Breaking the code – pixels, a surprising solution to neural network problems: the cerebellar input layer

Running head: Breaking the code: the cerebellar input layer

Mike Gilbert[a]* and 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

* Corresponding author; email: m.gilbert.1@bham.ac.uk

ABSTRACT

The cerebellum makes up 20–30 % of the brain by mass and volume in most vertebrates. 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 has ultimately failed to corroborate the predicted role of learning.

We ask, with a focus on the input layer of the cerebellar network:

1) If the agent of change is not synaptic memory, what is it? What is the physiological mechanism of the cerebellar network computation?

2) What is the computation, and how is information represented by internal signals activity – 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?

Our approach is to propose a reinterpretation of the rich anatomical and biophysical literature and test it with computer modelling. Modelling provides a test because information about computationally-relevant parameters are detailed, so we can replicate

networks in high resolution, including anatomical noise. Using this method, tested *in silico*, we found the evidence plausibly supports controversial conclusions. Using the same numbering:

1) The input layer computation is the functional result of cell morphologies and network architecture. Passing signals through a network provides the biological equivalent of mathematical operations, which run in series and in parallel.

2) Information is coded in statistics of internal firing rates. At this collective level, there are reliable statistical effects that conserve rate information between network layers.

3) Still more unexpectedly, multi-dimensional detail contained in input signals is converted to an extremely low resolution – pixelated – internal code. However, as a result, the internal code is intelligible by randomly sampling signals. That means downstream targets can receive different signals but the same code, providing an otherwise elusive mechanism to coordinate firing in microzones.

The proposals are bespoke because cell morphologies and the connections between them are unique to the cerebellum. However, the idea that anatomy and 3-dimensional network architecture implement mathematical operations unaided, if it holds water, is likely to have implications for coding in other brain regions.

Keywords: theory, computational neuroscience, cerebellum, network, code

53 # 1. INTRODUCTION

54 Most of the cerebellum by mass and volume is made up of the cerebellar cortex, which is
55 divided into two layers, an inner layer, the granular layer, and outer layer, the molecular
56 layer. The main source of glutamatergic input to the cerebellum is from mossy fibres.
57 Mossy fibres terminate in the granular layer, on granule cells and on Golgi cells, which
58 inhibit granule cells. The granule cell axon rises into the outer layer where it divides in
59 two to form parallel fibres. Parallel fibres lie parallel to the cerebellar surface and each
60 other. Parallel fibres contact Purkinje cells, which carry the only output of the cerebellar
61 cortex. Our focus is the granular layer.

62 In the traditional cerebellar model, the supervised learning model (1, an excellent
63 review), control of Purkinje cells by granule cell rates is displaced by learning coded as
64 parallel fibre synaptic weight changes (2-5). After training, when a Purkinje cell receives a
65 remembered pattern, it fires at a learned rate. It is implicit that a granule cell code is
66 redundant because control is surrendered to learning. Supervision is provided by
67 instruction signals which would, in principle, permit Purkinje cells to be trained and
68 controlled individually in this way.

69 We challenge that. Modelling physiological networks presents generic problems. It is
70 thought to be a safe starting point that the variable rate at which neurons discharge
71 action potentials – firing rates, sometimes called signal frequency – provides the
72 physiological equivalent of numbers that are used to code information, and that neural

networks provide the physiological implementation of computations which take firing rates as inputs. Relating that to physiology is not straightforward. Networks can be large and typically contain multiple cell types. Cell ratios – the ratio of cells of each type to other types – all differ. So do convergence and divergence ratios for each type onto other types they contact. Physiological networks are usually not anatomically separate, often sharing cells, sometimes most of them. Insentient networks don't 'know' what data values are a measure of, or what scale they are measured on, especially problematic if signals are mixed up. Biology is messy: most network parameter values are randomly variable.

