What if…? The granular layer, the input layer of the cerebellar network: a theory

Mike Gilbert[a]*, 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**

Mike Gilbert

m.gilbert.1@bham.ac.uk

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

ABSTRACT

It is thought that neuronal signal frequencies – firing rates – provide the physiological form of values that are used in neural network computations. The cerebellum is made up of repeating, anatomically overlapping networks. Neural network modelling has a long tradition of learning algorithms that are implemented physiologically by learned changes of synaptic transmission strength. In the cerebellum, the evidence remains inconclusive. We ask, with a focus on the input layer of the cerebellar network: (1) If the physiological instrument of the cerebellar network computation is not synaptic memory, what is it, and what is the computation? (2) What is the code? (3) How is that related to the matrix-like internal architecture of the cerebellar cortex, and the well-known functional organisation of the cortex into long, thin strips called microzones? We propose a reinterpretation of the evidence which plausibly supports controversial conclusions. Using the same numbering: (1) The granular layer computation is the functional result of anatomy. Simply passing signals through the entry layer sub-network provides the biological equivalent of mathematical operations, which run in series and in parallel. (2) Information is collectively coded in statistics of the firing rates of co-active cells. Coded this way, there are reliable statistical effects that conserve rate information when input to the cerebellum is converted to internal signals. (3) Still more unexpectedly, multi-dimensional detail contained in input signals is converted to an extremely low resolution – pixelated – internal code. This provides an otherwise elusive mechanism to coordinate firing of microzone-grouped Purkinje cells, which control the output cells of the cerebellar network. It is also a theoretical substrate for a path of rate codes that connect input to output of the cerebellar network. Since we argue that anatomy is the chief instrument of the granular layer computation, we can model it in detail, down to local anatomical noise. If this approach is sound, the idea that anatomy implements computations unaided is likely to have implications for coding in other brain regions.

1. INTRODUCTION

It is thought that neuronal signal frequencies – firing rates – provide the physiological form of 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 that contact between cells is adapted for faithful transmission of rate information (Cesana et al, 2010; Delvendahl & Hallermann, 2016; Moore & Blazis, 1989; Saviane & Silver, 2006), and 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). It is difficult to reconcile linear transmission with models that use experience to train synaptic weight adjustments.

Input rates and firing of a cell are connected by charge movement into and inside the cell. Using a classic biological neuron model, that would normally mean modelling the complex relationships of biophysical parameters that are normally experimentally measured on different scales. We explore the 'what if' scenario that all steps in intracellular transmission can be represented on a normalised scale. Put another way, we make the supposition that *data-coding* biophysical variables in fact have that relationship. Making that assumption, we found that anatomical (and other) evidence took on unexpected computational significance. Moreover, we found that internal signalling plausibly participates in network computations, and there is no need to rely on synaptic weights for a computational effect.

Our focus is the input layer of the cerebellar network. 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, 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 asked:

1) If the instrument of the cerebellar network computation is not synaptic weights, what is it?

2) How does the cerebellum encode information internally, in firing of granule cells? The traditional model does not need a code because learning transfers control of Purkinje cells from firing rates to synaptic weights. If control is instead by rates, as we propose, what is the code?

3) How does that help to explain cortical architecture, including the well-established organisation of Purkinje cells into microzones (Garwicz et al, 1998; Oscarsson, 1979; Ozden et al, 2009)?

Viewed without a learning prism, we find the evidence appears to support an unexpected interpretation. Using the same numbering:

1) Anatomy – cell morphologies and network architecture – feasibly provides the biological instrument of mathematical operations.

2) Granule cell rate information is collectively coded in the frequency distribution of firing rates. Code in this form means mossy fibre rate information is conserved by a statistical relationship within functionally defined spatial boundaries. This works for input in any permutation of active cells and rates. Anatomical noise does not impair computational performance.

3) Still more surprisingly, high-dimensional input data are converted to a low-resolution internal code. The smallest unit of granule cell encoded information has the dimensions of a long strip of the granular layer. In this sense, the internal code is pixelated. This has the asset that the internal code is intelligible by randomly sampling signals. That means in turn that Purkinje cells can receive different signals but the same code, providing a substrate for coherent firing of microzone-grouped Purkinje cells downstream.

This is a departure from conventional approaches. With that in mind, we provide physiological ballast in the shape of a review of evidence. Each main section of the paper begins with a sub-section of physiology. Derivation of parameters and the values they take (usually a range) is important to our approach and part of our method. Evidence of linear transmission is reviewed in the Discussion section.

Our main use of computational tools is to verify that the ideas are computationally viable, rather than as the chief language of the proposals. As we postulate that anatomy and network architecture provide the mechanism of computations, we are able to computer simulate the proposals in high-resolution detail because there is detailed anatomical information available.

2. MOSSY FIBRES

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. In this section, we derive, from firing rates received by a strip, rates received by a defined sector, and the number of copies of each signal, taking into account 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 sector. We will use that in two ways: to explain the computational function of termination at multiple locations, and to generate mossy fibre firing rates that the Golgi cell and granule cell models take as input in sections 3 and 4 respectively, with attention to noisy local detail.

2.1 Physiology

The molecular layer of the cerebellar cortex is organised functionally into long thin strips called microzones, around 100 times long as they are wide, populated by Purkinje cells and interneurons. Some 35 million parallel fibres pass through a microzone, at right angles, approaching from both sides, from a distance of up to 3 mm (Brand et al, 1976; Harvey & Napper, 1988; 1991; Mugnaini, 1983).

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 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.

So, mossy fibres do not terminate at a point but a region with dimensions: the field occupied by a terminal cluster. 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 & Heck, 2003), a 'field'. Microzone dimensions vary, but a mediolateral span of 150 µm is mid-range (Dean et al, 2010; Garwicz et al, 1998; Oscarsson, 1979; Ozden et al, 2009). A sagittal row of 75–100 fields (a 'rank') therefore has the dimensions of a microzone viewed from the cerebellar surface. A rank is not a microzone, however. A rank is a division of the granular layer; microzones are defined by their climbing fibre input, received in the molecular layer. Mossy fibre input to the cerebellum evoked by cutaneous stimulation largely shares the topographical organisation of climbing fibres (Apps & Garwicz, 2005; Apps & Hawkes, 2009; Ji & Hawkes, 1994), but with less clearly defined boundaries.

Figure 1 about here please

2.2 Derivation of the number of mossy fibres that innervate a rank

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

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

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

We work from an estimate of the total number of terminals per field, which is, in turn, derived from convergent estimates of the number of granule cells per field. (Estimate 1) There are an estimated 1.92×10^6 granule cells per µl (Harvey & Napper, 1988). The number of fields which fit in that volume is $(1{,}000/150)^2 \times (1{,}000/200) = \sim 222$, assuming a field depth of 150 µm (Sultan & Heck, 2003). The number of granule cells in a field is, therefore 1,920,000/222 = 8,649. (Estimate 2) An estimated 350,000 parallel fibres pass through a single Purkinje cell's territory (Harvey & Napper, 1988). Parallel fibres run about 3 mm in each direction from the granule cell soma (Brand et al, 1976; Mugnaini, 1983). The Purkinje cell dendritic arbour fills the molecular layer vertically and averages ~200 µm wide. Therefore, 6,000/150 = 40 fields fit in a row that provides parallel fibres that pass through a Purkinje cell dendritic territory, giving 350,000/40 = 8,750 granule cells per field.

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

Cluster size varies. We take the range as 4–12 terminals per cluster (Sultan & Heck, 2003). We assume the number varies at random and cluster sizes occur with equal probability. A maximum of 7 fits in a field (because a field shares the dimensions of the area that encloses a mean cluster). Fields are not anatomically bounded, so many

clusters straddle field limits. To reflect that, we assume there is a 1/7 probability that a 7-terminal cluster (or larger) has 7 terminals inside a field and also the same probability of any other number. A 6-terminal cluster has a 2/7 chance that all 6 terminals are inside and a 1/7 chance of any other number, and so on.

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

Table 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

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

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.

2.3 Derivation of mossy fibre input rates to a field

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

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

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

2.4 Mossy fibre terminal branching simulates sampling with replacement

Termination of a single mossy fibre at multiple locations in a rank is an anatomical solution to a computational problem. The problem is that anatomy must simulate random sampling with replacement. With replacement means, samples are independent.

In the commonly used example that samples are being drawn from a bag of coloured balls, the balls are replaced between samplings, so that taking them out does not affect the probability that subsequent samples include the same colours. Hence sampling 'with replacement'.

The equivalent problems facing mossy fibre anatomy is that all fields in a rank must receive an independent random sample of the entire distribution of mossy fibre input rates to the rank as a whole, simultaneously. The reason they must do so – we will argue – is functional. It is so that mossy fibre input to a whole rank is converted to the same granule cell code at all locations. This conversion is the subject of the rest of the paper.

If each mossy fibre could terminate in only one field, a mossy fibre 'sampled' by one field cannot also be sampled by any others. However, termination by a single mossy fibre in a random number of terminal branches, at multiple locations, removes that restriction. The physiological equivalent of random sampling is innervation of each field by a random sample of the mossy fibre population afferent to the whole rank. The result is an approximation of sampling with replacement.

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 an anatomical limit. But in practice even with real replacement the number of repeat samplings is constrained by a probability distribution. That is, at large numbers, the number of fields that can receive a signal is limited just as surely as if there was actual replacement.

We can test if terminal branching simulates replacement 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 4.

Table 4. 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

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), noting that data are scarce. 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.

2.5 Mossy fibre firing rates

Some very high mossy fibre rates can be generated in experimental conditions. However, it is thought the typical physiological range is 50–300 Hz (Jörntell & Ekerot, 2006) citing (van Kan et al, 1993). In the simulations, rates are generated in that range, within stated constraints on the shape of the distribution. Input rates to each field are derived from the population of input rates to a rank. Derivation uses the protocol described in this section, so that input to a field exactly reflects messy biological parameters, and random variation of anatomy between fields.

3. GOLGI CELLS

In this section, we propose a functional organisation of Golgi cells into 3-field groups. Ensemble-grouped Golgi cells all innervate substantially the whole of the middle field: convergence onto the middle field defines the group. Ensembles implement a computational conversion. We argue that anatomy implements ensemble computations. We test that by replicating an ensemble *in silico* using computationally relevant metrics of anatomy. Anatomical evidence provides parameter values, normally a range. Mossy fibre 'input' to the computation is to an ensemble territory – i.e., three fields – as distinct from only signals that are received by Golgi cells. Output is inhibition of middle-field granule cells. More exactly, there are 700 outputs, the number of glomeruli in a field, where inhibition of granule cells takes place. Ensembles are not anatomically bounded and merge functionally into rank-long strips which effectively form a single functional unit.

3.1 Physiology

Golgi cells are large interneurons whose cell bodies lie in the granular layer. They are most concentrated just below the Purkinje cell layer (Palay & Chan-Palay, 1974). 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 – mean range 650 +/- 179 µm by 180 +/- 40 µm (in mice: Barmack & Yakhnitsa, 2008) – and is the depth of the granular layer, which it fills vertically. Their density (Llinás & Negrello, 2015) and large, profusely ramified axonal field (Barmack & Yakhnitsa, 2008) means that Golgi

cell axonal territories overlap extensively in the sagittal direction. The Golgi cell population of a row of three fields all inhibit approximately the whole of the middle field. Convergence of their axonal fields defines the size and territory of the smallest functional Golgi cell unit – an ensemble. Their combined range is longer – five fields. But convergence on the middle field is functionally significant because it means all granule cells in the middle field receive inhibition from a random sample of a three-field unit.

There is a significant population of Golgi cells at deeper level. Deeper cells tend to have smaller cell bodies and to lack dendrites that extend into the molecular layer. However, the axons of smaller Golgi cells are indistinguishable from large Golgi cells (Palay & Chan-Palay, 1974).

4 Purkinje cells fit in a field, based on an estimate of the number that span a microzone (Dean et al, 2010). Estimates of the ratio of Golgi cells to Purkinje cells vary. '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). Estimates may be low where Golgi cells were identified by their large size, so that smaller cells at deeper level were not counted (Palay & 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, so 30 in an ensemble.

