mossy fibre anatomy is an anatomical solution to a computational problem

2.1 Hypothesis

Mossy fibre anatomy has important computational significance. It approximates the result of independent random sampling. The sampled population is a snapshot of the firing rates of mossy fibres which innervate a long, thin strip of the granular layer of the cerebellar cortex. A sample is a subset of rates received by a cell that receives contact from mossy fibres. Sampling is by Golgi cells and granule cells which populate the strip. We propose that the cerebellum exploits statistical effects called the central limit theorem. For that, strictly speaking samples must be independent, also termed sampling 'with replacement' (in the example often used to teach it, coloured balls are put back in the bag between samplings). Simultaneous sampling with replacement is not possible. However, mossy fibres simulate independence by terminal branching and termination of each branch in a cluster of terminals, we suggest.

2.2 Evidence

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 & Sugihara, 2013; Shinoda et al, 2000; Sultan & 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 & Heck, 2003; Wu et

114 al, 1999) measured from edge to edge of clusters (so they are separated by a greater
115 distance measured from centre to centre). Terminals themselves are intimately and
116 randomly intermingled with terminals of other mossy fibres.

117 Accordingly, mossy fibres do not terminate at a point but a region with dimensions:
118 the field occupied by a terminal cluster. A single mossy fibre gives rise to several
119 terminal branches and each of those to several terminals, such that a sagittal strip of
120 the granular layer receives multiple copies of every signal at each of multiple locations.
121 Cluster size is variable (because it depends on the number of terminals and how tightly
122 clustered they are), averaging 200 µm sagittally x 150 µm mediolaterally (Sultan &
123 Heck, 2003). We term the average dimensions of a cluster a 'field'. The granular layer
124 is anatomically seamless. A field is a nominal division that we use for modelling. A
125 sagittal row of 75–100 fields laid end to end has the dimensions viewed from the
126 cerebellar surface of a mid-range estimate of microzone dimensions (Dean et al, 2010;
127 Garwicz et al, 1998a; Oscarsson, 1979; Ozden et al, 2009). We term a row of fields a
128 'rank' (a 'strip' in the hypothesis section).

129 *Figure 1 about here*

130

131

132

2.3 Analysis

2.3.1 Terminal branching

Here, the sampled population is the population of mossy fibres that innervate a rank and a sample is the subpopulation that innervate a field. 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 therefore the number of terminal branches per mossy fibre had the same probability distribution, sampling by fields would approximate sampling with replacement. The practical result of terminal branching is a facsimile of independence.

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 on this is 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 & Heck, 2003). Estimated by this method, the mean is 4–5. Put another way, the anatomical cap on the number of fields that can receive a signal (set by the number of terminal branches of that cell) simulates the distributed probability of repeat samplings with actual replacement. The simulation reflects anatomy and anatomy simulates replacement.

2.3.2 Terminal clusters

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. Here, then, we have the same problem: if a branch ended in a single terminal, contact on any particular granule cell would mean it was beyond the range of most of the rest of the population. Mossy fibre termination in terminal clusters, and random intermingling of terminals (Fig 1), removes that restriction. Strictly speaking, this is not replacement and samples are not independent. But in practice, to approximate independence, it is only necessary for the number of repeat selections of a signal to have the same probability distribution as with real replacement. Terminal clustering has that result: anatomy mimics replacement.

A practical problem with the strategy of clustering is the risk that a random sample of terminals will contain two that arise from the same mossy fibre. 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 nearer 200 (Supplementary Materials, SM1).

2.3.3 Joint effect

As each field receives, at any given instant in time, a quasi-independent random sample of mossy fibre input rates to a rank, and efferent cells individually receive a quasi-independent random sample of input rates to a field, the joint effect is a biological facsimile of simultaneous random sampling with replacement by single cells of input rates to a whole rank. This is not a discrete event. If and to the extent that mossy fibres are continuously active during behaviour, or many of them are active at any time, it occurs continuously.

3. Golgi cells compute mossy fibre rates by working together; a long thin strip of Golgi cells is a single functional unit

3.1 Hypothesis

We propose that Golgi cells are organised into functional groups. The minimum size of a functional group is a sagittal row of three fields: an ensemble. Ensembles perform a computational conversion of mossy fibre rates into inhibition of granule cells. The effect of the computation is to synchronise inhibition at a strength that tracks the mean of mossy fibre rates. Mossy fibre terminals are ensheathed in a semi-permeable membrane, restricting diffusion of neurotransmitter spillover, a structure termed a glomerulus. We propose that there are two rounds of random sampling by an ensemble, sampling by Golgi cell basal dendrites of mossy fibre firing rates and sampling by glomeruli of Golgi cell firing rates. Golgi cell basal dendritic signalling is a sustained (and not discrete) event. Polarisation of a dendrite is, at any time, a straightforward function of the mean of afferent rates. Somatic charge – also a sustained state – reflects the mean of dendritic states. This generates a new data set with a narrower range – Golgi cell firing rates – which at a collective level is predicted by mossy fibre rates. This is in turn itself randomly sampled by glomeruli. Inhibition of granule cells is by intraglomerular spillover at a concentration that reflects the mean of afferent Golgi cell rates, generating a third data set, now with a narrowly focussed range and still reliably predicted by mossy fibre rates. Ensembles are neither

anatomically nor functionally discrete but conflate sagittally. Quasi-independent random sampling (by ensembles of mossy fibre rates) means that 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.

3.2 Evidence

3.2.1 Golgi cell dendritic signalling and somatic integration

It is part of our proposal that Golgi cell basal dendritic signals linearly reflect mossy fibre rates, so that mossy fibre rate information is conserved proportionally in Golgi cell firing rates. Discrete signals have the problem that data are corrupted by noise (interference by spatial and temporal integration that codes nothing). In theory, and we propose, this problem is overcome if signalling is a sustained state. In a snapshot of activity, a dendritic state codes the mean of afferent rates. The somatic state codes the mean of dendritic states. We arrive at that proposition from this evidence.

Golgi cells are large interneurons whose cell bodies and basal dendrites lie in the granular layer. Mossy fibres directly contact Golgi cell basal dendrites. Golgi cell firing frequency increases linearly with the amplitude of depolarising current (Dieudonné, 1998 p.845). 'Sensory-evoked Golgi-cell inhibition [of a granule cell] scales proportionally with the level of mossy fiber excitatory synaptic input' to the same cell,

242 such that inhibition reliably conveys mossy fibre rate information (Duguid et al, 2015).
243 See Supplementary Materials for more references for a linear relationship of input and
244 output rates.

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

254 This means that fast rate-proportional dendritic depolarisation is passively signalled by
255 large dendrites with little filtering to the Golgi cell soma. During behaviour in freely-
256 moving animals, mossy fibre activity is dense: a high fraction of mossy fibres are active
257 (Hartmann & Bower, 2001). Both mossy fibres (Arenz et al, 2008; Giovannucci et al,
258 2017; van Kan et al, 1993) and Golgi cells (Edgley & Lidierth, 1987) have been reported
259 to fire with a sustained, time-varying signature. During movement, the 'output of Golgi
260 cells varies smoothly as a function of input....They are capable of modulating their
261 output to reflect their input with little noise or signal distortion' (Moore & Blazis,

1989). In the behaving animal, we suggest, Golgi cell basal dendrites are in a sustained state of continuously variable rate-proportional polarisation under modulation by sustained inputs. Charge influx is proportional to the sum of input rates and so a linear function of the mean of afferent rates. In the ensemble computation, this is the biological form of calculating the sample means.

As charge transfer to the soma is passive, polarisation of the soma is a sustained state that reflects the mean of dendritic states (the consequence of passive, high-fidelity, continuous dendritic signalling). If the mean afferent rate falls, charge may temporarily flow back into dendrites, aiding a fast somatic response.

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.

3.2.2 Granule cell inhibition reflects the mean of afferent Golgi cell rates

It forms part of our proposal that granule cell inhibition scales with the mean firing rate of Golgi cells afferent to a glomerulus. We reach that conclusion from this evidence.

The Golgi cell axon branches 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, and fills

the granular layer vertically (Barmack & Yakhnitsa, 2008). Fine, beaded axon fibres enter glomeruli, where they contact and inhibit granule cells (Hámori & Szentágothai, 1966). 'In the adult rat cerebellum, each granule cell dendrite receives 2.6 ± 0.55 synaptic contacts from Golgi axon terminals' (Rossi & Hamann, 1998) citing (Jakab & Hámori, 1988).[1] However, the large majority of inhibition (98%) is by spillover (Duguid et al, 2012), where neurotransmitter released into the synaptic cleft spills out into the glomerular space. Golgi cells release GABA, an inhibitory neurotransmitter. This is detected by high-affinity $GABA_A$ receptors located perisynaptically and extrasynaptically on granule cells (Brickley et al, 1996; Nusser et al, 1998). Even synaptically-received signals have most of their effect (i.e., the large majority of charge transfer is mediated) via spillover.[2]

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

[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. & Hamann, M. (1998) Spillover-mediated transmission at inhibitory synapses promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry. *Neuron*, 20(4), 783-95.

afferent to its location. Intraglomerular GABA concentration is the biological form of taking the sample means. Golgi cells provide the only inhibitory input to granule cells.

3.3 Analysis

The fidelity of the relationship between mossy fibre rates and inhibition of granule cells depends on the number of samples and sample size at each step. The evidence is anatomical and available from the literature. A review appears in Supplementary Materials, SM2.

3.3.1 The Golgi cell ensemble computation

Golgi cells form an anatomically undivided carpet that seems to populate the granular layer seamlessly. A Golgi cell ensemble is the population of a sagittal row of three fields, a functional division. An ensemble is defined as those cells whose axonal range includes the whole or a substantial part of the middle field (in all three dimensions).

Quasi-independence of the random sample of mossy fibres that make contact on each Golgi cell basal dendrite, coupled with linear transmission, has a computational effect. 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.

This describes, in principle, the ensemble computation: each Golgi cell randomly samples mossy fibre firing rates and Golgi cell firing rates are in turn randomly sampled by glomeruli. Quasi-independence of sampling by Golgi cell basal dendrites is discussed in section 2. Independence of sampling by glomeruli is because a glomerulus in the middle field can receive innervation by any permutation of Golgi cells in the afferent ensemble.