One modelling approach is to simplify the parameters – for example, replacing cells with generic 'units'. Another is to model in abstract – how might an artificial system perform the computation a network may have? There is a theoretical reliance on synaptic weights to provide the mechanism of the network conversion. This has become unsafe. In the cerebellum, experimental evidence has ultimately failed to provide support for the expected role of long-term depression in motor learning (6-8) or for finely-graded modification of synaptic weights predicted by learning models (9, 10). There is accumulating evidence of linear transmission of rate-coded data, in conflict with models that use experience to train synaptic weight adjustments. For example, cerebellar neurons have been repeatedly found to fire at a rate that is a linear function of behavioural and task metrics, and this 'has been found at all levels of the cerebellar circuit' (1 p.239, citing 30 sources).

We asked, focussing on the input layer of the cerebellar network:

1) If the agent of the network conversion is not synaptic weights, what is the mechanism and what is the computation? If transmission is linear, what is the computational effect of cell and network anatomy?

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 functional anatomy? Also, how is it functionally related to the well-established organisation of Purkinje cells into microzones (11-13)?

Thinking about it in this way mitigates issues inherent in generic networks because the model is not generic. Anatomical data are rich. If anatomy implements computations, then knowing model parameters, we can model it. We are able to include detail that makes cells and networks unique, including random variability of parameter values (noise).

We find, using the same numbering:

1) Anatomy – cell morphologies and network input layer architecture – feasibly provides the biological equivalent of mathematical operations. Computational performance relies on linear transmission.

2) Granule cell rate information is collectively coded in statistical properties (mean and frequency distribution) of the firing rates of co-active cells. Code in this form means granule cells can read mossy fibre rates by taking random samples, and the granule cell code can in turn be read by randomly sampling granule cell rates. It also means mossy fibre rate information is conserved by a statistical relationship of external and internal signals. The relationship is independent of the permutation of active inputs and the permutation of the rates they each fire at. Anatomical noise does not impair computational performance.

3) Still more surprisingly, high-dimensional input data are converted to a low-resolution internal code. Information has spatial dimensions. The smallest unit of internal information is a strip with microzone dimensions (viewed from the cerebellar surface) but fuzzier boundaries. In this sense, the internal code is pixelated. The same information is contained in any random sample of signals inside those boundaries, however. This provides a substrate for mixed-up, multi-dimensional input signals to drive coherent firing of microzone-grouped cells downstream.

Since anatomy is the key instrument of information processing in this contention, and provides model parameters, each main section of the paper begins with a summary of the literature. Evidence of linear transmission is also discussed in context.

The structure of the paper follows the steps of the computation. It is accordingly in three parts headed ‘Mossy fibres’, ‘Golgi cells’ and ‘Granule cells’. The second and third parts

include a results section and our interpretation of the results. The theoretical nature of the work means there is no methods section. Our method is: (1) Propose the steps of the network input layer computation. (2) Present evidence that neurophysiology implements the steps. (3) Test it with a computer program that replicates the computationally relevant metrics of anatomy, and includes noise.

'Model' can be ambiguous. In artificial neural networks, for example, it is a mathematical statement of the proposal – the computation is synonymous with the model. *That is not our method.* In our approach, modelling provides a way to quantify the proposals so that we can test the computational effect of physiological parameters, if they are related the way we claim they are. We can do that in high-resolution detail because there is detailed anatomical information available. The steps of the computation are straightforward. The main function of computer modelling is a test.

The proposition of linear transmission of rate information comes with an advantage. Information transitions through several biophysical forms in transit even through a single cell. This has tended to divide modelling approaches between internal steps of signalling, to account for biophysical data, and network level signalling. Bringing them together is a knotty problem partly because different states are measured on different scales. Linear transmission makes the problem tractable. If firing rates are linearly related it is a plausible proposition that biophysical transitions are a chain of functionally linear steps regardless of the values they would take on a conventional scale. Indeed, non-linear

interim steps are improbable. Linear rate relationships accordingly provide a basis to represent all steps on a common scale. Linear intracellular transmission adds steps to the network computation. It forms part of our proposal that linear biophysical transitions exist for this computational effect.