Contact by mossy fibres on a Golgi cell is on basal dendrites. 4–6 basal dendrites arise from the Golgi cell soma. If we take it as 5, 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. It has been estimated that a Golgi cell 'cannot receive more than 10–12 MF [mossy fibre] inputs, given the total basolateral dendrites length and the density of glomeruli' (Pietrajtis & Dieudonné, 2013). Without more data it is challenging to estimate how many make contact. We feel two per dendrite (10 between an average of 5 dendrites) is low. Dendrites radiate from the cell body. A single dendrite has a range of 50–100 µm (Palay & Chan-Palay, 1974). The longest dimension of a field – which contains ~700 mossy fibre terminals – is 200 µm; on the other hand, there is little dendritic branching, so a dendrite occupies a path not a volume. We take the number as 20, 4 per dendrite.

Golgi cells inhibit granule cells in a structure termed a glomerulus (Eccles et al, 1964; Hámori & Szentágothai, 1966; Palay & Chan-Palay, 1974). Mossy fibre terminals are ensheathed in a semi-permeable membrane which restricts neurotransmitter diffusion. Excitation of granule cells by mossy fibres and inhibition by Golgi cells takes place here. Most of the Golgi cell axonal plexus – which is confined to the granular layer – is made up of very fine beaded fibres. These enter glomeruli, where they contact granule cell dendrites (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]

As far as is known, inhibition received by a glomerulus is from a random sample of Golgi cells afferent to a field. The convergence ratio – the mean number of Golgi cells with convergent input to a single glomerulus – is unknown, but somewhere in the range from nil to the number of Golgi cells whose axonal fields overlap at a point location. It depends on the ratio of Golgi cells to glomeruli and the divergence ratio, the mean number of glomeruli that receive innervation from a single Golgi cell. The divergence ratio is not known either but it is high. Each Golgi cell inhibits a large number of granule cells (Barmack & Yakhnitsa, 2008; D'Angelo et al, 2013). A single Golgi cell contacts thousands of granule cells (Hámori & Szentágothai, 1966). We take the convergence ratio as a random variable in the range 8–12:1. That is, we assume each glomerulus receives inhibition for around 1/3 of Golgi cells afferent to a field. (We test the effect of other ratios in Fig 3.)

3.2 Methods

This section describes the Golgi cell ensemble computation. It is thought that neuronal firing rates provide the physiological form of values that are used in network computations. We argue in this section that transmission through a Golgi cell ensemble is in biophysical steps that can be represented as mathematical operations. Each step is compounded of a large number of near-in-time computations. Information is represented at ensemble-wide level.

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

[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.

411 result. However, if the new distribution is then itself randomly sampled, and the process
412 is repeated, the results get tidier and narrower.

413 To apply that to Golgi cell ensembles, the first-sampled population is of mossy fibre rates
414 received by three fields. Three fields contain a total of 30 Golgi cells. Rates are sampled
415 by Golgi cell basal dendrites. At five dendrites per cell, the number of samples is 30 x 5 =
416 150. Sample size is the number of mossy fibres which make contact per dendrite, which
417 we take (above) to be 4.

418 The postsynaptic result – charge influx to dendrites – is a proportional function of input
419 rates. Otherwise, it would not convey rate information (Duguid et al, 2015). As far as we
420 know, there are no data for transformation of multiple mossy fibre input rates into Golgi
421 cell basal dendritic depolarisation. We take it as the mean. (We discuss evidence for
422 physiological conservation of rate information in section 3.5.2.) Taking the sample means
423 creates a new set of values (Fig 2 column 2). The frequency distribution of the new values
424 is in a more condensed range and centred on the mean of the sampled distribution. Basal
425 dendritic charge transfer is passive and little filtered (Cesana et al, 2010; Kanichay &
426 Silver, 2008). Integration of dendritic states at the Golgi soma is represented by taking
427 the mean of dendritic charge. This generates a third population of values (Fig 2 column
428 3). Golgi cell firing rates have a linear relationship with the amplitude of depolarising
429 current (Dieudonné, 1998). Thus, rate-proportional dendritic charge influx is converted
430 to Golgi cell firing rates.

431 The second stage of random sampling is of Golgi cell rates by glomeruli. Inhibition almost
432 exclusively by spillover (Brickley et al, 1996; Duguid et al, 2012; Nusser et al, 1998)
433 effectively increases sample size, as sample size is not the number of Golgi cells that
434 make direct contact on any particular granule cell, but the number afferent to a
435 glomerulus. Sample size is accordingly the number of Golgi cells with convergent input to
436 a glomerulus (taken as 8 to 12), and the number of samples is the number of glomeruli
437 per field (700). Rate-proportional spillover provides the physiological equivalent of
438 averaging afferent rates, represented by taking the sample means. This generates a final
439 population of values (Fig 2 columns 4 and 5), the output of the ensemble conversion.

440 3.3 Results

441 3.3.1 A single ensemble

442 The results for a single ensemble are shown in Figure 2. Neural information has a shifting
443 form. The ensemble conversion has intracellular as well as between-cell steps and
444 therefore transitions through forms that are normally measured on different scales. We
445 use a common scale because our interest is confined to the computational effect of
446 architecture through the steps of the conversion. Measuring data at different steps on
447 different scales would be arbitrary.

448 *Figure 2 about here please*

449 **3.3.1.1 Figure 2: the ensemble conversion**

450 Mossy fibre firing rates received by a rank were generated as described in section 2
451 within constraints on the frequency distribution (Fig 2, column 1). The distribution of
452 firing rates received by a row of three fields selected at random is shown (in red) in the
453 same panel.

454 In column 2, the red data are themselves randomly sampled – 150 samples, sample size
455 4, representing 30 Golgi cells x 5 basal dendrites, and contact on each dendrite by 4
456 mossy fibres – and the sample means plotted, representing dendritic charge states. The
457 range is narrower than the sampled population (in all rows), but data are not tightly
458 grouped.

459 In column 3, the sample means are themselves sampled (without replacement) – 30
460 samples, sample size 5, representing 30 Golgi cells with 5 dendrites each – and the
461 samples averaged (and plotted), representing somatic charge states. Somatic data are in
462 a narrower range than dendritic data but still not tightly grouped.

463 In the final step (column 4) somatic data are sampled 700 times, reflecting convergent
464 input from 3 fields of Golgi cells onto 700 glomeruli in the middle field. Sample size is 8–
465 12, randomly generated for each sample; physiological sample size is unknown. Column 5
466 shows the column 4 data as a histogram, for comparison with column 1.

467 The overall result is that mossy fibre signal frequencies received as input to a rank are
468 reliably converted to a narrow distribution of values representing inhibition of granule
469 cells. The same result is obtained regardless of the shape of the sampled distribution
470 (mossy fibre rates received by the rank as a whole). Ensembles convert input values to
471 synchronised inhibition of granule cells.

472 **3.3.1.2 Figure 3: unknown parameters**

473 We next explored the effect of varying model parameters for which physiological data
474 are unreported. We did that to verify that the conversion is not critically dependent on
475 our estimate of unknowns.

476 The results are shown in Figure 3. In 3A, the varied parameter is the number of mossy
477 fibres with convergent input to a single Golgi cell dendrite, and in 3B, it is the number of
478 Golgi cells with convergent input to a glomerulus. Both display the output of the
479 conversion, representing the mean of Golgi cell rates received by each glomerulus in the
480 field that receives ensemble output (so 700 values, the same as Figure 2, column 4). The
481 ratios are shown above each panel. In other figures, ratios are 4:1 (mossy fibres that
482 contact a Golgi cell basal dendrite) and 8–12:1 (Golgi cells that converge on a
483 glomerulus).

484 We found that increasing either ratio condenses the plotted data range (representing
485 tighter alignment of inhibition within a field), but there is a diminishing return. There is a
486 stronger effect of increasing a low ratio and a progressively weakening effect at higher
487 ratios. The only condition that returned a large range was a very low Golgi cell to
488 glomerulus ratio. So: varying unknown parameters did not reveal a reason to discard or
489 revise the model, provided that the physiological Golgi cell to glomerulus convergence
490 ratio is not unexpectedly low.

491 *Figure 3 about here please*

492 **3.3.2 A rank**

493 We simulated a sagittal row of 100 fields, a rank. The sampled distribution, representing
494 mossy fibre rates received by the whole row, was generated as described in section 2, as
495 before. This time we took output of the ensemble computation to be the mean of output
496 values (the 'glomerular field mean') – that is, for each field, the mean of values obtained
497 the same way as Figure 2, column 4. Within-field glomerular means are synchronised by
498 statistical constraint to a narrowly restricted range, so that taking the mean is a
499 reasonable measure to compare fields.

500 The results are shown in Figure 4. In Fig 4A and B, the top panel shows the mean of
501 mossy fibre rates received by each field in a row of 100, without copies (pink) and with
502 copies added ('mf field mean', red). Red data closely track the pink data, indicating that
503 adding copies does not skew the mean of firing rates received by a field.

504 The bottom long panel shows mf field mean (red) and glomerular field mean (blue), again
505 for 100 fields. As in Figure 2, the blue data are restricted to a focused range (the y axis
506 range is shorter than Figure 2). In Figure 4B, for example, SD of the blue data is ~3 Hz,
507 derived from mossy fibre rates with a range of ~190 Hz, SD 25 Hz. Within that range,
508 however, input and output vary randomly and independently from field to field, in a
509 bandwidth centred on the mean of mossy fibre rates received by the rank as a whole.

510 Panels 4C–E show the effect on the output of the rank conversion of changing the mean
511 of the sampled distribution without changing the shape or bandwidth.

512 Our findings are:

513 i) Field-to-field variability of the mf field mean is predominantly the result of the fact that
514 a field receives a random sample of mossy fibre rates, and not the effect of adding copies
515 to reflect terminal clustering;

516 ii) A rank of Golgi cells behaves like a single unit. Inhibition of the granule cell population
517 of a whole rank is tightly synchronised. We get to that conclusion in this way: glomerulus
518 field means are statistically constrained to a narrow range, and individual glomerulus
519 data (700 per field) are tightly grouped around the mean.

iii) The mf field mean and glomerular field mean were only related by having a narrowly constrained range with the same mean, also the mean of the sampled distribution (i.e., of mossy fibre rates received by the rank). This was unexpected. We expected noise to be compounded by the chain of chance inherent in a mechanism that exploits random sampling (twice). Instead, the mf field mean and the glomerular field mean varied independently (within a statistically constrained range).

iv) In all fields, the glomerular field mean was precisely, reliably and linearly coupled to the mean of the sampled distribution.

Summary. Our interpretation is that Golgi cells work together in functionally unitary sagittal strips. Inhibition of a rank is synchronised at an adjustable rate which is a linear function of the mean of mossy fibre rates received by the whole rank. Synchrony is an automated computational consequence of statistical effects of ensemble architecture.

Figure 4 about here please

3.4 Mossy fibre terminal clustering simulates sampling with replacement

Recoding exploits statistical effects (known as the central limit theorem) that depend on replacement. Samples taken with replacement are 'independent' – there is no effect of sampling on other samples. The physiological form of sampling is anatomical. A sample is made up of the firing rates of cells that contact a sampling unit – in this case a basal dendrite (not the whole cell). For independence, each sample must be equally likely to contain any rate in the sampled distribution. This is not anatomically achievable if mossy fibres each ended in a single terminal since clearly a single terminal cannot be in more than one place at once. Mossy fibre termination in terminal clusters, and random intermingling of terminals (Fig 1), means that a signal can be received by more than one target. Moreover, at local scale there is a near equal probability that this happens, because a sampling unit falls within or passes through a large majority of local cluster fields. It is a biological make-do, but the outcome approximates sampling with replacement.

Clustering itself creates a potential problem, however. Physiological contact must simulate sampling of mossy fibres afferent to a field, not terminals. That is, for the mechanism to work, sampling must be of rates received by a field, not signals (which include duplicates). To approximate the result of sampling rates only, a sampling unit must not receive contact from the same mossy fibre more than once. A result close to that is obtained with small sample size. It is an imperfect solution because it does not fully eliminate duplication, but duplication is unlikely, so rare. It is intuitive that small samples would contain duplication less often, and we can show mathematically with granule cells that sample size of 4 is small enough. If we assume that each dendrite of a granule cell has an equal chance of contact with each of m x 100 mossy fibre terminals, the probability of an nth contact with the same mossy fibre 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 granule cell dendrites and m is the number of terminals per cluster. Calculated in this way, around 95% of granule cells receive input to each dendrite from a different mossy fibre, and less than two in a thousand receive more than one duplicate of the same signal.

So: the conversion of mossy fibre rates into inhibition of granule cells in functionally unitary strips depends on anatomical mimicry of sampling with replacement. Sampling units are single Golgi cell dendrites, not the whole cell (so the number of samples per ensemble territory is 150 not 30). The sampled population is the population of mossy fibre input rates *to a whole rank*. Mossy fibre terminal branching and terminal clusters together have this result. The same statements apply to granule cells with the difference that the sampling unit is a whole cell.