In the simulation, we represented mossy fibre rates as values that were randomly generated within constraints on the range and frequency distribution. Constraints are stated (or shown) in figures where the data are used. This is a key element of the simulation because it incorporates random anatomical variability of mossy fibres, essential to assay whether real anatomy plausibly provides a substrate for the computational effect we claim it has. We describe it in detail in Supplementary Materials, SM1. The population of mossy fibre values represents a moment in time, a snapshot of input to a rank. We then derive, from firing rates received by a rank, rates received by each field, and the number of copies of each signal, taking into account the random variation of the number of terminals per cluster and clusters per cell, and the fact that most clusters fall partly inside and partly outside a field (not straightforward).

335 This provides the number of signals, rates and the number of copies received by each
336 of three fields – the first-sampled population. The computation is then relatively
337 straightforward. Three fields contain a total of 30 Golgi cells. Mossy fibre terminals are
338 randomly sampled by Golgi cell basal dendrites. Golgi cells have 4–6 basal dendrites
339 which we take as 5, so that the total number of samples is 150, 50 per field. Sample
340 size is the number of mossy fibres which make contact on each dendrite, which we
341 estimate is 4. The postsynaptic result – charge influx to dendrites – is a proportional
342 function of input rates. As far as we know, there are no data for the transformation of
343 multiple mossy fibre input rates into Golgi cell basal dendritic depolarisation. We
344 assume depolarisation is a linear function of the mean of afferent rates which we take
345 to be the mean itself. Taking the sample means generates a new set of values (Fig 2
346 column 2) with a smaller range and centred on the mean of the sampled distribution.
347 Basal dendritic charge transfer is passive and little filtered (Cesana et al, 2010;
348 Kanichay & Silver, 2008). Integration of dendritic states at the Golgi soma is
349 represented by taking the mean of dendritic charge. This generates a third population
350 of values (Fig 2 column 3). Golgi cell firing rates have a linear relationship with the
351 amplitude of depolarising current (Dieudonné, 1998). Thus, rate-proportional dendritic
352 charge influx is linearly converted to Golgi cell firing rates.

353 The second-sampled population is the population of Golgi cell rates. Sampling is by
354 glomeruli. Inhibition of granule cells is almost entirely via spillover (Brickley et al, 1996;
355 Duguid et al, 2012; Nusser et al, 1998), effectively increasing sample size, as sample

size is not the number of Golgi cells that make direct contact on any particular granule cell, but the number afferent to a glomerulus. Sample size is accordingly the number of Golgi cells with convergent input to a glomerulus (taken as 8 to 12), and the number of samples is the number of glomeruli per field (700; see Supplementary Materials SM1). Spillover is a physiological facsimile of averaging. We represented spillover by taking the sample means (the mean of Golgi cell input rates to each glomerulus), glomerulus by glomerulus. This generates a final population of values (Fig 2 columns 4 and 5), the output of the ensemble conversion.

Physiological output of the ensemble computation is inhibition of middle-field granule cells. More exactly, there are 700 outputs, the number of glomeruli per field. Mossy fibre input rates to three fields are converted to tightly-grouped outputs of the ensemble computation centred on the mossy fibre mean. The same result is obtained regardless of the shape and range of the frequency distribution of mossy fibre rates. Even a wide mossy fibre range is converted to tightly-grouped outputs (although narrowing the range of mossy fibre rates generates more narrowly-grouped outputs).

Our interest is confined to the computational effect of ensemble architecture through the steps of the conversion. Data generated by the model illustrate the relationship of the steps on a normalised scale (if we are correct about the mechanism). For this purpose, measuring data at different steps on different scales would be arbitrary.

Figure 2 about here please

3.3.2 Golgi cells in a long row of fields function as a single unit

We next repeated the ensemble conversion to derive inhibition of all fields in a sagittal row of 100 fields, a rank. As ensembles all perform the same computation and take (at any moment) what is functionally the same input, they should all return the same output. That is, our hypothesis predicts that ensembles conflate functionally into a rank-long unit. The results are shown in Figure 3.

Fields do not return exactly the same output, but the mean of outputs in each field is statistically constrained to a narrow range centred on the mean of mossy fibre input rates to the whole rank. By this reckoning, all granule cells in a rank receive tightly synchronised inhibition which linearly tracks the mean of mossy fibre rates. That is, Golgi cells in a rank act as a single functional unit. Field-to-field variation of output seems to reflect inherent uncertainty of the outcome of random sampling. That is, the limit on precision is computational, not biological.

Anatomical note: Golgi cells are most concentrated just below the Purkinje cell layer (Palay & Chan-Palay, 1974) but are also scattered at deeper level. The axons of deeper-lying Golgi cells are indistinguishable from the main population concentrated below the Purkinje cell layer (Palay & Chan-Palay, 1974). The simulation does not discriminate.

Figure 3 about here please.

4. Regulation of the density of active granule cells

4.1 Hypothesis

There is effectively an all-or-nothing result of the intraglomerular competition for influence of mossy fibre excitation and Golgi cell inhibition of granule cells. As a rank of Golgi cells functions as a single unit, all mossy fibre signals compete against the same threshold. The threshold is automatically adjusted so that no fewer than three mossy fibre signals are sufficient to depolarise a granule cell to firing threshold. This number is fixed by the Golgi cell computation, which automatically makes the necessary adjustment. The computational relationship of mossy fibre rates and inhibition has the effect of homeostatically regulating the fraction of the granule cell population that is active at any time. It does so by controlling the probability that a cell fires. The granule cell population of a field is large so that the proportion that fire is very near the expected number (the number predicted by probability). As the probability is controlled by inhibition, it is synchronised in a rank. Fine tuning may be provided by two feedback loops.

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

414 **4.2 Evidence**

415 This section builds on the evidence discussed in previous sections. Granule cells
416 receive input to each dendrite from a single mossy fibre. Granule cells have 3–5
417 dendrites, averaging 4 (the simulation assumes 4, but the result is the same with 3–5).
418 Contact is made in a glomerulus, as noted. Each dendrite of a single granule cell
419 extends into a different glomerulus. Estimates vary of the number of granule cells that
420 receive contact from a single terminal. 50 is mid-range (Gao et al, 2016; Jakab &
421 Hámori, 1988), but estimates vary significantly (Ritzau-Jost et al, 2014). 50 is consistent
422 with our earlier estimates of the number of terminals and granule cells in a field. A
423 granule cell receives contact from a quasi-independent random sample of mossy fibres
424 afferent to a field. The sampling unit is the whole cell since it receives only a single
425 input rate to each dendrite. Sample size is therefore 4, and the number of samples per
426 field is the number of granule cells, which we estimate at 8,750 (Supplementary
427 Materials SM4).

428 A granule cell needs a minimum number of active mossy fibre inputs to fire – thought
429 to be three (Billings et al, 2014; Jörntell & Ekerot, 2006). Certainly, one is too few
430 (Bengtsson & Jörntell, 2009; Chadderton et al, 2004; Duguid et al, 2012). We take it as
431 three. That not in itself sufficient, as each of those signals must also prevail over Golgi
432 cell inhibition mediated by spillover in the glomerulus. In each glomerulus, excitation
433 and inhibition compete for influence. The outcome of the competition is independent

434 of other glomeruli/terminals, even other terminals that provide input to the same
435 granule cell. It is unclear if there is always an emphatic winner, but a swing in the
436 dominant direction may be amplified by positive feedback, acting through presynaptic
437 receptors. Activation of presynaptic $GABA_B$ receptors on mossy fibres inhibits
438 glutamate release (Mitchell & Silver, 2000a), while activation of presynaptic mGluR2
439 receptors on Golgi cells inhibits GABA release (Mitchell & Silver, 2000b).

440 **4.3 Analysis**

441 The function of inhibition is to provide a synchronised inhibitory threshold which
442 automatically tracks mossy fibre rates. Only mossy fibre signals that prevail in the
443 glomerular competition contribute significantly to granule cell depolarisation.

444 Without regulation by Golgi cells, or if inhibition did not vary, we would expect
445 stronger mossy fibre signals to excite more granule cells, because all signals would
446 contribute to granule cell depolarisation, or more would do so. In the simulation – and
447 in fact, we suggest – the inhibitory threshold is set where no two signals are enough to
448 depolarise a granule cell to firing threshold. This effectively divides mossy fibre signals
449 into signals that are competitive and signals that aren't. A competitive signal is one
450 which is dominant in the glomerular competition, a slightly fuzzy threshold because
451 the outcome depends partly on which inhibitory 'rate' it happens to be paired with.

The probability that a given granule cell receives three competitive signals and therefore fires is a function of the percentage of mossy fibre signals that are competitive. As the percentage is unvarying (because inhibition is automated for that result), so is the probability. Because the number of granule cells is large even in a single field, the proportion that fire is near the expected number. 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, $e > i$.

We tested if the mechanism works by simulating it. 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 (the 'inhibitory threshold') at each of 4 dendrites of each of 8,750 granule cells in each of 100 fields. Mossy fibre input rates to each field are derived per section 3 and Supplementary Materials SM1, so taking into account, *inter alia*, that a field receives a

469 randomly variable number of copies of most signals. Inhibitory thresholds – outputs of
470 the ensemble conversion, 700 per field – are provided by the Golgi cell model.
471 Excitation and inhibition received by each glomerulus are a randomly selected pair of
472 inputs to a field. Pairing is used to identify the subset of granule cells that fire (cells
473 that receive at least 3 mossy fibre signals which exceed the inhibitory threshold).
474 Output of the model is the number of granule cells that fire in each field, the mean of
475 input rates they each receive (the sample means), and the mean of the sample means
476 in each field. The results are shown in Figure 4.

477 The simulation does not include a binomial function but derives output from input by
478 comparing excitatory to inhibitory input to each glomerulus, the way we claim it is
479 derived physiologically. The equation is for illustration: biology generates a result that
480 can also be obtained mathematically and in that sense is computational.