2. MOSSY FIBRES

2.1 Contents

A single mossy fibre has several, sagittally-aligned branch endings, each ending in a cluster of terminals. A sagittal strip of the granular layer therefore receives multiple copies of each mossy fibre signal at multiple locations. As far as we know, cerebellar theory has not previously addressed this.

We derive in this section the number mossy fibres that innervate a strip, the number that innervate defined local sectors, and the ratios of the number of those that contribute a single terminal, two and so on. That is, the number of input rates received be each sector, and the number of copies of each rate, taking into account random variation of the number of terminals per cluster, and the fact that most clusters fall partly inside and partly outside a sector. This information is used to generate mossy fibre signals that are received as inputs to the simulations of Golgi cell computations and recoding of mossy fibre into granule cell signals in sections 3 and 4. It is essential to very precisely simulate

anatomy to test that real networks can execute the computational function we say they have (and that the computation is noise-tolerant).

This reveals that the important function of terminal branching is to simulate sampling with replacement by sectors of mossy fibre input rates to a strip as a whole. Simulated replacement has the computational role in coding that sectors each receive the same mean of rates as the sampled distribution, notwithstanding relatively modest per sector sample size.

2.2 Anatomy

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 (14-17).

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 (18-21). Terminal clusters are sagittally aligned and separated by a randomly variable distance typically not less than 200 µm (18, 19) measured from edge to edge of clusters (so they are separated by a greater distance measured from centre to centre). Terminals are intimately and randomly intermingled with terminals of other mossy fibres.

191 So, mossy fibres do not terminate at a point but a region with dimensions: the field
192 occupied by a terminal cluster. Cluster size is variable (because it depends on the number
193 of terminals and how tightly clustered they are), averaging 200 µm sagittally x 150 µm
194 mediolaterally (19), a 'field' (a sector in section 2.1). A sagittal row of 75–100 fields (a
195 'rank') has the area of a microzone viewed from the cerebellar surface. Microzone
196 dimensions vary, but a mediolateral span of 150 µm is mid-range (5, 11-13). Where
197 microzones are functionally defined by the topography of termination of climbing fibres
198 on Purkinje cells, a rank is a division of the granular layer and has less sharply defined
199 boundaries. Mossy fibre input to the cerebellum evoked by cutaneous stimulation largely
200 shares the topographical organisation of climbing fibres (22-24).

201 *Figure 1 about here please*

202 2.3 Derivation of the size of the sampled population and sample size

203 Fields randomly sample mossy fibre input rates to a rank, we propose. To simulate that,
204 we need the size of the sampled population – the number of mossy fibre signals received
205 by a rank – and sample size, the number received by a field.

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

210 The number of mossy fibres afferent to a field has been estimated at 100 (19). However,
211 that number is a field-sized fraction of the general population. The actual number is
212 larger because mossy fibres end in a dispersed group (cluster) of terminals, and most
213 groups straddle field limits.

214 We work from an estimate of the total number of terminals per field, which is, in turn,
215 derived from convergent estimates of the number of granule cells per field. (Estimate 1)
216 There are an estimated 1.92 x 10^6 granule cells per µl (14). The number of fields which fit
217 in that volume is $(1{,}000/150)^2$ x (1,000/200) = ∼222, assuming a field depth of 150 µm
218 (19). The number of granule cells in a field is, therefore 1,920,000/222 = 8,649. (Estimate
219 2) An estimated 350,000 parallel fibres pass through a single Purkinje cell's territory (14).
220 Parallel fibres run about 3 mm in each direction from the granule cell soma (16, 17). The
221 Purkinje cell dendritic arbour fills the molecular layer vertically and averages ∼200 µm
222 wide. Therefore, 6,000/150 = 40 fields fit in a row that provides parallel fibres that pass
223 through a Purkinje cell dendritic territory, giving 350,000/40 = 8,750 granule cells per
224 field.