4.GRANULE CELLS

We next incorporated the Golgi cell ensemble conversion into a model that recodes mossy fibre signals into granule cell signals. It has been mooted that inhibition might be used to select (and change) the permutation of mossy fibre inputs to a granule cell (Marr, 1969), so that a single granule cell could in theory code information combined in different ways, or to adjust the number of inputs needed for a granule cell to fire (Albus, 1971), as part of a feedback-driven regulatory mechanism. We propose it has both a coding and regulatory function, and the outputs and mechanism are different from traditional proposals. **Coding.** Data are coded in the frequency distribution of granule cell rates; mean, shape and spread of the distribution are predicted by mossy fire rates; ranks operate as a single functional unit – any random sample of signals contained in the output of a whole rank codes the same information. **Regulation.** The function of the Golgi cell ensemble conversion is to maintain at a fixed number the minimum number of competitive mossy fibre signals needed for a single granule cell to fire, and thereby at a fixed proportion the number that do; it is principally this (and not feedback) that auto-regulates the fraction of granule cells that are active; regulation is local, operative at field level, so the fraction of granule cells that are active is the same in all fields.

Note: the idea that there may be an all-or-nothing result of the intraglomerular 'balance' of excitation and inhibition (section 4.3) has appeared previously (Gilbert & Miall, 2021a).

591 The idea that active granule cells may accordingly receive input rates with a narrowed
592 bandwidth at the higher end of the range also appears there.

593 4.1 Physiology

594 Derivation of computationally-relevant Golgi cell ensemble parameters was covered
595 earlier. This section briefly covers the derivation of additional parameters relevant to the
596 recoding computation.

597 Granule cells receive input to each dendrite from a single mossy fibre. Granule cells have
598 3–5 dendrites, averaging 4 (the simulation assumes 4, but the result is the same with 3–
599 5). Contact is made in a glomerulus, as noted. Each dendrite of a single granule cell
600 extends into a different glomerulus. Estimates vary of the number of granule cells that
601 receive contact from a single terminal. 50 is mid-range (Gao et al, 2016; Jakab & Hámori,
602 1988), although estimates vary (Ritzau-Jost et al, 2014). 50 is consistent with our earlier
603 estimates of the number of terminals and granule cells in a field. A granule cell receives
604 contact from a random sample of mossy fibres afferent to a field, and therefore receives
605 a random sample of mossy fibre rates. The sampling unit is the whole cell since it
606 receives only a single input rate to each dendrite. Sample size is therefore 4, and the
607 number of samples per field is the number of granule cells, which we estimate above at
608 8,750.

609 4.2 Methods

610 A granule cell needs a minimum number of active mossy fibre inputs to fire – thought to
611 be three (Billings et al, 2014; Jörntell & Ekerot, 2006). Certainly, one is too few
612 (Bengtsson & Jörntell, 2009; Chadderton et al, 2004; Duguid et al, 2012). We take it as
613 three. Three signals is not in itself sufficient, as each of those signals must also prevail
614 over spillover-mediated Golgi cell inhibition. Inhibition is provided by the Golgi cell
615 ensemble conversion, so it is synchronised in a rank. It is not exactly synchronous. A
616 single granule cell receives inhibition of each dendrite from an independent random
617 sample of Golgi cells afferent to the field. Nonetheless, it is in a focussed range so that all
618 mossy fibre signals compete against inhibition in a narrowly synchronous range.

619 In each glomerulus, excitation and inhibition compete for influence. Activation of
620 presynaptic $GABA_B$ receptors on mossy fibres inhibits glutamate release (Mitchell &
621 Silver, 2000a), while activation of presynaptic mGluR2 receptors on Golgi cells inhibits
622 GABA release (Mitchell & Silver, 2000b). The outcome of the competition is independent
623 of other mossy fibre terminals, even other terminals that provide input to the same
624 granule cell. It is unclear if there is always an emphatic winner, but it is possible, because
625 a swing in the dominant direction may be amplified by positive feedback. So, even if all
626 granule cells receive four active mossy fibre inputs, which we assume they do, only a
627 subset fire, those that receive at least three mossy fibre winners.

We term the mean of Golgi cell rates received by a single glomerulus the 'inhibitory threshold' and take it that a mossy fibre signal must exceed the inhibitory threshold to count towards the minimum number needed for a granule cell to fire. The simulation compares mossy fibre signal frequency with the inhibitory threshold at each of 4 dendrites of each of 8,750 granule cells in each of 100 fields. There are 700 inhibitory thresholds (outputs of the ensemble conversion) per field, derived in section 3, because there are 700 glomeruli. Excitation and inhibition are paired at random. The comparison is used to identify the subset of granule cells that fire (cells that receive at least 3 mossy fibre signals which exceed the inhibitory threshold). The simulation takes the mean of input rates received by each of those cells (the 'sample means'). Output of the model is accordingly the number of granule cells that fire in each field, the mean of input rates they each receive, and the average of the means for each field ('GC field mean'). The results are shown in Figures 5 and 6.

4.3 Results

We find:

i) The output of the recoding mechanism is the same in all fields in a rank (Fig 5 column 3, Fig 6 column 3). That is, at any moment, granule cell firing rates have the same frequency distribution. Accordingly, any random sample of the entire output of a rank has this distribution – a rank is functionally a single recoding unit.

ii) The frequency distribution of the sample means is narrower than the sampled distribution, normal, and has the same mean. The range of the sample means for active granule cells is even narrower because active cells receive mossy fibre rates in the narrowed range that exceeds the inhibitory threshold (Fig 5 column 2). The focus of the granule cell range depends on the range of the sampled distribution (Fig 6 column 2 rows 4–6).

iii) (i) and (ii) are independent of the sampled distribution (mossy fibre input rates to the rank as a whole) although a wide mossy fibre range weakens the granule cell focus.

iv) The mean of the granule cell frequency distribution follows the mean of the sampled distribution. Given a fixed shape and bandwidth of the mossy fibre distribution, it follows at a fixed distance (Fig 6 rows 1–3). Given a fixed shape and mean, increasing the mossy fibre range causes the same relative increase in the granule cell mean (Fig 6 rows 4–6).

Summary. Our interpretation is that recoding pixelates multidimensional mossy fibre input, converting it into a low-resolution internal code. This assumes that a granule cell fires at a rate that is a proportional function of the mean of input rates (see section 5.3 for a review of the evidence). Pixels are rank sized – there is no finer grade of resolution. Any random sample of signals inside pixel boundaries (over a modest minimum) codes the same information. There is a statistically conserved, reliable and precise relationship between the mean of mossy fibre rates received by the whole rank and the mean of a

random sample of granule cell rates. Auto-modulation of Golgi cell inhibition of granule cells means the number of granule cells that fire is independent of mossy fibre rates, and a stable percentage of the population.

Figures 5 and 6 about here

5. DISCUSSION

5.1 Executive summary

Neural network models often use acquired synaptic modification to implement computations. For example, learning algorithms use synaptic change to store experience. In the cerebellum, that has meant theories did not need to explain how granule cells code information, because control of Purkinje cells by rates was displaced by a learned response. We challenge that by proposing another way that input data are computed, and how physiology is adapted for that function, focussing on the granular layer of the cerebellar cortex.

We propose that Golgi cells work together as 3-field functional units – ensembles. Ensembles are not discrete but represent a minimum functional scale. Inhibition is synchronised between fields in a rank, so that granule cells in a rank receive synchronised inhibition at an adjustable rate which scales proportionally with the mean of the sampled population (mossy fibre input rates to the rank as a whole).

The ensemble conversion forms part of the mechanism of recoding that converts mossy fibre signals into a well-predicted frequency distribution of granule cell rates. A rank is a functional unit. The mean firing rate of granule cells in all fields is the same linear function of the mean of the sampled distribution. The output of a rank is the smallest unit of granule cell encoded information but not the smallest unit of code. The same information is coded in any random sample of granule cell firing rates, anywhere inside a rank. The result is a low-resolution granule cell code of rank-sized pixels.

5.2 Functional anatomy

Recoding exploits reliable effects of statistics and probability at large numbers. The anatomy of the termination of a single mossy fibre means a rank receives multiple copies of each input signal at multiple locations. The functional driver is to simulate independent random sampling. Sampling is by granule cells and Golgi cell basal dendrites. The sampled population is of mossy fibre input rates to the rank as a whole. Simulated independence means that the number of duplicates of a signal reflects the distributed probability of selection in the hypothetical scenario that samples were taken one at a time and selected rates were returned to the sampled population – i.e.,

sampling is with replacement. Sampling with replacement has statistical effects called the central limit theorem. This allows mossy fibre signals to be converted to an internal code with a precise and reliable relationship to input rates although single cells individually receive far too few signals to read the code or to themselves individually code anything.

The function of Golgi cells is to convert mossy fibres rates to synchronised inhibition of an entire rank of granule cells. Inhibition is at an automatically adjusted rate which tracks mossy fibre rates. Tracking ensures that mossy fibre signals which compete (with inhibition) successfully are limited to a regulated fraction, and that a fixed minimum number of competitive signals are needed to make a granule cell fire. This has, in turn, two functions. One relates to the granule cell code and the other to regulation of the number of granule cells that fire. *Code*. Granule cell rates are constrained to a twice-narrowed range, first by the central limit theorem and second because active cells are limited to those that receive mostly mossy fibre rates which are themselves a subset of the sampled distribution. *Regulation*. Because the number of granule cells in a field is large, the proportions of active and inactive granule cells are well predicted by the probability that a given cell fires. That probability depends on the fraction of mossy fibre inputs to a field that are competitive, given a granule cell must receive a fixed minimum number of competitive signals to fire.

5.3 Pixels offer a substrate of control of Purkinje cells

A granule cell code in this form provides a plausible substrate to drive coherent firing of microzone-grouped cells downstream which populate a parallel strip. A microzone is around 100 Purkinje cells long and 4 wide, and lies at right angles across the path of parallel fibres, as noted. An estimated 350,000 parallel fibres pass through a single Purkinje cell territory of which around half make contact (Harvey & Napper, 1988; 1991). So, all Purkinje cells in a microzone receive contact from a different random sample of parallel fibres. Most of them receive none of the same signals. Even two Purkinje cells lying directly on beam receive contact from only (a number near to) half the same parallel fibres. However, since the output of recoding is pixelated, and all Purkinje cells randomly sample granule cell signals, all Purkinje cells receive the same code, with the same timing.

5.4 Evidence of linear transmission

5.4.1 Feed-forward inhibition of granule cells

Direct evidence for or against linear conservation of rate information at each step of Golgi cell internal transmission is not available at the biophysical scale, as far as we know. However, evidence is available at cell level.

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

Contact on Golgi cells by mossy fibres is multi-synaptic (Palay & Chan-Palay, 1974), contributing to a reliable, rapid (submillisecond) response (low quantal release per synapse but collectively low uncertainty). 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 & Silver, 2008), as expected for large-diameter dendrites. Contact is in glomeruli. A slow decay component associated with transmission by spillover is weak or absent, suggesting a modest spillover component of transmission. Possible reasons are individually modest release probability (DiGregorio et al, 2002) and lower-affinity AMPA receptors (DiGregorio et al, 2007).

'Sensory-evoked Golgi-cell inhibition [of a granule cell] scales proportionally with the level of mossy fiber excitatory synaptic input' to the same cell, such that inhibition reliably conveys mossy fibre rate information (Duguid et al, 2015). 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 & 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 & Blazis, 1989).

So: mossy fibre rate information is conserved proportionally in Golgi cell firing rates. We extrapolate that the steps in between are also linear, because it is more likely than a linear outcome of non-linear steps.

There is a further stage in the pathway to inhibition of granule cells. As noted, inhibition is almost exclusively by GABA spillover (Brickley et al, 1996; Duguid et al, 2012; Nusser et al, 1998). If spillover is proportional to Golgi cell rates, we would expect intraglomerular diffusion to provide the physiological equivalent of averaging, so that inhibition of granule cells reflects the mean of afferent rates.