481 The test indicates that regulation by this mechanism is computationally viable if the
482 granule cell code is sparse (Fig 4 column 3), i.e., a low percentage (1–2%) of granule
483 cells is active. Above that, the mechanism performs poorly. Even sparsely coded, the
484 relative standard deviation varies significantly from field to field, predicting the
485 existence of an additional mechanism that provides local fine tuning. We discuss
486 candidates in the Discussion section.

487 The results are the same with any (tested) distribution of mossy fibre rates and also
488 whether mossy fibres are all active or some are active and others are not. Active cells

489 can be in any random permutation, subject to there being a minimum number that are
490 active at any time.

491 *Figure 4 about here please.*

492

493 **5. Granule cell-encoded information is pixellated**

494 **5.1 Hypothesis**

495 Granule cells individually 'independently' randomly sample mossy fibre input rates to a
496 rank (as discussed). Rate proportional charge entry to dendrites (which receive a
497 dominant excitatory signal), and integration at the granule cell some, provides the
498 physical form of taking the sample means. The granule cell firing rate is a linear
499 function of depolarising somatic charge. Derived in this way, granule cell rates are
500 collectively well predicted and statistically constrained to a narrower range than mossy
501 fibre rates. The granule cell population of a rank is a single functional unit. At any given
502 instant, granule cell firing rates which make up the output of a rank have a normal
503 distribution whose mean is linearly related to the mean of the sampled distribution
504 (mossy fibre rates which make up the input to a rank). In this way, granule cell activity
505 reliably and precisely conserves mossy fibre rate information. Mossy fibre rate data are
506 'processed' as if they are assumed to be measured on a common scale: the cerebellum
507 does not discriminate. Inside rank boundaries, the number and permutation of active

508 mossy fibres and the rates they fire at individually all code nothing (therefore also true
509 of granule cells). The output of a rank is the smallest unit of internally coded
510 information, and in that sense is pixellated. However, it is unnecessary to receive all of
511 it to receive the information. Granule cell firing rates are randomly distributed among
512 active granule cells (which are themselves randomly and uniformly dispersed among
513 the population). So, the granule cell code is contained in any random sample of
514 granule cell signals, over a minimum number. Therefore, target cells downstream can
515 receive the same code without receiving any of the same signals.

516 **5.2 Evidence that supports linear mossy fibre-granule cell transmission**

517 This section builds on the evidence cited previously. The hypothesis includes a claim
518 that the granule cell firing rate is a linear function of the mean of input rates. Here we
519 present evidence for this claim.

520 The short and equal length of granule cell dendrites would suggest light and equal
521 filtering. The mossy fibre-granule cell connection has multiple adaptations to transmit
522 high frequency signals (Delvendahl & Hallermann, 2016) across a wide bandwidth.
523 Faithful transmission of rate information is reported (Rancz et al, 2007; Ritzau-Jost et
524 al, 2014). Vesicle release and replenishment are fast (Saviane & Silver, 2006) and
525 postsynaptic AMPA receptors operate in their linear range (Sargent et al, 2005), where
526 they are resistant to desensitisation (DiGregorio et al, 2007). Conversion of
527 depolarising somatic charge into granule cell firing is a simple function such that

528 granule cells 'have a relatively linear and uncomplicated conversion of depolarisation
529 level to spike rate' (Bengtsson & Jörntell, 2009 p.2393, citing Jorntell and Ekerot 2006
530 and D'Angelo et al 1998).

531 In summary, there is a *prima facie* case in evidence that granule cell firing rates are
532 some linear function of input rates to the cell. However, there is not a consensus. We
533 review evidence against this conclusion in the Discussion section.

534 **5.3 Analysis**

535 **5.3.1 Test results**

536 We ran the same simulation as section 4, this time calculating the mean of input rates
537 to each active granule cell (the sample means), and the mean of the sample means in
538 each field (the 'field means'). The results are shown in Figures 5 and 6.

539 **Figure 5.** The shape and range of the frequency distribution of mossy fibre rates
540 received by a rank is unreported as far as we know. Mossy fibre rates are thought to
541 operate in a physiological range of around 250 Hz. In Figure 5 we simulate input to a
542 rank with this range – so wider range than we would expect concurrent rates to have
543 that converge on a rank – to 'stress test' the recoding computation.

544 In the top four rows, around 30% of granule cells fire in each field. The narrowing and
545 predicted shape of the granule cell frequency distribution is clearly seen and the field
546 means approach a straight line, the test that a rank operates as a single functional unit.

547 In the bottom four rows the recoding computation is modified by the addition of a
548 coefficient which sparsens the granule cell code to around 1% active. The narrowing of
549 the granule cell frequency distribution is still clear and the field means still
550 approximate a straight line. The in-field frequency distribution of granule cell rates is
551 more ragged, except the bottom row, where the mossy fibre distribution is itself
552 normal.

553 **Figure 6.** We next tested if precision is increased with a narrower mossy fibre
554 distribution, and if the granule cell mean and mossy fibre mean have a linear
555 relationship. In the top three rows of Figure 6, where mossy fibre rates are normally
556 distributed (SD 12.5 Hz), the granule cell distribution is very narrow (and normal if
557 rates are counted in narrower bins) and the field means are very close to a straight
558 line. The field means follow the mean of mossy fibre rates at a precisely fixed distance.

559 In the bottom three rows – where mossy fibre rates have a random uniform
560 distribution – we vary the range of the mossy fibre distribution. In all conditions the
561 field means remain almost a straight line and a fixed distance from the mossy fibre
562 mean. The distance depends on the mossy fibre range.

563 **5.3.2 Conclusions**

564 The recoding computation is robust. With a normal distribution or (even only
565 relatively) modest range of mossy fibre rates, the functional unity of a rank is observed

in all conditions, and there is a linear relationship of granule cell rates and mossy fibre rates, even with a sparse granule cell code.

In almost all conditions, granule cell rates in all fields share the same mean and frequency distribution. That is to say, a whole rank codes the same information. In that sense, recoding pixelates multidimensional mossy fibre input, converting it into a low-resolution internal code.

The same information is encoded in a much smaller number of signals. This can be a small fraction of active granule cells, indeed *any* small fraction. In that sense, the code is homogenised – firing rates are randomly distributed among active cells. If active granule cells are a quasi-random sample of the population, targets randomly sample granule cell rates.

Figures 5 and 6 about here please.

6. Discussion

6.1 Summary

We challenge the traditional cerebellar learning model by postulating that there is an unaided computational effect of the combination of linear transmission and bespoke functional anatomy. Because anatomy implements a biological proxy of computations,

584 and our knowledge of anatomy is detailed, we can simulate the computation in high
585 resolution.

586 In brief, we find that Golgi cells work together in anatomically overlapping ensembles
587 to convert mossy fibre signals into inhibition of granule cells. As all ensembles in a strip
588 perform the same computation and take (at any moment) the same input, ensembles
589 conflate functionally into strip-sized functional units. As a result, granule cells in a strip
590 (getting on for a million) receive synchronised inhibition which tracks the mossy fibre
591 mean. Feed-forward inhibition of granule cells by this pathway controls an adjustable
592 threshold which homeostatically regulates the probability that granule cells fire,
593 thereby regulating the fraction that are active. Like the Golgi cell ensemble conversion,
594 the conversion of mossy fibre signals into granule cell signals is functionally unitary
595 within a strip: granule cell encoded information is pixellated. While pixels are the
596 smallest unit of information, they are not the smallest unit of granule cell code.
597 Granule cell rate information is coded in any random sample of granule cell signals.

598 **6.2 The whole population of a microzone receives the same code**

599 Pixellation offers a substrate of coherent control of microzone-grouped Purkinje cells
600 downstream. Microzones are only a few cells wide (Garwicz et al, 1998a; Garwicz et al,
601 1998b; Oscarsson, 1979; Ozden et al, 2009) and lie at right angles to parallel fibres. At
602 a mid-range estimate (noting that microzone dimensions may vary regionally), a
603 microzone is populated by around 400 Purkinje cells, lying 4 deep, 40 µm apart (Palay

604 & Chan-Palay, 1974). The Purkinje cell dendritic field is large, filling the molecular layer
605 vertically, and severely flattened in the plane orthogonal to parallel fibres. An
606 estimated 350,000 parallel fibres pass through a Purkinje cell dendritic territory, of
607 which half make contact, at an average of 1.24 synapses each (i.e., one and sometimes
608 two) (Harvey & Napper, 1988; 1991; Napper & Harvey, 1988a; b). Even if only a small
609 fraction of parallel fibres is active, a single Purkinje cell still receives contact from
610 hundreds that are active at any time. If 1% of parallel fibres are active, for example, a
611 Purkinje cell receives, at any instant, ~1,750 signals. So, a sagittal strip of Purkinje cells
612 all comfortably receive enough signals to receive the same information
613 notwithstanding that almost all of them receive none of the same signals.

614 **6.3 Fine regulation of granule cell activity by feedback**

615 **6.3.1 Parallel fibre feedback**

616 Inhibition of granule cells is thought to receive modulation by feedback via parallel
617 fibres received by dendrites which some Golgi cells extend into the molecular layer. It
618 is a long-standing idea that feedback may contribute to self-regulation of the fraction
619 of co-active parallel fibres (Albus, 1971; Marr, 1969). We agree (Gilbert & Rasmussen,
620 2024). Our working hypothesis is that the strength of an effect by this pathway is
621 modest (because fine adjustments are sufficient for its function). Modulation of Golgi
622 cell rates by single parallel fibre signals is weak (Dieudonné, 1998), although sub-linear
623 integration is mitigated by gap junctions (Vervaeke et al, 2012). Also, only a modest

624 number of parallel fibre signals are received at any time. A small fraction of parallel
625 fibres which pass through a Golgi cell territory make contact (~1,200 out of a number
626 in the hundreds of thousands (Cesana et al, 2013)). Of those, only some are active at
627 any time, and if the parallel fibre code is sparse, it is a small fraction. For example, if
628 1% of parallel fibres are active, it is 12. Also, not all Golgi cells have apical dendrites
629 (Palay & Chan-Palay, 1974). Taken together, this would suggest (inconclusively) that
630 the effect of parallel fibre inputs on Golgi cell firing is relatively weak, and control by
631 direct input from mossy fibres is the main form of regulation. We envisage that the
632 feedback pathway makes fine, stabilising adjustments to an already sparse parallel
633 fibre code.