225 Granule cells have 3–5 dendrites, averaging 4, and receive input to each of them from a
226 single mossy fibre. Estimates of divergence of a single mossy fibre terminal onto granule
227 cells vary (25). It is unclear how much this reflects biological diversity versus differences
228 in opinion and methods. Here we use a mid-range estimate of 1:50 (26, 27). Assuming a
229 mossy fibre terminal receives a single granule cell dendrite from each of d granule cells,

230 and given an average of 4 dendrites per granule cell and z granule cells per field, the total
231 number of mossy fibre terminals in a field is N = (z x 4)/d = 700 given z = 8,750 and d =
232 50. A convergent estimate is provided by multiplying the per field number of mossy
233 fibres by the average number of terminals per cluster.

234 Cluster size varies. We take the range as 4–12 terminals per cluster (19). We assume the
235 number varies at random and cluster sizes occur with equal probability. A maximum of 7
236 fits in a field (because a field shares the dimensions of the area that encloses a mean
237 cluster). Fields are not anatomically bounded, so many clusters straddle field limits. To
238 reflect that, we assume there is a 1/7 probability that a 7-terminal cluster (or larger) has
239 7 terminals inside a field and also the same probability of any other number. A 6-terminal
240 cluster has a 2/7 chance that all 6 terminals are inside and a 1/7 chance of any other
241 number, and so on.

242 Table 1 shows the derivation of the expected (i.e., predicted by probability) ratios of
243 terminals contributed to a field by each cluster size (R1), not counting terminals that lie
244 outside when a cluster straddles field boundaries; and the ratios of mossy fibres
245 contributing n = 1–7 terminals (R2), regardless of cluster size. R1 ratios are given by the
246 sum of numbers in each row. So, for example, the ratio of cells that end in 8 terminals to
247 cells that end in 5 is 28:25. R2 ratios are given by the number of appearances of each
248 number in the grid – so the expected ratio of cells which contribute 7 terminals to the
249 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

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

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.

2.4 Derivation of mossy fibre input rates to a field

Terminal clustering means a field receives multiple copies of most signals. This section lists the steps in the simulation that generate mossy fibre firing rates received by a field. As noted, at local level, anatomical parameters are random variables. Random variables are randomly generated (in steps 1, 2 and 4 below) with the stated probabilities.

In the first step, the simulation randomly generates, for each field, the cluster size ratio. This is the ratio 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

295 probability. In step two, the ratio of cluster sizes is weighted to reflect the greater
296 probability that large clusters contribute more terminals to a field (with probabilities
297 reflecting the expected ratio of terminals contributed by each cluster size). In step 3,
298 ratios are converted to the fraction of terminals contributed by each cluster size. When
299 multiplied by the total number of terminals, this gives the number of terminals
300 contributed by each cluster size. In step four, the values obtained in step three are
301 divided up between cells that contribute a single terminal, 2 terminals, 3 and so on, to
302 reflect straddling of field limits. The proportions give the number of mossy fibres afferent
303 to a simulated field that contribute a single terminal, two, three and so on up to seven,
304 the maximum.

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

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

314

2.5 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 the granular layer computation exploits statistical effects which require that a Golgi cell at any location of a rank must randomly sample the entire distribution of mossy fibre input rates to the whole rank. The same statement applies to granule cells. There is an additional requirement. This is that samples are independent, meaning, all samples have an equal probability that they contain any given rate picked from the sampled distribution. If samples were being drawn from a bag of coloured balls, balls would need to be replaced between samplings, so that, as samples were withdrawn, it did not change what was left in the bag. Hence sampling 'with replacement'. The physiological equivalent of random sampling is contact from a random sample of the afferent cell population. All samples are 'selected' concurrently. There is no replacement. Physiology must therefore simulate replacement. It is worth stating clearly that this is not a simplification but a hypothesis.

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 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 would be limited just as surely if there was actual replacement as it is in reality by anatomy.

That means 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 (19, 28), noting that data are limited. Estimated by this method, the mean is 4–5. Put another way, the anatomical cap

352 on the number of fields that can receive a signal (set by the number of terminal branches
353 of that cell) is reflected in the simulation by the probability distribution. The simulation
354 reflects anatomy and anatomy simulates replacement.