5.4.2 Excitation of granule cells

If granule cells collectively code mossy fibre rate information in the form we propose, we would expect mossy fibre to granule cell transmission to be adapted to conserve rate

775 information. The short and equal length of granule cell dendrites would suggest light and
776 equal filtering. The mossy fibre-granule cell connection has multiple adaptations to
777 transmit high frequency signals (Delvendahl & Hallermann, 2016) across a wide
778 bandwidth. Faithful transmission of rate information is reported (Rancz et al, 2007;
779 Ritzau-Jost et al, 2014). Vesicle release and replenishment are fast (Saviane & Silver,
780 2006) and postsynaptic AMPA receptors operate in their linear range (Sargent et al,
781 2005), where they are resistant to desensitisation (DiGregorio et al, 2007). Conversion of
782 depolarising somatic charge into granule cell firing is a simple function such that granule
783 cells 'have a relatively linear and uncomplicated conversion of depolarisation level to
784 spike rate' (Bengtsson & Jörntell, 2009 p.2393, citing Jorntell and Ekerot 2006 and
785 D'Angelo et al 1998).

786 However, linear transmission is contentious because it has been reported that
787 transmission strength may be variable. There is variation between granule cells (and
788 trials) of EPSCs following mossy fibre stimulation *in vitro* (Sargent et al, 2005), leading to
789 the suggestion that fast vesicular release may allow precise signalling with
790 heterogeneous synaptic weights. However, in this study stimulation was typically not
791 above 2 Hz to avoid short-term depression observed at higher frequencies, and GABA
792 was blocked.

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

800 Still, there *are* weights. Different strength and short-term dynamics of the mossy fibre to
801 granule cell connection have been reported in vestibular slices, with $GABA_A$ receptors
802 blocked. The amplitude of the response, measured at the soma, is reported to be
803 statistically correlated to the source of the mossy fibre that is stimulated (Chabrol et al,
804 2015), suggesting that the strength of the connection depends on the source.

805 However, the presence of weights does not necessarily mean that the function is to
806 modulate firing rates. Indeed, the authors of the study don't suggest that, but propose
807 that it may temporally code the source of inputs because different combinations have a
808 different response onset time. An alternative, which they also mention, is that some
809 inputs may be in a supporting role, reminiscent of 'driver' and 'modulatory' signals in the
810 thalamus and cortex (Sherman & Guillery, 2011). Feasibly, signals in a supporting role
811 must be present for the postsynaptic granule cell to fire, but do not affect the granule
812 cell firing rate.

813 The evidence is inconclusive. There is plausible evidence for linear transmission.

5.5 Evidence of a sustained state of elevated activity

We simulate a moment in time, a snapshot of activity. It is intuitive to think of signals as discrete episodes with a duration. However, that would mean the combined effect of a set of signals would still include time-variable elements, because integration would partly depend on their relative timing and duration. We propose that the cerebellum side-steps that problem by primarily using continuous variables for internal signalling, and that Golgi cells and glomeruli are in a sustained state during periods of activity.

During behaviour in freely-moving animals, mossy fibre activity is dense: a high fraction of mossy fibres are active (Hartmann & Bower, 2001). Both mossy fibres (Arenz et al, 2008; Giovannucci et al, 2017; van Kan et al, 1993) and Golgi cells (Edgley & Lidierth, 1987) have been reported to fire with a sustained, time-varying signature, though it can depend on modality and the manner of stimulation (Holtzman et al, 2006). This would suggest – and we propose – that in the behaving animal Golgi cells receive time-varying but sustained input signals. The large Golgi cell soma, where dendritic signals are integrated, is in this contention in a state of continuously variable change. 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 & Lidierth, 1987). Golgi cells fire spontaneously (Ruigrok et al, 2011), at a rate that is modulated by somatic charge, so that downstream, a glomerulus receives convergent input at continuously variable rates. 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).

On the face of it, there could still be a timing issue. Signalling by spillover is sometimes assumed to be ambient and slow. However, the action of glomerular GABA spillover has a fast phasic component – not as fast as synaptic transmission (~1 ms) but with a rise time of only a few milliseconds (Mapelli et al, 2014). Unlike the spiky appearance of synaptically-induced inhibitory postsynaptic currents, spillover (Duguid et al, 2015) generates a sustained outward current. The response time to a changing rate of multiple active inputs should be faster than the experimentally isolated response to a discrete signal, but this is untested, as far as we know. Our provisional conclusion is that glomeruli do not cause a special timing issue.

5.6 A sparse granule cell code is economic and quiet

The recoding mechanism provides regulation of the fraction of active granule cells. In learning models, low density of active granule cells increases storage capacity for pattern memory. In our model, Purkinje cells are controlled by rates, not memory, so if we are correct, and density is low, it is for other reasons.

A low fraction saves energy. Granule cells account for more than half the cells in the brain. The axons of granule cells are unmyelinated. The energy cost of transmission is therefore higher than myelinated axons. A sparse code is accordingly potentially a large energy saving. Granule cells fire in short bursts. A single burst is at a single firing rate. Modulation of firing of parallel fibre targets is achieved by the turnover of active parallel fibres. A low fraction of granule cells firing in short bursts is a large saving.

Stability mitigates noise. Collective neural activity contains multiple variables all potentially with a competing influence on downstream targets. Non-coding variables are noise, potentially capable of corrupting transmission. If nothing is coded in the fraction of active granule cells, fluctuation of the number is noise. A stable regulated percentage of active granule cells therefore importantly eliminates a redundant variable.

Stability may be enhanced by granule cell ascending axon contact on Golgi cell basal dendrites (Cesana et al, 2010). A Golgi cell receives ~400 functional synaptic contacts from this source, around 25% of the granule cells that a Golgi cell receives contact from in total. 'During sustained mf [mossy fibre] discharge, synapses from local GrCs [granule cells] contribute equivalent charge to GoCs [Golgi cells] as mf synapses' and 'display similarly fast kinetics' (Cesana et al, 2013), providing the substrate for potent local feedback. Fast feedback would provide a mechanism to increase stability of the fraction of active granule cells at a local level.

5.7 Feedback inhibition of granule cells

Inhibition of granule cells probably receives modulation by feedback via parallel fibres received by dendrites which some Golgi cells extend into the molecular layer. It is a long-standing idea that feedback may contribute to self-regulation of the fraction of co-active parallel fibres (Albus, 1971; Marr, 1969). Our expectation is that the strength of an effect by this pathway is modest (because fine adjustments are sufficient for its function), consistent with a leading role in regulation of feed-forward inhibition (i.e., driven by direct mossy fibre input to basal dendrites).

Modulation of Golgi cell rates by single parallel fibre signals is weak (Dieudonné, 1998), although sub-linear integration is mitigated by gap junctions (Vervaeke et al, 2012). Also, only a modest number of parallel fibre signals are received at any time. A small fraction of parallel fibres which pass through a Golgi cell territory make contact (~1,200 out of a number in the hundreds of thousands (Cesana et al, 2013)). Of those, only some are active at any time, and if the parallel fibre code is sparse, it is a small fraction. For example, if 1% of parallel fibres are active, it is 12. As noted, not all Golgi cells have apical dendrites, although the axonal plexus of those cells does not differ from cells that do (Palay & Chan-Palay, 1974), so the collective effect of parallel fibre feedback is diluted because a significant population of Golgi cells don't receive it. Taken together, this would suggest (inconclusively) that the effect of parallel fibre inputs on Golgi cell firing is relatively weak, and control by direct input from mossy fibres is the main form of control.

We envisage that the feedback pathway makes fine, stabilising adjustments to an already sparse parallel fibre code.

5.8 Input from other sources?

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

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

Our alternative solution is that resolution of high-dimensional input data is sacrificed at a surprisingly profligate level. The cerebellum disregards source and modality data below a physical scale which mirrors topography. It is therefore unnecessary to discriminate below that scale. Mapping studies show that mossy fibre signals (evoked by cutaneous stimulation) share the topographical organisation of climbing fibres (Apps & Garwicz, 2005; Apps & Hawkes, 2009; Ji & Hawkes, 1994). At sub-topographical dimensions, signal source and type are simply discounted. This has the flexible advantage that networks do not need to be adapted to the source or type of input they receive – anatomically, networks are interchangeable, so network design can be modular.

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.

GC field mean The mean of input rates to active granule cells, calculated for each cell, averaged for each field.

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

mf field mean The mean of mossy fibre signal frequencies received as input to a field, including copies.

rank A sagittal row of 100 fields.

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 inputs needed for a granule cell to fire.

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.

946 **REFERENCES**

947 Albus, J. S. (1971) A Theory of Cerebellar Function. *Mathematical Biosciences*, 10(1-2), 25-61.
948 Albus, J. S. (1989) The Marr and Albus theories of the cerebellum-two early models of associative
949 memory, *Digest of Papers. COMPCON Spring 89. Thirty-Fourth IEEE Computer Society*
950 *International Conference: Intellectual Leverage*. Feb. 27 1989-March 3 1989.
951 Apps, R. & Garwicz, M. (2005) Anatomical and physiological foundations of cerebellar
952 information processing. *Nat Rev Neurosci*, 6(4), 297-311.
953 Apps, R. & Hawkes, R. (2009) Cerebellar cortical organization: a one-map hypothesis. *Nat Rev*
954 *Neurosci*, 10(9), 670-81.
955 Arenz, A., Silver, R. A., Schaefer, A. T. & Margrie, T. W. (2008) The contribution of single synapses
956 to sensory representation in vivo. *Science*, 321(5891), 977-80.
957 Barmack, N. H. & Yakhnitsa, V. (2008) Functions of interneurons in mouse cerebellum. *J Neurosci*,
958 28(5), 1140-52.
959 Bengtsson, F. & Jörntell, H. (2009) Sensory transmission in cerebellar granule cells relies on
960 similarly coded mossy fiber inputs. *Proc Natl Acad Sci U S A*, 106(7), 2389-94.
961 Billings, G., Piasini, E., Lorincz, A., Nusser, Z. & Silver, R. A. (2014) Network structure within the
962 cerebellar input layer enables lossless sparse encoding. *Neuron*, 83(4), 960-74.
963 Brand, S., Dahl, A. L. & Mugnaini, E. (1976) The length of parallel fibers in the cat cerebellar
964 cortex. An experimental light and electron microscopic study. *Exp Brain Res*, 26(1), 39-58.
965 Brickley, S. G., Cull-Candy, S. G. & Farrant, M. (1996) Development of a tonic form of synaptic
966 inhibition in rat cerebellar granule cells resulting from persistent activation of GABAA receptors. *J*
967 *Physiol*, 497 (Pt 3), 753-9.
968 Brunel, N., Hakim, V., Isope, P., Nadal, J. P. & Barbour, B. (2004) Optimal information storage and
969 the distribution of synaptic weights: perceptron versus Purkinje cell. *Neuron*, 43(5), 745-57.
970 Cesana, E., Dieudonné, S., Isope, P., Bidoret, C., D'Angelo, E. & Forti, L. (2010) Novel granule cell-
971 Golgi cell excitatory input in the cerebellar granular layer, *FENS Abstr*.
972 Cesana, E., Pietrajtis, K., Bidoret, C., Isope, P., D'Angelo, E., Dieudonne, S. & Forti, L. (2013)
973 Granule cell ascending axon excitatory synapses onto Golgi cells implement a potent feedback
974 circuit in the cerebellar granular layer. *J Neurosci*, 33(30), 12430-46.
975 Chabrol, F. P., Arenz, A., Wiechert, M. T., Margrie, T. W. & DiGregorio, D. A. (2015) Synaptic
976 diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nat*
977 *Neurosci*, 18(5), 718-27.
978 Chadderton, P., Margrie, T. W. & Hausser, M. (2004) Integration of quanta in cerebellar granule
979 cells during sensory processing. *Nature*, 428(6985), 856-60.
980 D'Angelo, E., Solinas, S., Mapelli, J., Gandolfi, D., Mapelli, L. & Prestori, F. (2013) The cerebellar
981 Golgi cell and spatiotemporal organization of granular layer activity. *Front Neural Circuits*, 7, 93.
982 Dean, P., Porrill, J., Ekerot, C. F. & Jörntell, H. (2010) The cerebellar microcircuit as an adaptive
983 filter: experimental and computational evidence. *Nat Rev Neurosci*, 11(1), 30-43.
984 Delvendahl, I. & Hallermann, S. (2016) The Cerebellar Mossy Fiber Synapse as a Model for High-
985 Frequency Transmission in the Mammalian CNS. *Trends Neurosci*, 39(11), 722-737.
986 Dieudonné, S. (1998) Submillisecond kinetics and low efficacy of parallel fibre-Golgi cell synaptic
987 currents in the rat cerebellum. *J Physiol*, 510 (Pt 3), 845-66.