634 **6.3.2 Granule cell ascending axon feedback**

635 While the range of variation of the number of active granule cells is small, it is less
636 stable expressed as relative standard deviation. Stability may be enhanced by granule
637 cell ascending axon contact on Golgi cell basal dendrites (Cesana et al, 2010). A Golgi
638 cell receives ~400 functional synaptic contacts from this source, around 25% of the
639 granule cells that a Golgi cell receives contact from in total. During sustained mossy
640 fibre discharge, synapses from local granule cells contribute equivalent charge to Golgi
641 cells as mossy fibre synapses and 'display similarly fast kinetics' (Cesana et al, 2013),
642 providing the substrate for short-latency local feedback that adds fine adjustment to
643 feed-forward regulation. The result should be fast modulation of inhibition that is

644 inversely proportional to the local number of active granule cells, increasing precision
645 (i.e., reducing variability from time to time and place to place). We do not include local
646 feedback in the simulation.

647 **6.4 Evidence possibly counter to linear mossy fibre-granule cell transmission**

648 There is variation between granule cells (and trials) of EPSCs following mossy fibre
649 stimulation *in vitro* (Sargent et al, 2005), leading to the suggestion that fast vesicular
650 release may allow precise signalling with heterogeneous synaptic weights. However, in
651 this study stimulation was typically not above 2 Hz to avoid short-term depression
652 observed at higher frequencies, and GABA was blocked.

653 An alternative interpretation is that during behaviour, at physiological mossy fibre
654 rates, there is glomerular build-up of intraglomerular glutamate, and signalling by
655 spillover becomes important, as it is for inhibitory signals (Duguid et al, 2012; Mapelli
656 et al, 2014). Spillover from multiple release sites (200-400 at a single mossy fibre
657 terminal) is plausibly sufficient to mitigate variability of vesicular size, release
658 probability, and the number of release points at synapses made on any individual cell.
659 Indeed, short-term synaptic plasticity may have the function of increasing the relative
660 influence of spillover.

661 Still, there *are* weights. Different strength and short-term dynamics of the mossy fibre
662 to granule cell connection have been reported in vestibular slices, with $GABA_A$

receptors blocked. The amplitude of the response, measured at the soma, is reported to be statistically correlated to the source of the mossy fibre that is stimulated (Chabrol et al, 2015), suggesting that the strength of the connection depends on the source.

However, the presence of weights does not necessarily mean that the function is to modulate firing rates. Rather, it is suggested that the function may temporally code the source of inputs because different combinations have different response onset times. An alternative, which they also mention, is that some inputs may be in a supporting role, reminiscent of 'driver' and 'modulatory' signals in the thalamus and cortex (Sherman & Guillery, 2011). Feasibly, signals in a supporting role must be present for the postsynaptic granule cell to fire, but do not affect the granule cell firing rate.

We include in Supplementary Materials at SM5 some notes on other aspects of linear transmission. We also review other possible sources of control of Golgi cells.

6.5 A new solution to an old problem

A perennial issue for network theory is that biological networks do not seem to have much information about the signals they receive. This is analogous to working with data recorded on different scales of measurement but not knowing what the scales are, or what values represent. Input to the cerebellum is eclectic and in multiple

682 modalities. In some cases, a single granule cell can receive multi-modal input. In the
683 traditional cerebellar model, supervision solves this problem. It is unnecessary for the
684 cerebellum itself to discriminate because it is an assumption that externally sourced
685 instruction signals train the correct response.

686 We suggest an alternative. Our solution is that resolution of high-dimensional input
687 data is sacrificed at a surprisingly profligate level. The cerebellum disregards source
688 and modality data below a physical scale which mirrors topography. It is therefore
689 unnecessary to discriminate below that scale. Mapping studies show that mossy fibre
690 signals (evoked by cutaneous stimulation) share the topographical organisation of
691 climbing fibres (Apps & Garwicz, 2005; Apps & Hawkes, 2009; Ji & Hawkes, 1994).

692 At sub-topographical dimensions, signal source and type are simply discounted. This
693 has the flexible advantage that networks do not need to be adapted to the source or
694 type of input they receive – anatomically, networks are interchangeable, so network
695 design can be modular. As a result, it is hugely flexible, allowing the almost unchanging
696 basic wiring to survive dramatic upheavals of the vertebrate phenotype.

697

Glossary

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

field mean The mean of the granule cell sample means; the granule cell sample means are, in each field, the mean of mossy fibre rates received by each granule cell.

glomerulus field mean The mean of inhibitory thresholds in a field.

inhibitory threshold The mean of Golgi cell rates received as convergent input to a glomerulus. A mossy fibre signal must exceed this value to count towards the minimum number of active mossy fibre inputs needed for a granule cell to fire.

rank A sagittal row of 100 fields.

Grants

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

Funding sources had no involvement in the study.

Acknowledgement

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

REFERENCES

Albus, J. S. (1971) A Theory of Cerebellar Function. *Mathematical Biosciences*, 10(1-2), 25-61.

Albus, J. S. (1989) The Marr and Albus theories of the cerebellum-two early models of associative memory, *Digest of Papers. COMPCON Spring 89. Thirty-Fourth IEEE Computer Society International Conference: Intellectual Leverage*. Feb. 27 1989-March 3 1989.

Apps, R. & Garwicz, M. (2005) Anatomical and physiological foundations of cerebellar information processing. *Nat Rev Neurosci*, 6(4), 297-311.

Apps, R. & Hawkes, R. (2009) Cerebellar cortical organization: a one-map hypothesis. *Nat Rev Neurosci*, 10(9), 670-81.

Arenz, A., Silver, R. A., Schaefer, A. T. & Margrie, T. W. (2008) The contribution of single synapses to sensory representation in vivo. *Science*, 321(5891), 977-80.

Barmack, N. H. & Yakhnitsa, V. (2008) Functions of interneurons in mouse cerebellum. *J Neurosci*, 28(5), 1140-52.

Bengtsson, F. & Jörntell, H. (2009) Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. *Proc Natl Acad Sci U S A*, 106(7), 2389-94.

Billings, G., Piasini, E., Lorincz, A., Nusser, Z. & Silver, R. A. (2014) Network structure within the cerebellar input layer enables lossless sparse encoding. *Neuron*, 83(4), 960-74.

Brickley, S. G., Cull-Candy, S. G. & Farrant, M. (1996) Development of a tonic form of synaptic inhibition in rat cerebellar granule cells resulting from persistent activation of GABAA receptors. *J Physiol*, 497 (Pt 3), 753-9.

Brunel, N., Hakim, V., Isope, P., Nadal, J. P. & Barbour, B. (2004) Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell. *Neuron*, 43(5), 745-57.

Cesana, E., Dieudonné, S., Isope, P., Bidoret, C., D'Angelo, E. & Forti, L. (2010) Novel granule cell-Golgi cell excitatory input in the cerebellar granular layer, *FENS Abstr*.

Cesana, E., Pietrajtis, K., Bidoret, C., Isope, P., D'Angelo, E., Dieudonne, S. & Forti, L. (2013) Granule cell ascending axon excitatory synapses onto Golgi cells implement a potent feedback circuit in the cerebellar granular layer. *J Neurosci*, 33(30), 12430-46.

Chabrol, F. P., Arenz, A., Wiechert, M. T., Margrie, T. W. & DiGregorio, D. A. (2015) Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nat Neurosci*, 18(5), 718-27.

Chadderton, P., Margrie, T. W. & Hausser, M. (2004) Integration of quanta in cerebellar granule cells during sensory processing. *Nature*, 428(6985), 856-60.

Dean, P., Porrill, J., Ekerot, C. F. & Jörntell, H. (2010) The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nat Rev Neurosci*, 11(1), 30-43.

Delvendahl, I. & Hallermann, S. (2016) The Cerebellar Mossy Fiber Synapse as a Model for High-Frequency Transmission in the Mammalian CNS. *Trends Neurosci*, 39(11), 722-737.

Dieudonné, S. (1998) Submillisecond kinetics and low efficacy of parallel fibre-Golgi cell synaptic currents in the rat cerebellum. *J Physiol*, 510 (Pt 3), 845-66.

DiGregorio, D. A., Rothman, J. S., Nielsen, T. A. & Silver, R. A. (2007) Desensitization properties of AMPA receptors at the cerebellar mossy fiber granule cell synapse. *J Neurosci*, 27(31), 8344-57.