355 2.6 The range of mossy fibre firing rates

356 Some very high mossy fibre rates can be generated in experimental conditions. However,
357 it is thought the typical physiological range is 50–300 Hz (29) citing (30). In the
358 simulations, rates are generated at random in that range, within stated constraints on
359 the shape of the distribution. Figures display the sampled distribution.

360

361 3. GOLGI CELLS

362 3.1 Contents

363 We propose a functional organisation of Golgi cells into 3-field groups. Ensembles
364 implement a computational conversion of mossy fibre firing rates received by an
365 ensemble territory into synchronised inhibition of granule cells in the middle field.
366 Ensembles are not anatomically bounded and merge functionally into sagittal strips so
367 that the unit of conversion is a rank. Ensemble-grouped Golgi cells all innervate
368 substantially the whole of the middle field: convergence onto the middle field defines the
369 group.

370 ‘Input’ is to an ensemble territory – i.e., three fields – which is sampled by Golgi cells.
371 Output is inhibition of middle-field granule cells. More exactly, there are 700 outputs, the
372 number of glomeruli, that mediate inhibition. Some of the steps of the computation are
373 intracellular. It forms part of the proposal that in physiological practice the only variables
374 with significant influence are ensemble architecture and firing rates. A linear relationship
375 of input and output is evidenced. Anatomy provides parameters and the values they
376 take, next.

377 **3.2 Anatomy**

378 Anatomy implements ensemble computations. We test that by replicating an ensemble
379 *in silico* using computationally relevant metrics of anatomy. Anatomical evidence
380 provides parameter values, normally a range.

381 Golgi cells are large interneurons whose cell bodies lie in the granular layer. They are
382 most concentrated just below the Purkinje cell layer (31). The Golgi cell axon branches
383 profusely, giving rise to a dense plexus. The axonal field is sagittally elongated, that is, in
384 the direction of the long axis of a microzone – mean range 650 +/- 179 µm by 180 +/- 40
385 µm (in mice: 32) – and is the depth of the granular layer, which it fills vertically. Their
386 density (33) and large, profusely ramified axonal field (32) means that Golgi cell axonal
387 territories overlap extensively in the sagittal direction. The Golgi cell population of a row
388 of three fields all inhibit approximately the whole of the middle field. Convergence of
389 their axonal fields defines the size and territory of the smallest functional Golgi cell unit –

390 an ensemble – we suggest. Their combined range is longer – five fields. But convergence
391 on the middle field is functionally significant because it means all granule cells in the
392 middle field receive inhibition from a random sample of a three-field unit.

393 There is a significant population of Golgi cells at deeper level. Deeper cells tend to have
394 smaller cell bodies and to lack dendrites that extend into the molecular layer. However,
395 the axons of smaller Golgi cells are indistinguishable from large Golgi cells (31).

396 4 Purkinje cells fit in a field, based on an estimate of the number that span a microzone
397 (5). Estimates of the ratio of Golgi cells to Purkinje cells vary. 'In man, the ratio [of
398 Purkinje cells to Golgi cells] is 1:1.5, while in the monkey and cat it is almost 1:1.9 and in
399 the rat 1:3.3' (34). Estimates may be low where Golgi cells were identified by their large
400 size, so that smaller cells at deeper level were not counted (31 p.121). We take it that
401 there are 10 Golgi cells per field, a ratio of 1:2.5 with 4 Purkinje cells, so 30 in an
402 ensemble.

403 Contact by mossy fibres on a Golgi cell is on basal dendrites. 4–6 basal dendrites arise
404 from the Golgi cell soma. If we take it as 5, a population of ensemble grouped Golgi cells
405 has a combined total of 30 x 5 = 150 dendrites. The number of mossy fibres which
406 contact a Golgi cell is unreported. It has been estimated that a Golgi cell 'cannot receive
407 more than 10–12 MF [mossy fibre] inputs, given the total basolateral dendrites length
408 and the density of glomeruli' (35). Without more data it is challenging to estimate how
409 many make contact. Two per dendrite (10 between an average of 5 dendrites) is low.