988 DiGregorio, D. A., Nusser, Z. & Silver, R. A. (2002) Spillover of glutamate onto synaptic AMPA
989 receptors enhances fast transmission at a cerebellar synapse. *Neuron*, 35(3), 521-33.
990 DiGregorio, D. A., Rothman, J. S., Nielsen, T. A. & Silver, R. A. (2007) Desensitization properties of
991 AMPA receptors at the cerebellar mossy fiber granule cell synapse. *J Neurosci*, 27(31), 8344-57.
992 Duguid, I., Branco, T., Chadderton, P., Arlt, C., Powell, K. & Hausser, M. (2015) Control of
993 cerebellar granule cell output by sensory-evoked Golgi cell inhibition. *Proc Natl Acad Sci U S A*,
994 112(42), 13099-104.
995 Duguid, I., Branco, T., London, M., Chadderton, P. & Hausser, M. (2012) Tonic inhibition enhances
996 fidelity of sensory information transmission in the cerebellar cortex. *J Neurosci*, 32(32), 11132-43.
997 Dumoulin, A., Triller, A. & Dieudonné, S. (2001) IPSC kinetics at identified GABAergic and mixed
998 GABAergic and glycinergic synapses onto cerebellar Golgi cells. *J Neurosci*, 21(16), 6045-57.
999 Eccles, J. C., Ito, M. & Szentágothai, J. (1967) *The cerebellum as a neuronal machine*. Berlin, New
1000 York etc.: Springer-Verlag.
1001 Eccles, J. C., Llinás, R. & Sasaki, K. (1964) Golgi cell inhibition in the cerebellar cortex. *Nature*, 204,
1002 1265-6.
1003 Edgley, S. A. & Lidierth, M. (1987) The discharges of cerebellar Golgi cells during locomotion in
1004 the cat. *J Physiol*, 392, 315-32.
1005 Ekerot, C. F. & Jörntell, H. (2001) Parallel fibre receptive fields of Purkinje cells and interneurons
1006 are climbing fibre-specific. *Eur J Neurosci*, 13(7), 1303-10.
1007 Eyre, M. D. & Nusser, Z. (2016) Only a Minority of the Inhibitory Inputs to Cerebellar Golgi Cells
1008 Originates from Local GABAergic Cells. *eNeuro*, 3(2).
1009 Fujita, M. (1982) Adaptive filter model of the cerebellum. *Biol Cybern*, 45(3), 195-206.
1010 Galliano, E., Mazzarello, P. & D'Angelo, E. (2010) Discovery and rediscoveries of Golgi cells. *J*
1011 *Physiol*, 588(Pt 19), 3639-55.
1012 Gao, Z., Proietti-Onori, M., Lin, Z., Ten Brinke, M. M., Boele, H. J., Potters, J. W., Ruigrok, T. J.,
1013 Hoebeek, F. E. & De Zeeuw, C. I. (2016) Excitatory Cerebellar Nucleocortical Circuit Provides
1014 Internal Amplification during Associative Conditioning. *Neuron*, 89(3), 645-57.
1015 Garwicz, M., Ekerot, C. F. & Jorntell, H. (1998) Organizational principles of cerebellar neuronal
1016 circuitry. *News in Physiological Sciences*, 13, 28-34.
1017 Gilbert, M. & Miall, R. C. (2021a) How and Why the Cerebellum Recodes Input Signals: An
1018 Alternative to Machine Learning. *Neuroscientist*, DOI: 10.1177/1073858420986795.
1019 Giovannucci, A., Badura, A., Deverett, B., Najafi, F., Pereira, T. D., Gao, Z., Ozden, I., Kloth, A. D.,
1020 Pnevmatikakis, E., Paninski, L., De Zeeuw, C. I., Medina, J. F. & Wang, S. S. (2017) Cerebellar
1021 granule cells acquire a widespread predictive feedback signal during motor learning. *Nat*
1022 *Neurosci*.
1023 Hámori, J. & Szentágothai, J. (1966) Participation of Golgi neuron processes in the cerebellar
1024 glomeruli: an electron microscope study. *Exp Brain Res*, 2(1), 35-48.
1025 Hartmann, M. J. & Bower, J. M. (2001) Tactile responses in the granule cell layer of cerebellar
1026 folium crus IIa of freely behaving rats. *J Neurosci*, 21(10), 3549-63.
1027 Harvey, R. J. & Napper, R. M. (1988) Quantitative study of granule and Purkinje cells in the
1028 cerebellar cortex of the rat. *J Comp Neurol*, 274(2), 151-7.
1029 Harvey, R. J. & Napper, R. M. (1991) Quantitative studies on the mammalian cerebellum. *Prog*
1030 *Neurobiol*, 36(6), 437-63.

1031 Holtzman, T., Rajapaksa, T., Mostofi, A. & Edgley, S. A. (2006) Different responses of rat
1032 cerebellar Purkinje cells and Golgi cells evoked by widespread convergent sensory inputs. *J*
1033 *Physiol*, 574(Pt 2), 491-507.
1034 Hull, C. & Regehr, W. G. (2012) Identification of an inhibitory circuit that regulates cerebellar
1035 Golgi cell activity. *Neuron*, 73(1), 149-58.
1036 Isope, P. & Barbour, B. (2002) Properties of unitary granule cell-->Purkinje cell synapses in adult
1037 rat cerebellar slices. *J Neurosci*, 22(22), 9668-78.
1038 Jakab, R. L. & Hámori, J. (1988) Quantitative morphology and synaptology of cerebellar glomeruli
1039 in the rat. *Anat Embryol (Berl)*, 179(1), 81-8.
1040 Jelitai, M., Puggioni, P., Ishikawa, T., Rinaldi, A. & Duguid, I. (2016) Dendritic excitation-inhibition
1041 balance shapes cerebellar output during motor behaviour. *Nat Commun*, 7, 13722.
1042 Ji, Z. & Hawkes, R. (1994) Topography of purkinje cell compartments and mossy fiber terminal
1043 fields in lobules ii and iii of the rat cerebellar cortex: Spinocerebellar and cuneocerebellar
1044 projections. *Neuroscience*, 61(4), 935-954.
1045 Jörntell, H. & Ekerot, C. F. (2006) Properties of somatosensory synaptic integration in cerebellar
1046 granule cells in vivo. *J Neurosci*, 26(45), 11786-97.
1047 Kanichay, R. T. & Silver, R. A. (2008) Synaptic and cellular properties of the feedforward inhibitory
1048 circuit within the input layer of the cerebellar cortex. *J Neurosci*, 28(36), 8955-67.
1049 Lange, W. (1974) Regional differences in the distribution of golgi cells in the cerebellar cortex of
1050 man and some other mammals. *Cell Tissue Res*, 153(2), 219-26.
1051 Llinás, R. R. & Negrello, M. N. (2015) *Cerebellum*, 2015. Available online:
1052 http://www.scholarpedia.org/article/Cerebellum [Accessed.
1053 Mapelli, L., Solinas, S. & D'Angelo, E. (2014) Integration and regulation of glomerular inhibition in
1054 the cerebellar granular layer circuit. *Front Cell Neurosci*, 8, 55.
1055 Marr, D. (1969) A theory of cerebellar cortex. *J Physiol*, 202(2), 437-70.
1056 Miles, F. A., Fuller, J. H., Braitman, D. J. & Dow, B. M. (1980) Long-term adaptive changes in
1057 primate vestibuloocular reflex. III. Electrophysiological observations in flocculus of normal
1058 monkeys. *J Neurophysiol*, 43(5), 1437-76.
1059 Mitchell, S. J. & Silver, R. A. (2000a) GABA spillover from single inhibitory axons suppresses low-
1060 frequency excitatory transmission at the cerebellar glomerulus. *J Neurosci*, 20(23), 8651-8.
1061 Mitchell, S. J. & Silver, R. A. (2000b) Glutamate spillover suppresses inhibition by activating
1062 presynaptic mGluRs. *Nature*, 404(6777), 498-502.
1063 Moore, J. W. & Blazis, D. E. J. (1989) 11 - Simulation of a Classically Conditioned Response: A
1064 Cerebellar Neural Network Implementation of the Sutton–Barto–Desmond Model, in Byrne, J. H.
1065 & Berry, W. O. (eds), *Neural Models of Plasticity*Academic Press, 187-207.
1066 Mugnaini, E. (1983) The length of cerebellar parallel fibers in chicken and rhesus monkey. *J Comp*
1067 *Neurol*, 220(1), 7-15.
1068 Nusser, Z., Sieghart, W. & Somogyi, P. (1998) Segregation of different GABAA receptors to
1069 synaptic and extrasynaptic membranes of cerebellar granule cells. *J Neurosci*, 18(5), 1693-703.
1070 Oscarsson, O. (1979) Functional units of the cerebellum - sagittal zones and microzones. *Trends in*
1071 *Neurosciences*, 2, 143-145.
1072 Ozden, I., Sullivan, M. R., Lee, H. M. & Wang, S. S. (2009) Reliable coding emerges from
1073 coactivation of climbing fibers in microbands of cerebellar Purkinje neurons. *J Neurosci*, 29(34),
1074 10463-73.
1075 Palay, S. L. & Chan-Palay, V. (1974) *Cerebellar cortex : cytology and organization*. Berlin: Springer.

1076 Pietrajtis, K. & Dieudonné, S. (2013) Golgi Neurons, in Manto, M., Schmahmann, J. D., Rossi, F.,
1077 Gruol, D. L. & Koibuchi, N. (eds), *Handbook of the Cerebellum and Cerebellar Disorders*.
1078 Dordrecht: Springer Netherlands, 829-852.
1079 Rancz, E. A., Ishikawa, T., Duguid, I., Chadderton, P., Mahon, S. & Hausser, M. (2007) High-fidelity
1080 transmission of sensory information by single cerebellar mossy fibre boutons. *Nature*, 450(7173),
1081 1245-8.
1082 Rasmussen, A., Zucca, R., Jirenhed, D. A., Johansson, F., Ortenblad, C., Svensson, P. & Hesslow, G.
1083 (2014) Golgi cell activity during eyeblink conditioning in decerebrate ferrets. *Cerebellum*, 13(1),
1084 42-5.
1085 Raymond, J. L. & Medina, J. F. (2018) Computational Principles of Supervised Learning in the
1086 Cerebellum. *Annu Rev Neurosci*, 41, 233-253.
1087 Ritzau-Jost, A., Delvendahl, I., Rings, A., Byczkowicz, N., Harada, H., Shigemoto, R., Hirrlinger, J.,
1088 Eilers, J. & Hallermann, S. (2014) Ultrafast action potentials mediate kilohertz signaling at a
1089 central synapse. *Neuron*, 84(1), 152-163.
1090 Ros, H., Sachdev, R. N., Yu, Y., Sestan, N. & McCormick, D. A. (2009) Neocortical networks entrain
1091 neuronal circuits in cerebellar cortex. *J Neurosci*, 29(33), 10309-20.
1092 Rossi, D. J. & Hamann, M. (1998) Spillover-mediated transmission at inhibitory synapses
1093 promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry. *Neuron*,
1094 20(4), 783-95.
1095 Rossi, D. J., Hamann, M. & Attwell, D. (2003) Multiple modes of GABAergic inhibition of rat
1096 cerebellar granule cells. *J Physiol*, 548(Pt 1), 97-110.
1097 Ruigrok, T. J., Hensbroek, R. A. & Simpson, J. I. (2011) Spontaneous activity signatures of
1098 morphologically identified interneurons in the vestibulocerebellum. *J Neurosci*, 31(2), 712-24.
1099 Sargent, P. B., Saviane, C., Nielsen, T. A., DiGregorio, D. A. & Silver, R. A. (2005) Rapid vesicular
1100 release, quantal variability, and spillover contribute to the precision and reliability of
1101 transmission at a glomerular synapse. *J Neurosci*, 25(36), 8173-87.
1102 Saviane, C. & Silver, R. A. (2006) Fast vesicle reloading and a large pool sustain high bandwidth
1103 transmission at a central synapse. *Nature*, 439(7079), 983-7.
1104 Schonewille, M., Gao, Z., Boele, H. J., Veloz, M. F., Amerika, W. E., Simek, A. A., De Jeu, M. T.,
1105 Steinberg, J. P., Takamiya, K., Hoebeek, F. E., Linden, D. J., Huganir, R. L. & De Zeeuw, C. I. (2011)
1106 Reevaluating the role of LTD in cerebellar motor learning. *Neuron*, 70(1), 43-50.
1107 Sherman, S. M. & Guillery, R. W. (2011) Distinct functions for direct and transthalamic
1108 corticocortical connections. *J Neurophysiol*, 106(3), 1068-77.
1109 Shinoda, Y. & Sugihara, I. (2013) Axonal Trajectories of Single Climbing and Mossy Fiber Neurons
1110 in the Cerebellar Cortex and Nucleus, in Manto M., S. J. D., Rossi F., Gruol D.L., Koibuchi N. (ed),
1111 *Handbook of the Cerebellum and Cerebellar Disorders*Springer, Dordrecht.
1112 Shinoda, Y., Sugihara, I., Wu, H. S. & Sugiuchi, Y. (2000) The entire trajectory of single climbing
1113 and mossy fibers in the cerebellar nuclei and cortex. *Prog Brain Res*, 124, 173-86.
1114 Sultan, F. & Heck, D. (2003) Detection of sequences in the cerebellar cortex: numerical estimate
1115 of the possible number of tidal-wave inducing sequences represented. *J Physiol Paris*, 97(4-6),
1116 591-600.
1117 ten Brinke, M. M., Boele, H. J., Spanke, J. K., Potters, J. W., Kornysheva, K., Wulff, P., AC, I. J.,
1118 Koekkoek, S. K. & De Zeeuw, C. I. (2015) Evolving Models of Pavlovian Conditioning: Cerebellar
1119 Cortical Dynamics in Awake Behaving Mice. *Cell Rep*, 13(9), 1977-88.

van Kan, P. L., Gibson, A. R. & Houk, J. C. (1993) Movement-related inputs to intermediate cerebellum of the monkey. *J Neurophysiol*, 69(1), 74-94.