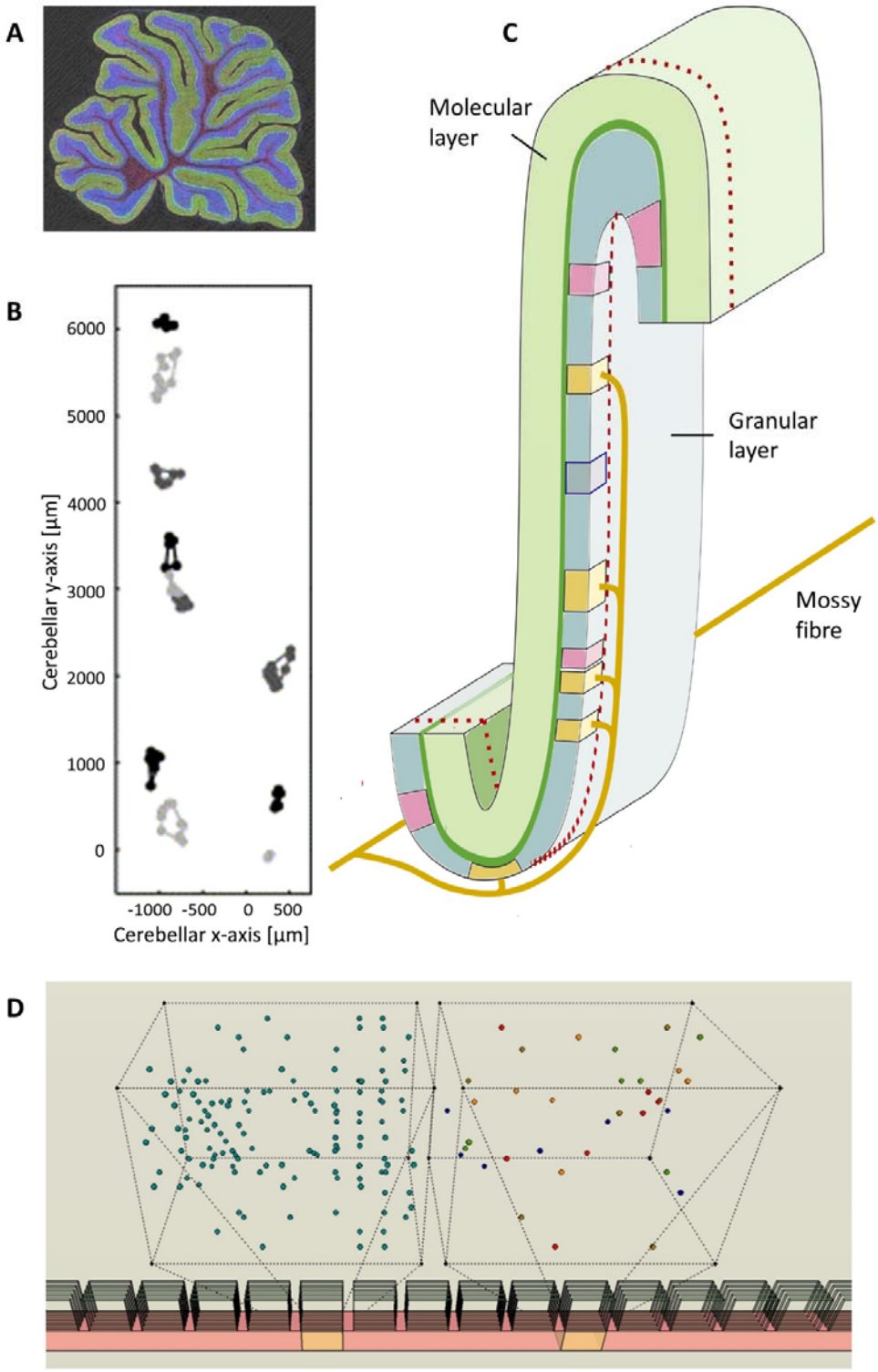
764 Duguid, I., Branco, T., Chadderton, P., Arlt, C., Powell, K. & Hausser, M. (2015) Control of
765 cerebellar granule cell output by sensory-evoked Golgi cell inhibition. *Proc Natl Acad Sci U S A*,
766 112(42), 13099-104.
767 Duguid, I., Branco, T., London, M., Chadderton, P. & Hausser, M. (2012) Tonic inhibition
768 enhances fidelity of sensory information transmission in the cerebellar cortex. *J Neurosci*,
769 32(32), 11132-43.
770 Eccles, J. C., Ito, M. & Szentágothai, J. (1967) *The cerebellum as a neuronal machine*. Berlin,
771 New York etc.: Springer-Verlag.
772 Eccles, J. C., Llinás, R. & Sasaki, K. (1964) Golgi cell inhibition in the cerebellar cortex. *Nature*,
773 204, 1265-6.
774 Edgley, S. A. & Lidierth, M. (1987) The discharges of cerebellar Golgi cells during locomotion in
775 the cat. *J Physiol*, 392, 315-32.
776 Ekerot, C. F. & Jörntell, H. (2001) Parallel fibre receptive fields of Purkinje cells and
777 interneurons are climbing fibre-specific. *Eur J Neurosci*, 13(7), 1303-10.
778 Fujita, M. (1982) Adaptive filter model of the cerebellum. *Biol Cybern*, 45(3), 195-206.
779 Gao, Z., Proietti-Onori, M., Lin, Z., Ten Brinke, M. M., Boele, H. J., Potters, J. W., Ruigrok, T. J.,
780 Hoebeek, F. E. & De Zeeuw, C. I. (2016) Excitatory Cerebellar Nucleocortical Circuit Provides
781 Internal Amplification during Associative Conditioning. *Neuron*, 89(3), 645-57.
782 Garwicz, M., Ekerot, C. F. & Jorntell, H. (1998a) Organizational principles of cerebellar neuronal
783 circuitry. *News in Physiological Sciences*, 13, 28-34.
784 Garwicz, M., Jörntell, H. & Ekerot, C. F. (1998b) Cutaneous receptive fields and topography of
785 mossy fibres and climbing fibres projecting to cat cerebellar C3 zone. *J Physiol*, 512 (Pt 1), 277-
786 93.
787 Gilbert, M. & Miall, R. C. (2021a) How and Why the Cerebellum Recodes Input Signals: An
788 Alternative to Machine Learning. *Neuroscientist*, DOI: 10.1177/1073858420986795.
789 Gilbert, M. & Rasmussen, A. (2024) Gap Junctions May Have A Computational Function In The
790 Cerebellum: A Hypothesis. *The Cerebellum*.
791 Giovannucci, A., Badura, A., Deverett, B., Najafi, F., Pereira, T. D., Gao, Z., Ozden, I., Kloth, A.
792 D., Pnevmatikakis, E., Paninski, L., De Zeeuw, C. I., Medina, J. F. & Wang, S. S. (2017) Cerebellar
793 granule cells acquire a widespread predictive feedback signal during motor learning. *Nat*
794 *Neurosci*.
795 Hámori, J. & Szentágothai, J. (1966) Participation of Golgi neuron processes in the cerebellar
796 glomeruli: an electron microscope study. *Exp Brain Res*, 2(1), 35-48.
797 Hartmann, M. J. & Bower, J. M. (2001) Tactile responses in the granule cell layer of cerebellar
798 folium crus IIa of freely behaving rats. *J Neurosci*, 21(10), 3549-63.
799 Harvey, R. J. & Napper, R. M. (1988) Quantitative study of granule and Purkinje cells in the
800 cerebellar cortex of the rat. *J Comp Neurol*, 274(2), 151-7.
801 Harvey, R. J. & Napper, R. M. (1991) Quantitative studies on the mammalian cerebellum. *Prog*
802 *Neurobiol*, 36(6), 437-63.
803 Holtzman, T., Rajapaksa, T., Mostofi, A. & Edgley, S. A. (2006) Different responses of rat
804 cerebellar Purkinje cells and Golgi cells evoked by widespread convergent sensory inputs. *J*
805 *Physiol*, 574(Pt 2), 491-507.
806 Isope, P. & Barbour, B. (2002) Properties of unitary granule cell-->Purkinje cell synapses in
807 adult rat cerebellar slices. *J Neurosci*, 22(22), 9668-78.

808 Jakab, R. L. & Hámori, J. (1988) Quantitative morphology and synaptology of cerebellar
809 glomeruli in the rat. *Anat Embryol (Berl)*, 179(1), 81-8.
810 Jelitai, M., Puggioni, P., Ishikawa, T., Rinaldi, A. & Duguid, I. (2016) Dendritic excitation-
811 inhibition balance shapes cerebellar output during motor behaviour. *Nat Commun*, 7, 13722.
812 Ji, Z. & Hawkes, R. (1994) Topography of purkinje cell compartments and mossy fiber terminal
813 fields in lobules ii and iii of the rat cerebellar cortex: Spinocerebellar and cuneocerebellar
814 projections. *Neuroscience*, 61(4), 935-954.
815 Jörntell, H. & Ekerot, C. F. (2006) Properties of somatosensory synaptic integration in
816 cerebellar granule cells in vivo. *J Neurosci*, 26(45), 11786-97.
817 Kanichay, R. T. & Silver, R. A. (2008) Synaptic and cellular properties of the feedforward
818 inhibitory circuit within the input layer of the cerebellar cortex. *J Neurosci*, 28(36), 8955-67.
819 Mapelli, L., Solinas, S. & D'Angelo, E. (2014) Integration and regulation of glomerular inhibition
820 in the cerebellar granular layer circuit. *Front Cell Neurosci*, 8, 55.
821 Marr, D. (1969) A theory of cerebellar cortex. *J Physiol*, 202(2), 437-70.
822 Mitchell, S. J. & Silver, R. A. (2000a) GABA spillover from single inhibitory axons suppresses
823 low-frequency excitatory transmission at the cerebellar glomerulus. *J Neurosci*, 20(23), 8651-8.
824 Mitchell, S. J. & Silver, R. A. (2000b) Glutamate spillover suppresses inhibition by activating
825 presynaptic mGluRs. *Nature*, 404(6777), 498-502.
826 Moore, J. W. & Blazis, D. E. J. (1989) 11 - Simulation of a Classically Conditioned Response: A
827 Cerebellar Neural Network Implementation of the Sutton–Barto–Desmond Model, in Byrne, J.
828 H. & Berry, W. O. (eds), *Neural Models of Plasticity*Academic Press, 187-207.
829 Napper, R. M. & Harvey, R. J. (1988a) Number of parallel fiber synapses on an individual
830 Purkinje cell in the cerebellum of the rat. *J Comp Neurol*, 274(2), 168-77.
831 Napper, R. M. & Harvey, R. J. (1988b) Quantitative study of the Purkinje cell dendritic spines in
832 the rat cerebellum. *J Comp Neurol*, 274(2), 158-67.
833 Nusser, Z., Sieghart, W. & Somogyi, P. (1998) Segregation of different GABAA receptors to
834 synaptic and extrasynaptic membranes of cerebellar granule cells. *J Neurosci*, 18(5), 1693-703.
835 Oscarsson, O. (1979) Functional units of the cerebellum - sagittal zones and microzones.
836 *Trends in Neurosciences*, 2, 143-145.
837 Ozden, I., Sullivan, M. R., Lee, H. M. & Wang, S. S. (2009) Reliable coding emerges from
838 coactivation of climbing fibers in microbands of cerebellar Purkinje neurons. *J Neurosci*, 29(34),
839 10463-73.
840 Palay, S. L. & Chan-Palay, V. (1974) *Cerebellar cortex : cytology and organization*. Berlin:
841 Springer.
842 Rancz, E. A., Ishikawa, T., Duguid, I., Chadderton, P., Mahon, S. & Hausser, M. (2007) High-
843 fidelity transmission of sensory information by single cerebellar mossy fibre boutons. *Nature*,
844 450(7173), 1245-8.
845 Raymond, J. L. & Medina, J. F. (2018) Computational Principles of Supervised Learning in the
846 Cerebellum. *Annu Rev Neurosci*, 41, 233-253.
847 Ritzau-Jost, A., Delvendahl, I., Rings, A., Byczkowicz, N., Harada, H., Shigemoto, R., Hirrlinger, J.,
848 Eilers, J. & Hallermann, S. (2014) Ultrafast action potentials mediate kilohertz signaling at a
849 central synapse. *Neuron*, 84(1), 152-163.
850 Rossi, D. J. & Hamann, M. (1998) Spillover-mediated transmission at inhibitory synapses
851 promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry.
852 *Neuron*, 20(4), 783-95.