410 Dendrites radiate from the cell body. A single dendrite has a range of 50–100 µm (31).
411 The longest dimension of a field – which contains ~700 mossy fibre terminals – is 200
412 µm; on the other hand, there is little dendritic branching, so a dendrite occupies a path
413 not a volume. We take the number as 20, 4 per dendrite.

414 Golgi cells inhibit granule cells in a structure termed a glomerulus (31, 36, 37). Mossy
415 fibre terminals are ensheathed in a semi-permeable membrane which restricts
416 neurotransmitter diffusion. Excitation of granule cells by mossy fibres and inhibition by
417 Golgi cells takes place here. Most of the Golgi cell axonal plexus – which is confined to
418 the granular layer – is made up of very fine beaded fibres. These enter glomeruli, where
419 they contact granule cell dendrites (37). 'In the adult rat cerebellum, each granule cell
420 dendrite receives 2.6 ± 0.55 synaptic contacts from Golgi axon terminals' (38) citing (27).[1]
421 However, the large majority of inhibition (98%) (39) is by spillover, where
422 neurotransmitter released into the synaptic cleft spills out into the glomerular space.
423 Golgi cells release GABA, an inhibitory neurotransmitter. This is detected by high-affinity
424 $GABA_A$ receptors located perisynaptically and extrasynaptically on granule cells (40, 41).
425 Even synaptically-received signals have most of their effect (i.e., the large majority of
426 charge transfer is mediated) via spillover.[2]

[1] Data are sparse. The original data are an estimate derived from visual inspection of slices thirty years ago so far without later corroboration, except from Rossi and Hamann 1998, who derive a lower estimate but propose the disparity is because their estimate is too low.
[2] Because of the long duration of IPSCs generated by spillover (recorded in slices), the total charge carried is three times that of IPSCs generated by directly connected terminals. 38. Rossi DJ, Hamann M. Spillover-mediated transmission at inhibitory synapses promoted by high affinity alpha6 subunit GABA(A) receptors and glomerular geometry. Neuron. 1998;20(4):783-95.

427 As far as is known, inhibition received by a glomerulus is from a random sample of Golgi
428 cells afferent to a field. The number of Golgi cells with convergent input to a single
429 glomerulus is unknown. There is high divergence. Each Golgi cell inhibits a large number
430 of granule cells (32, 42). A single Golgi cell contacts thousands of granule cells (37). Given
431 a fixed cell ratio, the convergence ratio scales proportionally with divergence. We take
432 the ratio as a random variable in the range 8–12:1 (we test the effect of other ratios in
433 Fig 3).

434 3.3 The Golgi cell ensemble computation

435 It is thought that discharge rates of neurons provide the physiological form of values that
436 are used in network computations. That is, firing rates provide the biological equivalent
437 of numbers, and transitions that change rates can be represented mathematically.

438 We argue in this section that transmission through a Golgi cell ensemble is in biophysical
439 steps that can be represented as mathematical operations. Steps are a sequence; a single
440 step is a compound of multiple operations reflecting events that occur in parallel.
441 Information is represented at ensemble-wide level. Each step takes a population of
442 values and converts it to another population of values. Ensembles exploit statistical
443 effects of random sampling. There are two stages of random sampling. Ensembles
444 perform better than single cells because there are more layers of sampling and more
445 samples per layer.

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

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

The postsynaptic result – charge influx to dendrites – is a proportional function of input rates. Otherwise, it would not convey rate information (43). As far as we know, there are no data for transformation of multiple mossy fibre input rates into Golgi cell basal dendritic depolarisation. We take it as the mean. (We discuss evidence for physiological conservation of rate information in the next section 3.5.2.) Taking the sample means creates a new set of values (Fig 2 column 2). The frequency distribution of the new values is in a more condensed range and centred on the mean of the sampled distribution. Basal dendritic charge transfer is passive and little filtered (44, 45). Integration of dendritic

466 states at the Golgi soma is represented by taking the mean of dendritic charge. This
467 generates a third population of values (Fig 2 column 3). Golgi cell firing rates have a linear
468 relationship with the amplitude of depolarising current (46). Thus, rate-proportional
469 dendritic charge influx is converted to Golgi cell firing rates.