Vervaeke, K., Lorincz, A., Gleeson, P., Farinella, M., Nusser, Z. & Silver, R. A. (2010) Rapid desynchronization of an electrically coupled interneuron network with sparse excitatory synaptic input. *Neuron*, 67(3), 435-51.

Vervaeke, K., Lorincz, A., Nusser, Z. & Silver, R. A. (2012) Gap junctions compensate for sublinear dendritic integration in an inhibitory network. *Science (New York, N.Y.)*, 335(6076), 1624-1628.

Vos, B. P., Volny-Luraghi, A. & De Schutter, E. (1999) Cerebellar Golgi cells in the rat: receptive fields and timing of responses to facial stimulation. *Eur J Neurosci*, 11(8), 2621-34.

Walter, J. T. & Khodakhah, K. (2006) The linear computational algorithm of cerebellar Purkinje cells. *J Neurosci*, 26(50), 12861-72.

Walter, J. T. & Khodakhah, K. (2009) The advantages of linear information processing for cerebellar computation. *Proc Natl Acad Sci U S A*, 106(11), 4471-6.

Welsh, J. P., Yamaguchi, H., Zeng, X. H., Kojo, M., Nakada, Y., Takagi, A., Sugimori, M. & Llinas, R. R. (2005) Normal motor learning during pharmacological prevention of Purkinje cell long-term depression. *Proc Natl Acad Sci U S A*, 102(47), 17166-71.

Wu, H. S., Sugihara, I. & Shinoda, Y. (1999) Projection patterns of single mossy fibers originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei. *J Comp Neurol*, 411(1), 97-118.

1141 **Fig 1. Mossy fibre termination in a rank.** There is random interdigitation of branch
1142 endings of mossy fibres co-afferent to a rank, and at local level random intermingling of
1143 terminals of mossy fibres co-afferent to a field. **(A)** Sagittal section of the rat cerebellum
1144 (obtained by multi-photon microscopy, courtesy Tom Deerinck). The outer layer, the
1145 molecular layer, of the cerebellar cortex is stained green and the inner layer, the granular
1146 layer, is blue. **(B)** Output of a procedure used to group mossy fibre terminal into clusters
1147 (Sultan & Heck, 2003 fig 6), reproduced with permission. All terminals belong to one
1148 mossy fibre, which terminates in two strips. Clusters are shown in different shades of
1149 grey. **(C)** Schematic of an area of the folded cerebellar cortex. The strip inside the red
1150 dotted line is the width and depth of a microzone but only part of its length. Termination
1151 of mossy fibre signals is not at a point but at several randomly spaced, sagittally-aligned
1152 locations. Yellow zones are terminal locations of a single mossy fibre. Size of locations
1153 varies depending partly on the number of terminals in each cluster. Pink zones are
1154 terminal locations of a second mossy fibre. In reality, zones overlap extensively and
1155 terminals are intermingled. A mossy fibre mediolateral axon may give rise to several
1156 sagittal branches, of which here only one is shown. The zone enclosed by a blue line has
1157 the dimensions of a 'field'. A field is a nominal division of the granular layer with the
1158 average dimensions of a mossy fibre terminal cluster, which we use to quantify granular
1159 layer computations. A sagittal row of 100 fields is a 'rank'. **(D)** Section of a microzone
1160 viewed obliquely from the cerebellar surface, to illustrate the size and random
1161 intermingling of mossy fibre terminals. The molecular layer is represented by squares of
1162 about the dimensions of a Purkinje cell territory. The wire frame boxes are an exploded
1163 view of the orange fields. The right box shows randomly distributed terminals of 5 mossy
1164 fibres (each a different colour). We estimate that ~180 mossy fibres contribute terminals
1165 to a field, and a field contains a total of around 700 terminals. Proportions of terminal
1166 size to field size are preserved. The left box shows terminals of 20 mossy fibres.

1167

1168 **Fig 2. Golgi cell ensemble conversion. Column 1** We randomly generated (3,645) mossy
1169 fibre rates with a uniform, normal, twin-peaked or discontinuous distribution
1170 representing input to a row of 100 fields. Rates were randomly sampled to represent
1171 input to individual fields and copies added to reflect variable size of terminal clusters and
1172 that most clusters straddle field boundaries. Red data represent copies-included signals
1173 received by a row of three fields. **Column 2** The first-column red data were randomly
1174 sampled 150 times, sample size 4, representing mossy fibre input to a 3-field population
1175 of 30 Golgi cells, each with 5 basal dendrites (the mean), and each of those receiving
1176 contact from 4 mossy fibres. Data are the sample means, representing dendritic
1177 depolarisation. **Column 3** We next randomly sampled column 2 data without
1178 replacement 30 times, sample size 5, and took the sample means, representing somatic
1179 integration of dendritic signals. **Column 4** We then took 700 random samples of column 3
1180 data, sample size 8–12, reflecting convergent input from 3 fields of Golgi cells onto 700
1181 glomeruli in the middle field (to which a three-field ensemble has output). Data are the
1182 sample means. **Column 5** The column 4 data are presented as a histogram.

1183

1184 **Fig 3. Testing unreported parameters shows that reliability and precision are chiefly**
1185 **impacted at different steps in the computation.** All data were obtained in the same way
1186 as Fig 2 column 4, but we tested the effect of varying ratios for which physiological
1187 parameters have not been reported. Graphed data are accordingly the means of Golgi
1188 cell input rates to each of 700 glomeruli, the population of the middle field of a Golgi cell
1189 ensemble territory. Broken dark blue line: the mean of the means. The sampled
1190 distribution was generated in the same way as Fig 2 row 4. (The sampled distribution is
1191 the population of mossy fibre input rates to a row of 100 fields which includes the
1192 ensemble territory.) Yellow line: the mean of the sampled distribution. Varied ratios are
1193 the convergence ratio of mossy fibres onto a single Golgi cell basal dendrite (4:1 in Fig 2)
1194 and the convergence ratio of Golgi cells onto a single glomerulus is (a random number in
1195 the range 8–12:1 in Fig 2). In **(A)** we varied the first and in **(B)** the second. The ratios are
1196 shown above each panel. y axes cover a shorter range than Fig 2. A narrower range
1197 represents tighter within-field synchrony of inhibition received by granule cells. Findings
1198 were: (1) In both cases, there is a diminishing return of increasing the ratio. There is a
1199 modest return of higher ratios than Fig 2, selected to be physiologically plausible. (2) The
1200 blue mean varies in a small range around the yellow mean, but within that, varies at
1201 random each time the simulation is run. In (A), it varies from panel to panel because
1202 mossy fibre input rates to the simulated ensemble territory are resampled each time. In
1203 (B), that step in the computation is performed only once and the output used for all
1204 panels. Here, the blue mean is the same in all panels, but moves randomly when the
1205 simulation is rerun, in the same range as (A). So, reliability (reproducibility of the
1206 glomerular field mean) is impacted at the step of sampling of mossy fibre rates by Golgi
1207 cells, but not sampling of Golgi cell rates by glomeruli. Precision (narrowness of the
1208 glomerular range) is impacted by both ratios, but chiefly at the second step.

1209

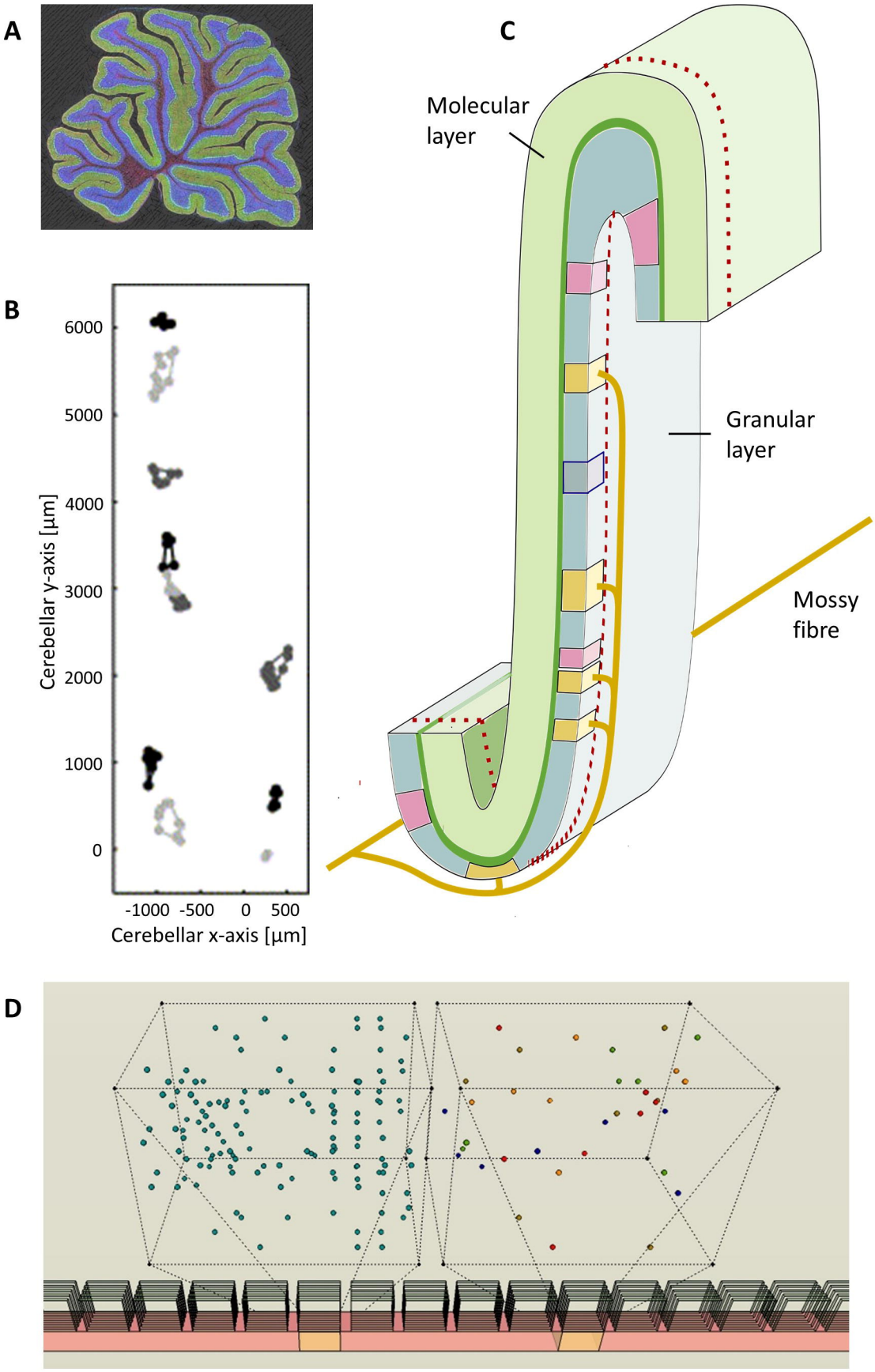
1210 **Fig 4. A whole rank of Golgi cells work as a functional unit. Synchronised inhibition of**
1211 **granule cells is precisely and linearly related to the mean of mossy fibre rates. (A)** and
1212 **(B)** Simulation of conversion by Golgi cells of mossy fire signals into inhibition of granule
1213 cells in a sagittal row of 100 fields, taking the mossy fibre distribution on the left as input.
1214 The top panel compares the mean of mossy fibre input rates to each field (pink) with the
1215 mean with copy signals added to reflect mossy fibre terminal clustering (red). The dashed
1216 grey line is the mean of the sampled distribution (i.e., the mean of mossy fibre input
1217 rates to the whole rank). The bottom panel compares the copies added data (red) with
1218 the glomerular field mean (blue line). SD of inhibitory thresholds is shown in pale blue.
1219 The y scale of the long panels is Hz. Results are: (1) In all top panels, red data closely track
1220 the pink data, indicating that adding copies does not skew the mean of firing rates
1221 received by a field. (2) Blue data across all fields are in a narrowly circumscribed range.
1222 The degree of precision depends on the bandwidth of the sampled distribution although
1223 the conversion works with any input distribution. A normal mossy fibre distribution with
1224 a 2SD range of over 100 Hz is converted to a blue range of ~3 Hz. (3) Fields that sample
1225 the same input data (that data being: mossy fibre input rates to a rank) return the same
1226 output. This has the physiological counterpart that inhibition of granule cells is
1227 synchronised in a rank. (4) The blue range is centred on the grey dashed line – the mean
1228 of mossy fibre rates received by the rank as a whole – and varies independently of the
1229 red data. That is, field-to-field variation of mean mossy fibre input rates doesn't flow
1230 through to an effect on blue data, the output of the system. (5) We also tested the effect
1231 of setting sample size 'optimally' (i.e., higher, at 24:1) for mossy fibres that contact a
1232 single Golgi cell dendrite, and Golgi cells that inhibit a single glomerulus. This only very
1233 modestly improved performance: higher ratios do not return a notably narrower blue
1234 range (than the range when parameters are set at known or expected physiological
1235 values). The results are not shown. **(C)–(E)** The same simulation, sampling the mossy
1236 fibre distribution on the left. The blue data are the glomerular field means, i.e., mean
1237 inhibition of granule cells in each field (inhibition of granule cells is condensed into a
1238 small range by the ensemble conversion, so we use the means to compare fields). y axes
1239 of panels on the right of each pair have the same range as x axes of panels on the left.
1240 Results: inhibition of granule cells is precisely and linearly coupled to the mean of the
1241 sampled distribution.