853 Rossi, D. J., Hamann, M. & Attwell, D. (2003) Multiple modes of GABAergic inhibition of rat
854 cerebellar granule cells. *J Physiol*, 548(Pt 1), 97-110.
855 Ruigrok, T. J., Hensbroek, R. A. & Simpson, J. I. (2011) Spontaneous activity signatures of
856 morphologically identified interneurons in the vestibulocerebellum. *J Neurosci*, 31(2), 712-24.
857 Sargent, P. B., Saviane, C., Nielsen, T. A., DiGregorio, D. A. & Silver, R. A. (2005) Rapid vesicular
858 release, quantal variability, and spillover contribute to the precision and reliability of
859 transmission at a glomerular synapse. *J Neurosci*, 25(36), 8173-87.
860 Saviane, C. & Silver, R. A. (2006) Fast vesicle reloading and a large pool sustain high bandwidth
861 transmission at a central synapse. *Nature*, 439(7079), 983-7.
862 Schonewille, M., Gao, Z., Boele, H. J., Veloz, M. F., Amerika, W. E., Simek, A. A., De Jeu, M. T.,
863 Steinberg, J. P., Takamiya, K., Hoebeek, F. E., Linden, D. J., Huganir, R. L. & De Zeeuw, C. I.
864 (2011) Reevaluating the role of LTD in cerebellar motor learning. *Neuron*, 70(1), 43-50.
865 Sherman, S. M. & Guillery, R. W. (2011) Distinct functions for direct and transthalamic
866 corticocortical connections. *J Neurophysiol*, 106(3), 1068-77.
867 Shinoda, Y. & Sugihara, I. (2013) Axonal Trajectories of Single Climbing and Mossy Fiber
868 Neurons in the Cerebellar Cortex and Nucleus, in Manto M., S. J. D., Rossi F., Gruol D.L.,
869 Koibuchi N. (ed), *Handbook of the Cerebellum and Cerebellar Disorders*Springer, Dordrecht.
870 Shinoda, Y., Sugihara, I., Wu, H. S. & Sugiuchi, Y. (2000) The entire trajectory of single climbing
871 and mossy fibers in the cerebellar nuclei and cortex. *Prog Brain Res*, 124, 173-86.
872 Sultan, F. & Heck, D. (2003) Detection of sequences in the cerebellar cortex: numerical
873 estimate of the possible number of tidal-wave inducing sequences represented. *J Physiol Paris*,
874 97(4-6), 591-600.
875 ten Brinke, M. M., Boele, H. J., Spanke, J. K., Potters, J. W., Kornysheva, K., Wulff, P., AC, I. J.,
876 Koekkoek, S. K. & De Zeeuw, C. I. (2015) Evolving Models of Pavlovian Conditioning: Cerebellar
877 Cortical Dynamics in Awake Behaving Mice. *Cell Rep*, 13(9), 1977-88.
878 van Kan, P. L., Gibson, A. R. & Houk, J. C. (1993) Movement-related inputs to intermediate
879 cerebellum of the monkey. *J Neurophysiol*, 69(1), 74-94.
880 Vervaeke, K., Lorincz, A., Nusser, Z. & Silver, R. A. (2012) Gap junctions compensate for
881 sublinear dendritic integration in an inhibitory network. *Science*, 335(6076), 1624-28.
882 Vos, B. P., Volny-Luraghi, A. & De Schutter, E. (1999) Cerebellar Golgi cells in the rat: receptive
883 fields and timing of responses to facial stimulation. *Eur J Neurosci*, 11(8), 2621-34.
884 Walter, J. T. & Khodakhah, K. (2006) The linear computational algorithm of cerebellar Purkinje
885 cells. *J Neurosci*, 26(50), 12861-72.
886 Walter, J. T. & Khodakhah, K. (2009) The advantages of linear information processing for
887 cerebellar computation. *Proc Natl Acad Sci U S A*, 106(11), 4471-6.
888 Welsh, J. P., Yamaguchi, H., Zeng, X. H., Kojo, M., Nakada, Y., Takagi, A., Sugimori, M. & Llinas,
889 R. R. (2005) Normal motor learning during pharmacological prevention of Purkinje cell long-
890 term depression. *Proc Natl Acad Sci U S A*, 102(47), 17166-71.
891 Wu, H. S., Sugihara, I. & Shinoda, Y. (1999) Projection patterns of single mossy fibers
892 originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei. *J Comp
893 Neurol*, 411(1), 97-118.

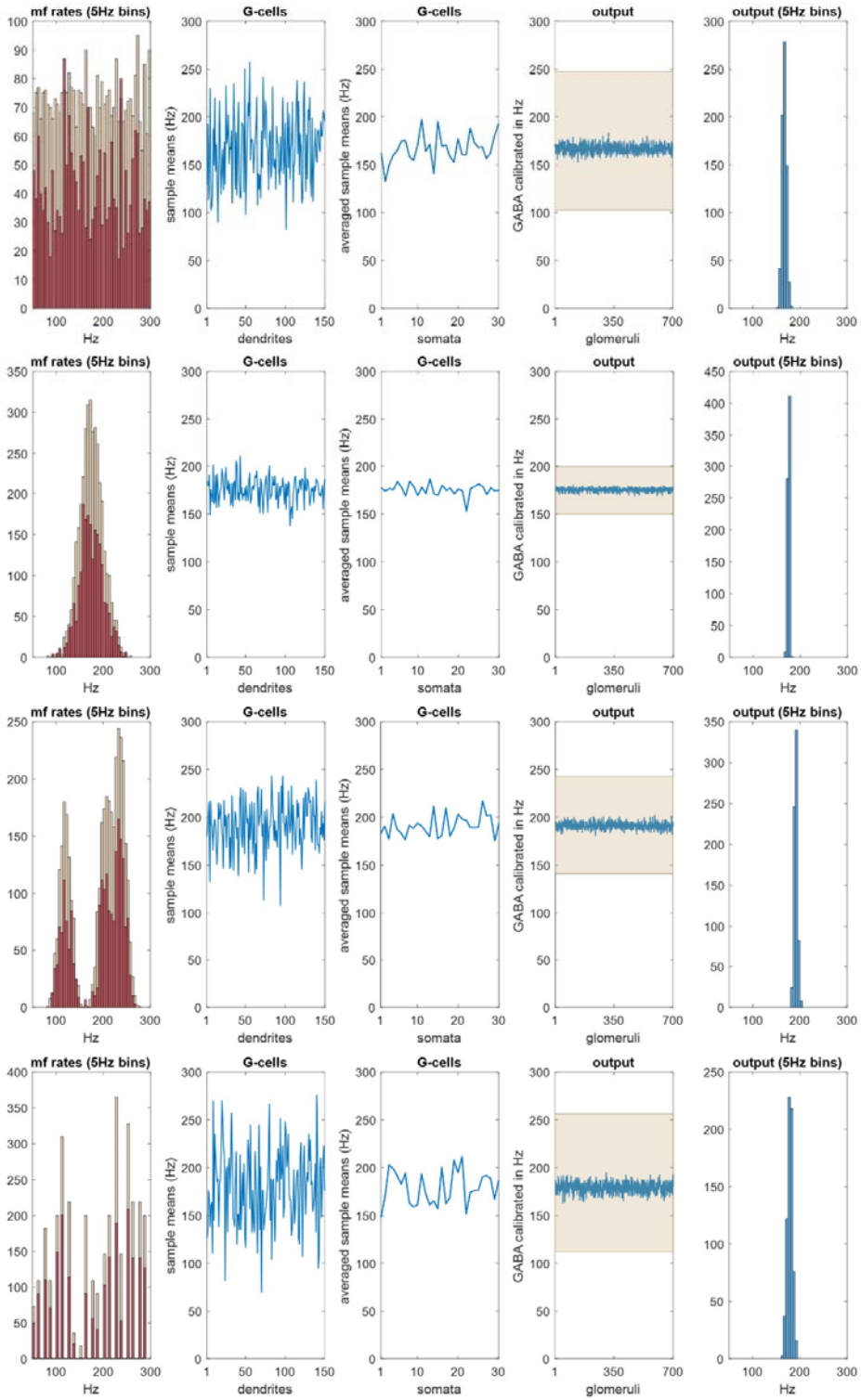
894

895


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

896

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


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

925

Fig 2. Golgi cell ensemble conversion. Column 1 We randomly generated (3,645) values representing mossy fibre rates with a uniform, normal, twin-peaked or discontinuous distribution, simulating input to a row of 100 fields. Rate data were randomly sampled to represent input to individual fields and copies added to reflect variable size of terminal clusters and that most clusters straddle field boundaries. Red data represent copies-included signals received by a row of three fields. The makeup of input to each field is an important feature of the simulation; see main text and Supplementary Materials SM1 for derivation of rates. **Column 2** The first-column red data were randomly sampled 150 times, sample size 4, representing mossy fibre input to a 3-field population of 30 Golgi cells, each with 5 basal dendrites (the mean), and each of those receiving contact from 4 mossy fibres. Data shown are the sample means, representing dendritic depolarisation. **Column 3** We next randomly sampled column 2 data *without* replacement 30 times, sample size 5, and took the sample means, representing somatic integration of dendritic signals. **Column 4** We then took 700 random samples of column 3 data, sample size 8–12, reflecting convergent input from 3 fields of Golgi cells onto 700 glomeruli in the middle field (which receives the output of a three-field ensemble), and took the sample means, representing GABA spillover, which has a linear relationship with the mean of afferent Golgi cell rates. Inhibition of granule cells takes place in the glomerulus and is almost exclusively mediated by spillover. **Column 5** The column 4 data are presented as a histogram. **All rows** Column 1 mossy fibre data are converted to narrowly-focussed inhibition of granule cells. The mean of the column 4/5 data is linearly related to the mean of the sampled distribution. This result is independent of the shape of the mossy fibre distribution.