470 The second stage of random sampling is of Golgi cell rates by glomeruli. Inhibition almost
471 exclusively by spillover (39-41) effectively increases sample size, as sample size is not the
472 number of Golgi cells that make direct contact on any particular granule cell, but the
473 number afferent to a glomerulus. Sample size is accordingly the number of Golgi cells
474 with convergent input to a glomerulus (taken as 8 to 12), and the number of samples is
475 the number of glomeruli per field (700). Rate-proportional spillover provides the
476 physiological equivalent of averaging afferent rates, represented by taking the sample
477 means. This generates a final population of values (Fig 2 columns 4 and 5), the output of
478 the ensemble conversion.

479 3.4 Results

480 3.4.1 A single ensemble

481 The results for a single ensemble are shown in Figure 2. Neural information has a shifting
482 form. The ensemble conversion has intracellular as well as between-cell steps and
483 therefore transitions through forms that are normally measured on different scales. We
484 use a common scale because our interest is confined to the computational effect of
485 architecture through the steps of the conversion, assuming a linear relationship of forms.

486 *Figure 2 about here please*

487 **3.4.1.1 An ensemble converts randomly sampled mossy fibre rates into synchronised**
488 **inhibition of granule cells**

489 Outputs of the simulation of a Golgi cell ensemble are graphed in Figure 2. Mossy fibre
490 firing rates received by a rank were generated as described in section 2 within
491 constraints on the frequency distribution (column 1). The distribution of firing rates
492 received by a row of three fields selected at random is shown (in red) in the same panel.

493 In column 2, the red data are themselves randomly sampled – 150 samples, sample size
494 4, representing 30 Golgi cells x 5 basal dendrites, and contact on each dendrite by 4
495 mossy fibres – and the sample means plotted, representing dendritic charge states. The
496 range is narrower than the sampled population (in all rows), but data are not tightly
497 grouped.

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

502 In the final step (column 4) somatic data are sampled 700 times, reflecting convergent
503 input from 3 fields of Golgi cells onto 700 glomeruli in the middle field. Sample size is 8–
504 12, randomly generated for each sample; physiological sample size is unknown. Column 5
505 shows the column 4 data as a histogram, for comparison with column 1.

506 The result is that mossy fibre signal frequencies received as input to a rank are reliably
507 converted to a narrow distribution of values representing inhibition of granule cells, in a
508 randomly selected field. The same result is obtained regardless of the shape of the
509 sampled distribution. Ensembles convert input values to synchronised inhibition of
510 granule cells in the middle field.

511 **3.4.1.2 Unknown parameters**

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

515 The results are shown in Figure 3. In 3a, the varied parameter is the number of mossy
516 fibres with convergent input to a single Golgi cell dendrite, and in 3b, it is the number of
517 Golgi cells with convergent input to a glomerulus. Both display the output of the
518 conversion, representing the mean of Golgi cell rates received by each glomerulus in the
519 field that receives ensemble output (so 700 values, the same as Figure 2, column 4). The
520 ratios are shown above each panel. In other figures, ratios are 4:1 (mossy fibres that
521 contact a Golgi cell basal dendrite) and 8–12:1 (Golgi cells that converge on a
522 glomerulus).

523 We found that increasing the convergence ratio condenses the plotted data range
524 (representing tighter alignment of inhibition within a field), but there is a diminishing
525 return. There is a stronger effect of increasing a low ratio and a progressively weakening

effect at higher ratios. The only condition that returned a large range was a very low Golgi cell to glomerulus ratio. So: varying unknown parameters did not reveal a reason to discard or revise the model, provided that the physiological Golgi cell to glomerulus convergence ratio is not unexpectedly low.

Figure 3 about here please

3.4.2 A rank

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

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

The bottom long panel shows mf field mean (red) and glomerular field mean (blue), again for 100 fields. As in Figure 2, the blue data are restricted to a focused range (the y axis range is shorter than Figure 2). In Figure 4B, for example, the SD of output is ~3 Hz,

546 derived from mossy fibre rates with a range of ~190 Hz, SD 25 Hz. Within that range,
547 however, input and output vary randomly and independently from field to field, in a
548 bandwidth centred on the mean of mossy fibre rates received by the rank as a whole.