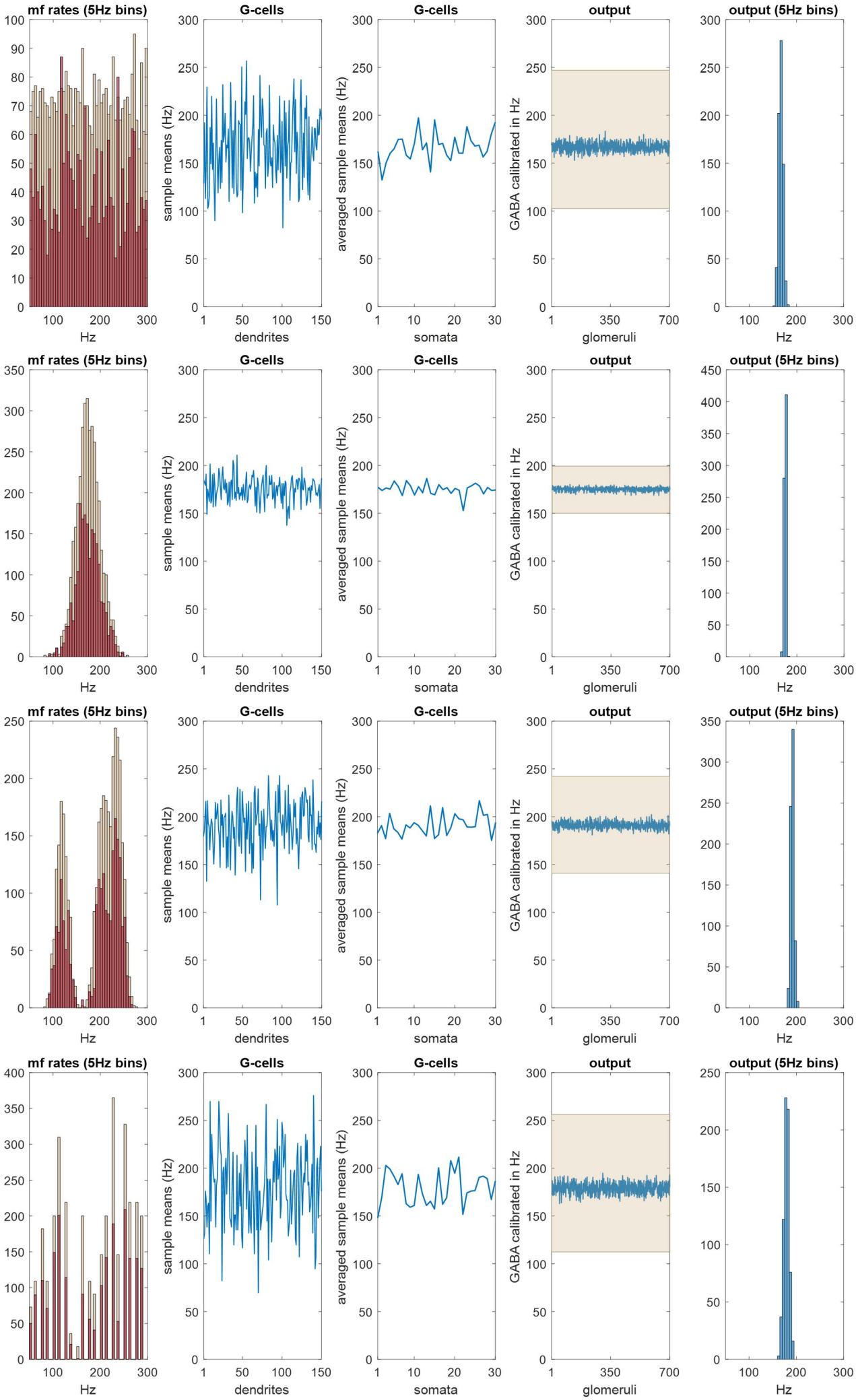
1242

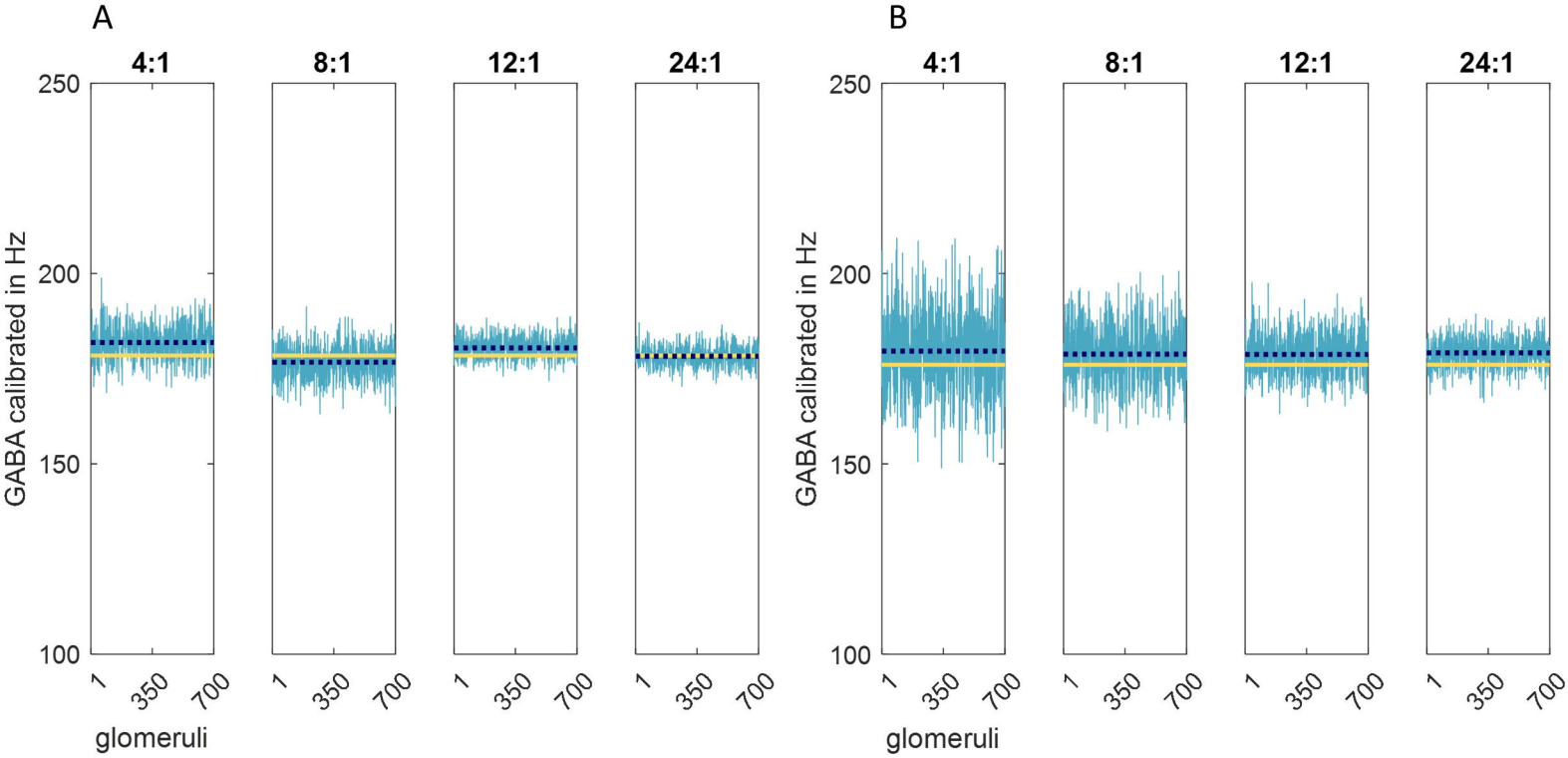
1243

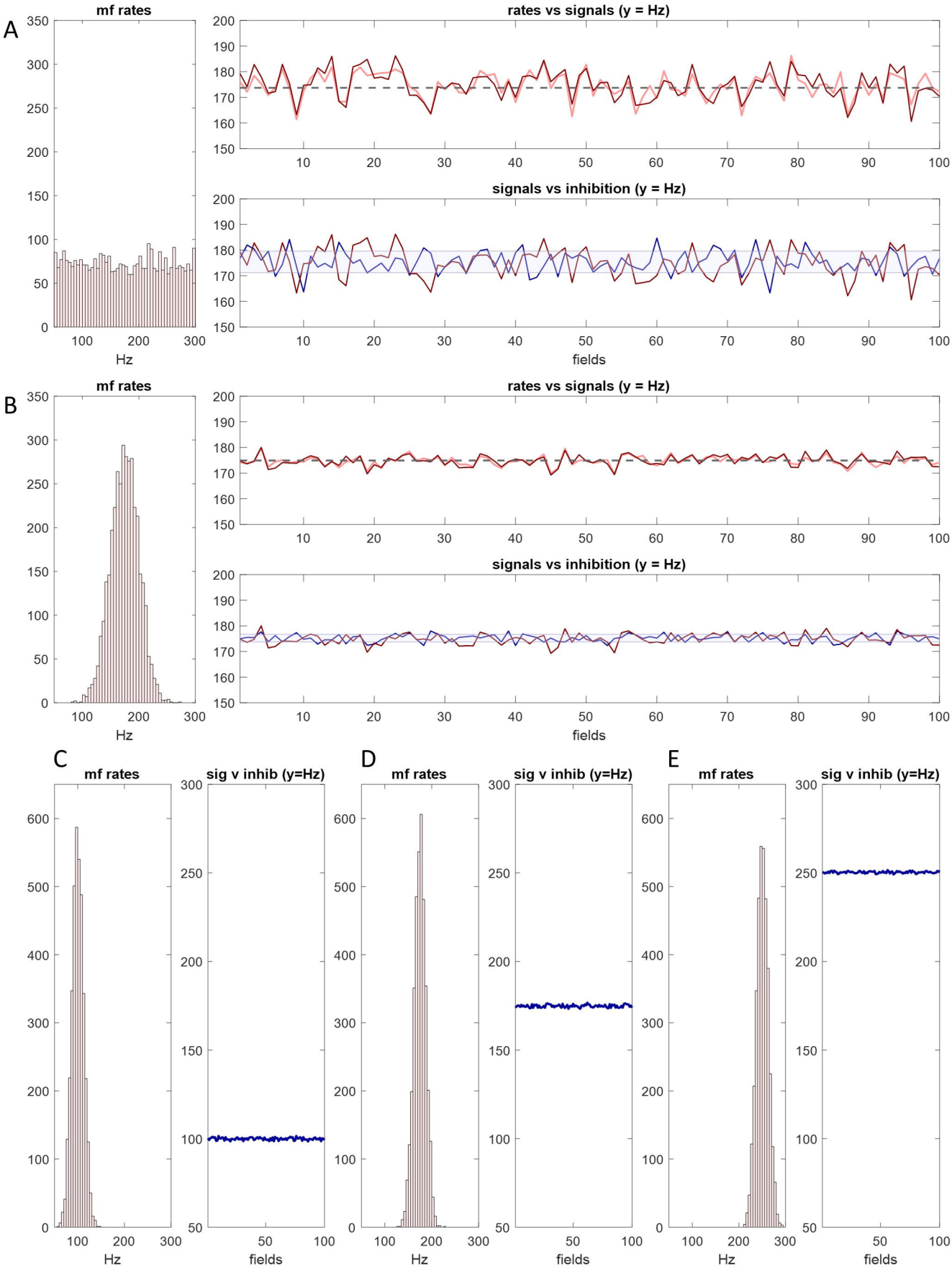
Fig 5. Incorporating the Golgi cell conversion into a model that recodes input to the cerebellum into internal signals. We incorporated the Golgi cell ensemble conversion into a simulation of recoding in the cerebellar granular layer. Recoding converts input to the cerebellum from mossy fibres into firing of granule cells. **Column 1.** We randomly generated (3,645) mossy fibre rates with a uniform, discontinuous or normal distribution representing 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 field (field 2, selected at random), with copies added, as an example. y axes are scaled to the data. **Column 2.** Red data for each field were randomly sampled 8,750 times, sample size 4, representing a per field population of 8,750 granule cells, each with 4 dendrites, which each receive contact from a single mossy fibre terminal. The field 2 sample means are displayed as a histogram (pale green, top three rows only; there are the same number of green data in each panel). Blue data represent the subgroup of granule cells that fire. In the top three rows no adjustment is made to inhibition of granule cells. In the bottom three rows it is adjusted (by adding a constant) to generate sparse code, i.e., a low percentage (around 1%) of active granule cells. Note that y axes in the bottom three rows are rescaled. **Column 3.** The y scale is Hz. The mean (line) and SD (shading) of the red and blue data for each field in a row of 100 fields. Blue line data are GC field means. GC field means are limited to a statistically severely constrained range regardless of the shape and range of the sampled distribution, although a narrower sampled range yields greater precision. The significance is that a whole rank recodes as a single unit. Any random sample of granule cells has the same mean and range, which is centred on the mean, and has a frequency distribution which is 'pulled' towards normal. **Column 4.** The percentage of granule cells that fire in each field. The white line and shaded area in the top three panels are mean and SD respectively. In the sparse code condition, the percentage of active granule cells in each field is limited to a severely constrained range.