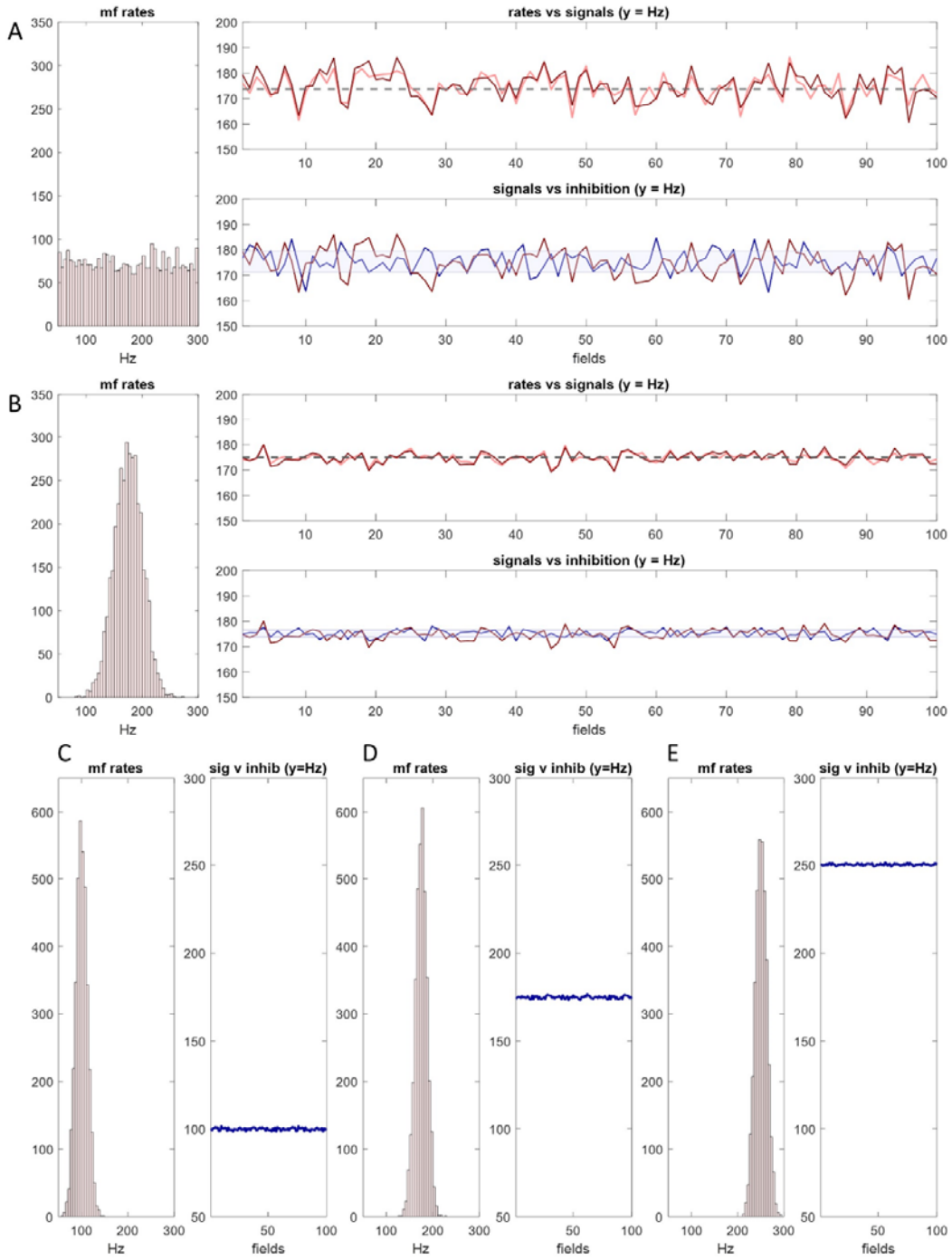


Fig 3. The Golgi cell population of a rank functions as a single unit. A Simulation of Golgi cells in a sagittal row of 100 fields, taking the mossy fibre distribution on the left as input (see Fig 2 for derivation of mossy fibre rates). **Top right**: pink: the mean of

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 row. **Bottom right**: red: same as top right; blue: mean of inhibitory thresholds (the 'glomerulus field mean'); pale blue: SD of inhibitory thresholds. **All** y = Hz. **B** Same as (A) but sampling a different mossy fibre distribution.

A–B interpretation

1. 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.

2. Blue data across all fields are in a narrowly circumscribed range. The degree of precision of the ensemble computation (how tightly grouped the data are) depends on the bandwidth of the sampled distribution (mossy fibre rates). A normal mossy fibre distribution with a 2SD range of over 100 Hz is converted to a blue range of ~3 Hz.

3. Fields that sample the same input data return the same output: inhibition of granule cells is synchronised in a rank.

4. 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 fields randomly sample mossy fibre input rates to a rank and Golgi cells randomly sample input rates to a field is not compounded – random variation of the red data doesn't flow through to bigger random variation of the blue data (in fact the blue data have a slightly smaller range).

5. We also tested the effect of setting sample size 'optimally' (i.e., higher, at 24:1, not shown here; see SM2) for mossy fibres that contact a single Golgi cell dendrite, and Golgi cells that inhibit a single glomerulus. This only very modestly improved performance: higher ratios return a barely narrower blue range, indicating that precision of the computation is not limited by anatomy.

C The same simulation. Left: the sampled distribution (mossy fibre input rates to a rank); right: the glomerulus 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. **D** and **E** The same, moving the sampled distribution to the right.

C–D interpretation Inhibition of granule cells is precisely and linearly coupled to the mean of the sampled distribution.

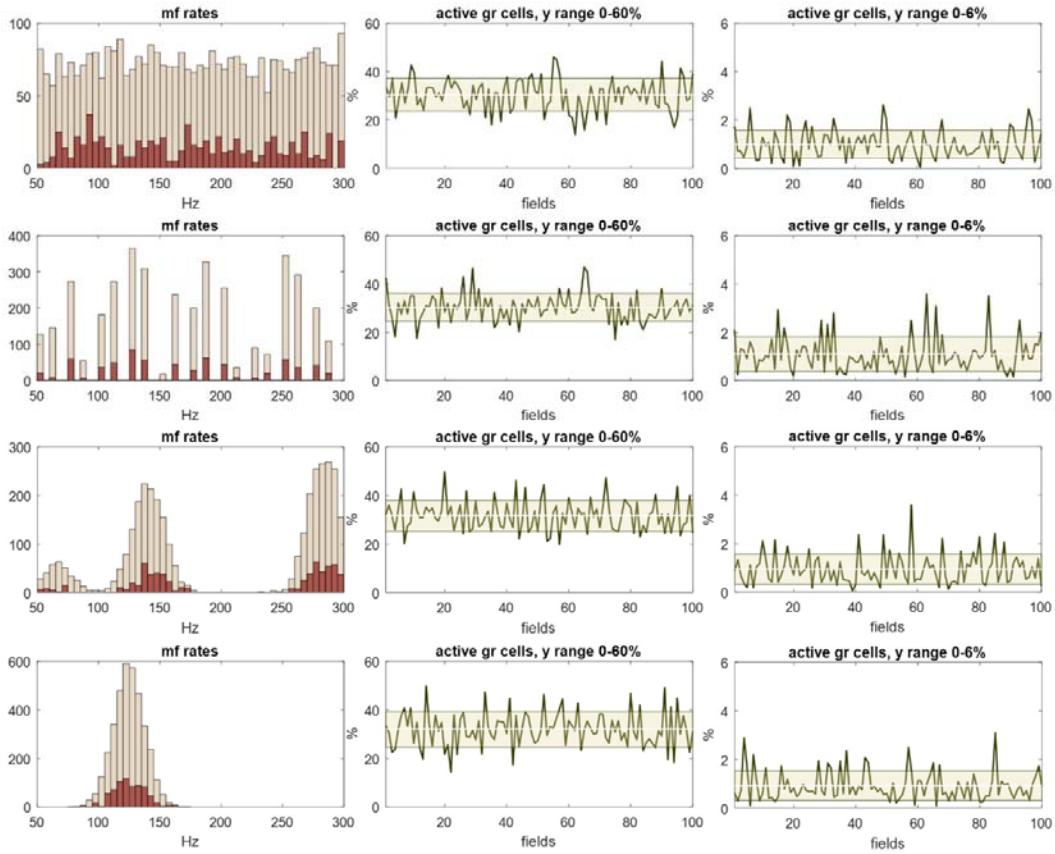
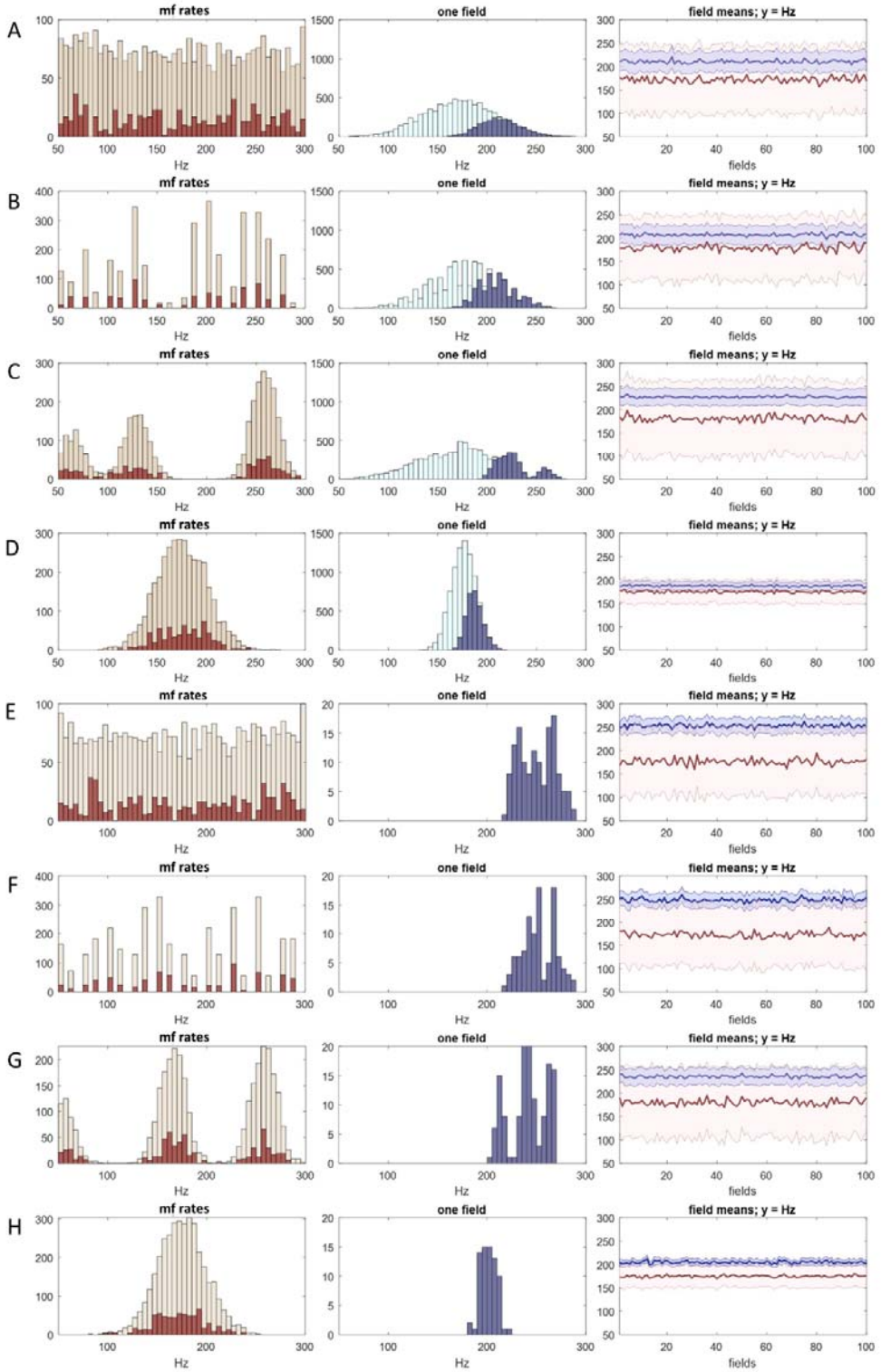


Fig 4. Regulation of granule cell activity by the feed-forward pathway (mossy fibres-Golgi cells-granule cells) We incorporated the Golgi cell ensemble conversion into a simulation of recoding that converts mossy fibre signals (the main input to the cerebellum) into an internal code (granule cell signals). One of the outputs of recoding is regulation of the fraction of granule cells that are concurrently active. **Column 1 rows 1 to 4, respectively** Mossy fibre rates were randomly generated with a uniform, discontinuous, three-peak or normal distribution of firing rates, to represent input to a row of 100 fields. Input to each field was derived in the same way as before. 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 all rows** 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 all rows** The same, with an adjustment (a co-efficient) to sparsen granule cell activity. Column 3 y axes have 1/10 the range of column 2. The value of the coefficient depends on the shape and bandwidth of the sampled distribution, suggesting that those parameters are not time variable

1007 physiologically. The percentage of active cells in each field is condensed into a narrow
1008 range but relative SD suggests fine tuning is provided by an additional mechanism(s).

1009



1010

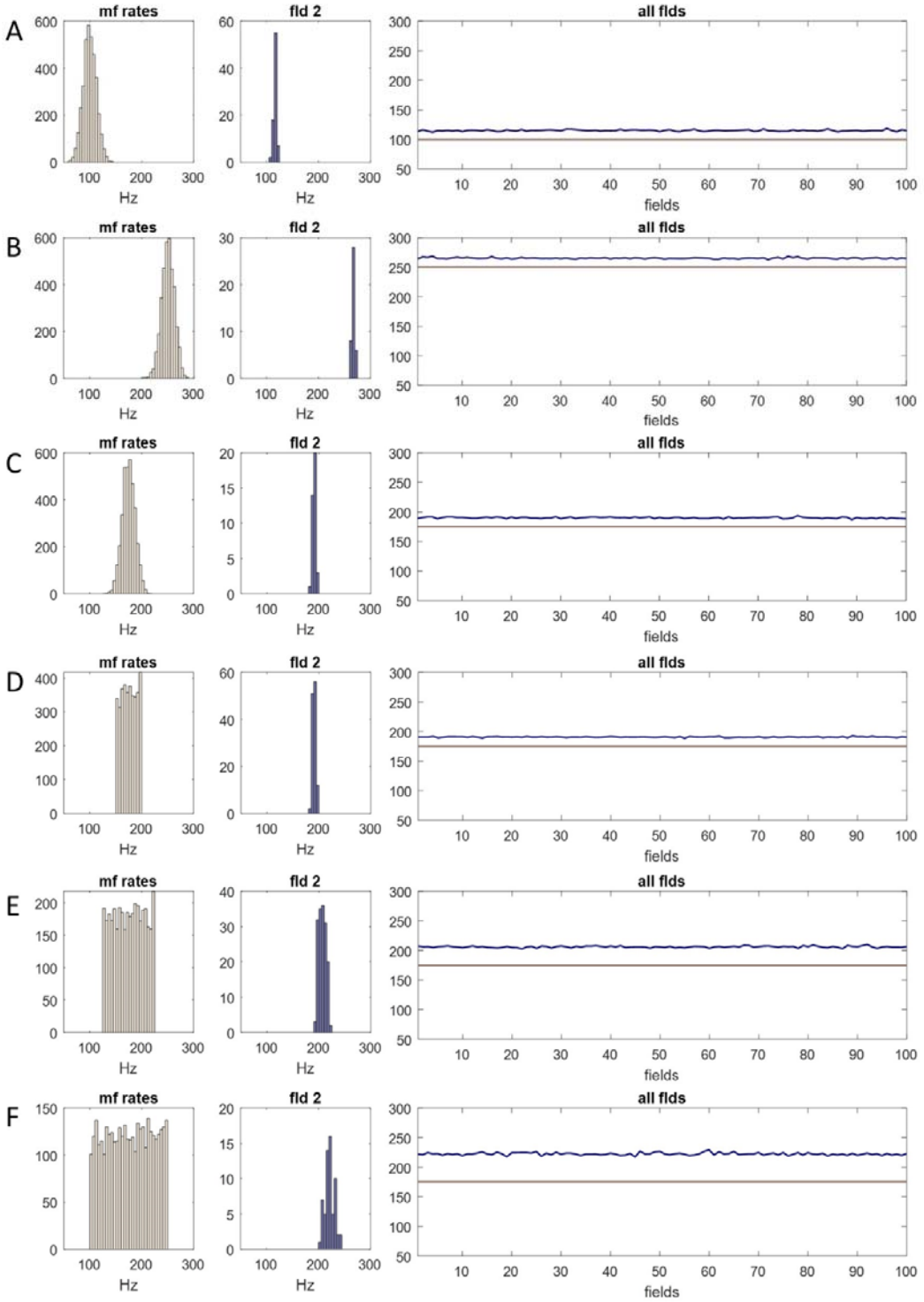
Fig 5. A long sagittal row of fields recodes as a single unit. Column 1 Mossy fibre input rates to a pixel and to fields were generated in the same way as earlier figures. Red data are rates received by a single field with duplicate signals added. **Column 2** Red data in column 1 were randomly sampled 8,750 times, sample size 4, representing a per field population of 8,750 granule cells, each with 4 dendrites. Pale green: sample means – the mean of input rates to each granule cell; blue: sample means of the subset of granule cells which meet the conditions to fire. In rows A–D, around 30% fire. In rows E–H, around 1% fire; y axes are rescaled and green data are omitted. **Column 3** Red: mean of mossy fibre input rates to a row of 100 fields, including duplicate signals; pink: SD of mossy fibre input rates to each field; blue: field means (i.e., for each field, the mean of the blue data in column 2); pale blue: SD of granule cell rates in each field.

Interpretation

1. The output of all fields – column 3, blue data – is statistically constrained to a narrow range. This result is independent of the shape of the sampled distribution (mossy fibre rates in column 1) and of the fact that sparse code has an untidy in-field distribution (E–H, column 2, because the number of active granule cells in a single field is relatively modest).

2. We note previously that fine tuning of regulation of the number of active granule cells in each field is likely to be provided by feedback. Variability of the number in the simulation causes the blue data to also be more variable (because the number of active simulated granule cells in some fields is low, which skews the mean). Fine-tuned regulation should cause the column 3 blue data to more nearly approach a straight line, parallel to the mean of the red data: a sagittal row of fields recodes as a single unit.

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



1040

1041 **Fig 6. Granule cell firing rates collectively code mossy fibre rate information. Column**
1042 **1** We generated mossy fibre rates as before with either a normal (SD 12.5) or uniform
1043 distribution representing input to a row of 100 fields, with copies added – the sampled
1044 distribution. **Column 2** Frequency distribution of the mean of mossy fibre input rates
1045 to each of the granule cells that fire in a randomly selected field, obtained with the
1046 recoding model in the same way as in the bottom 4 rows of Fig 5. **Column 3** Grey line:
1047 the mean of the sampled distribution; blue line: field means. **A–C** The fields means are
1048 precisely and linearly related to the mean of mossy fibre rates. Blue data approach a
1049 straight line which is parallel to the grey line at a mean distance of 15 Hz (SD 1.21 Hz)
1050 in all rows. **D–F** If the shape and mean of a uniform frequency distribution of mossy
1051 fibre rates are unchanging, the distance between the blue and grey lines is linearly
1052 related to the range.

1053