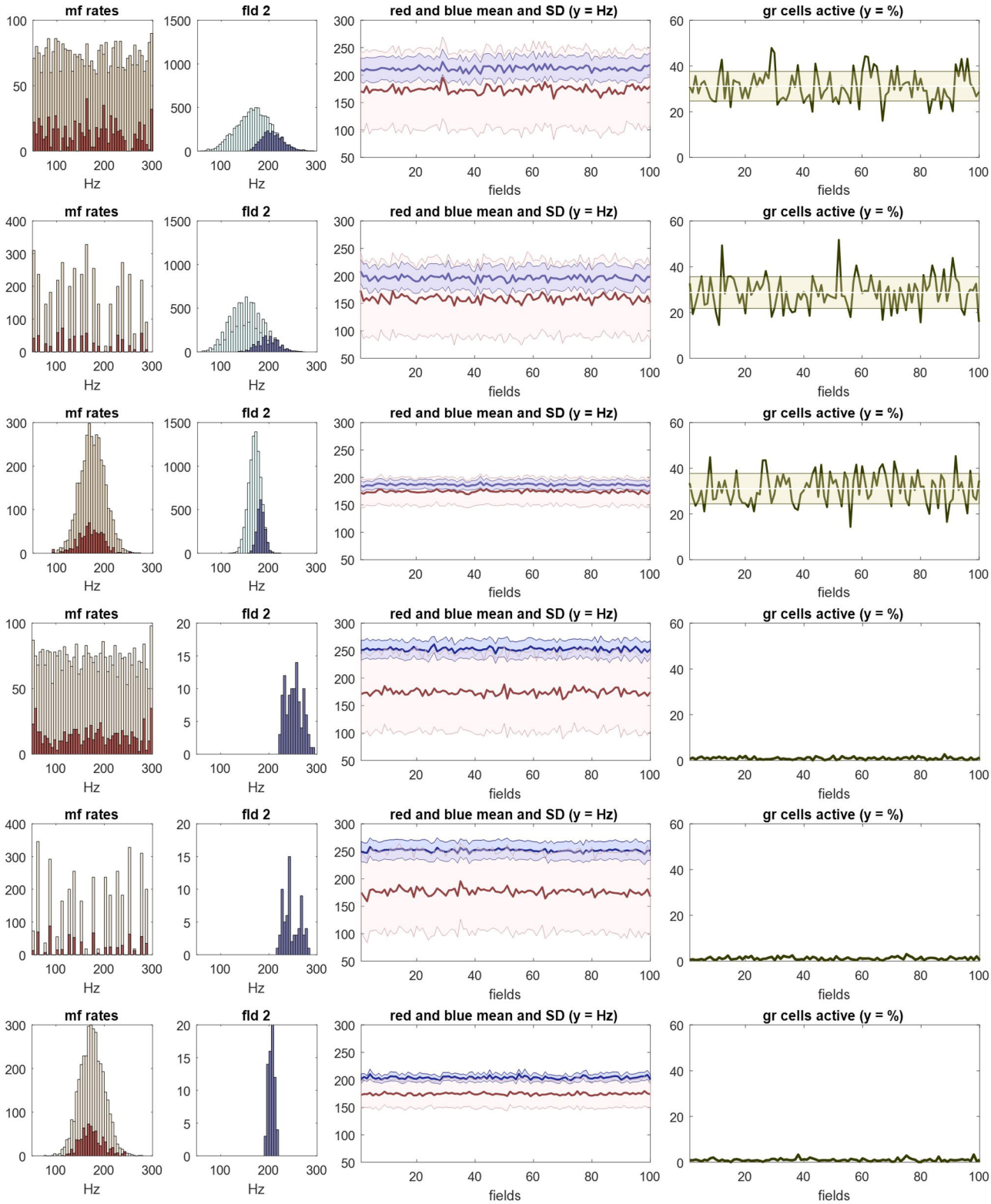
Fig 6. Granule cell rates collectively code mossy fibre rate information. Column 1 We randomly generated (3,645) mossy fibre rates per section 2 with either a normal (SD 12.5) or uniform distribution representing input to a row of 100 fields, with copies added – the sampled distribution. **Column 2** Frequency distribution of the mean of mossy fibre input rates to each of the granule cells that fire in a randomly selected field, obtained with the recoding model 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: GC field means. **Column 4** The percentage of active granule cells in each field. **(A)–(C)** The blue data are precisely and linearly related to the mean of mossy fibre rates. Blue data track the mean of the sampled distribution at a mean distance of 15 Hz (SD 1.21 Hz) in all rows. This relationship is independent of mossy fibre rates if the shape and bandwidth of the distribution of mossy fibre rates are approximately fixed. There is also no effect of changing mossy fibre rates on the percentage of granule cells that fire. **(D)–(F)** If instead the shape and mean of the sampled distribution are unchanging, there is a linear relationship with the bandwidth. Increasing the bandwidth causes the same relative increase in the mean distance between the mossy fibre mean and GC field means. There is no effect on the percentage of granule cells that fire.

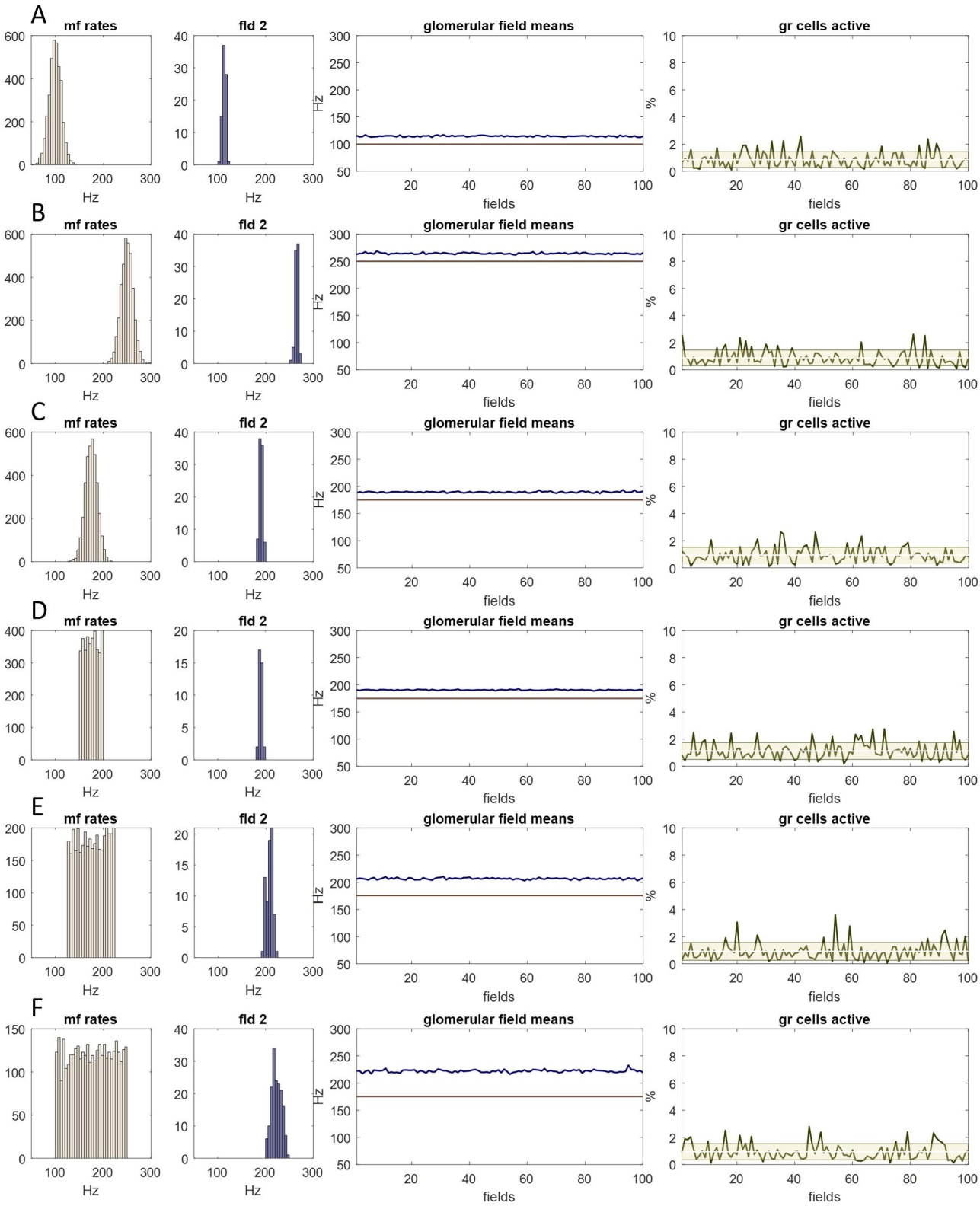

A
B
C
D
Molecular layer
Granular layer
Mossy fibre
Cerebellar y-axis [µm]
Cerebellar x-axis [µm]
6000
5000
4000
3000
2000
1000
0
-1000
-500
0
500


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


A
B
4:1
8:1
12:1
24:1
250
200
150
100
GABA calibrated in Hz
1
350
700
glomeruli


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


mf rates
fld 2
red and blue mean and SD (y = Hz)
gr cells active (y = %)
Hz
fields


A
mf rates
fld 2
glomerular field means
gr cells active
Hz
fields
%
B
C
D
E
F