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

551 Our findings are:

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

555 ii) A rank of Golgi cells behaves like a single unit. Inhibition of the granule cell population
556 of a whole rank is synchronised. Glomerulus field means are statistically constrained to a
557 narrow range.

558 iii) The mf field mean and glomerular field mean were only related by having a narrowly
559 constrained range with the same mean, which was also the mean of the sampled
560 distribution (i.e., of mossy fibre rates received by the rank). This was unexpected. We
561 expected noise to be compounded by the chain of random samplings, such that a shift
562 (deviation from the mean) of the mf field mean would tend to cause a shift in the same
563 direction of the glomerular field mean. Instead, they varied independently. Indeed, the
564 output range was modestly narrower than the input range.

565 iv) The glomerular field mean is precisely and linearly coupled to the mean of the
566 sampled distribution.

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

571 *Figure 4 about here please*

572 3.5 Functional anatomy

573 The ensemble computation generates central-limit-theorem-predicted output. For that,
574 physiology must create conditions that simulate sampling with replacement, and which
575 control for other variables. Control for other variables means, the output of the
576 conversion must be exclusively the effect of ensemble architecture (which includes cell
577 morphology), and not other variables, because other variables are noise. Other variables
578 come in three types: mechanics of transmission that may mean it is not linear, time, and
579 other sources of control of Golgi cells. This section considers sampling with replacement,
580 linear transmission, and time.

581 3.5.1 Mossy fibre terminal clustering simulates sampling with replacement

582 Recoding exploits statistical effects (known as the central limit theorem) that depend on
583 replacement. Samples taken with replacement are 'independent' – there is no effect of

584 sampling on other samples. The physiological form of sampling is anatomical. A sample is
585 made up of the firing rates of cells that contact a sampling unit – in this case a Golgi cell
586 dendrite (not the whole cell; for granule cells, a unit is the whole cell). For independence,
587 each sample must be equally likely to contain any rate in the sampled distribution. This is
588 not anatomically achievable if mossy fibres each ended in a single terminal since clearly a
589 single terminal cannot be in more than one place at once. By terminating anywhere, a
590 cell has removed itself from the population that might contact other cells. Mossy fibre
591 termination in terminal clusters, and random intermingling of terminals (Fig 1), means
592 that a signal can be received by more than one target. Moreover, at local scale there is a
593 near equal probability that this happens, because a sampling unit falls within or passes
594 through a large majority of local cluster fields. It is a biological make-do, but the outcome
595 approximates sampling with replacement.

596 Clustering itself creates a potential problem, however. Physiological contact must
597 simulate sampling of mossy fires afferent to a field, not terminals, i.e., the population of
598 rates, not the population of signals, which includes duplicates. To approximate the result
599 of sampling rates only, a single sampling unit must not receive contact from the same
600 mossy fibre more than once. A result close to that is obtained with small sample size. It is
601 an imperfect solution because it does not fully eliminate duplication, but duplication is
602 unlikely, so rare. It is intuitive that small samples would contain duplication less often,
603 and we can show mathematically with granule cells that sample size of 4 is small enough.
604 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 mossy fibres that innervate a whole strip. 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.

3.5.2 Linear transmission

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. However, evidence is available at cell and anatomical level.

Golgi cells fire autonomously. Under anaesthesia, firing falls into a slow oscillating pattern (47). Under excitatory synaptic input, however, this disappears (48). Anaesthesia 'has a strong influence on spontaneous activity of Golgi cells' (35). Discharge at low rates with irregular timing under anaesthesia (49) is replaced by higher rates with more regular timing without anaesthesia (50-52). (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 (31), 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 (44, 45), 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 (53) and lower-affinity AMPA receptors (54).

643 ‘Sensory-evoked Golgi-cell inhibition [of a granule cell] scales proportionally with the
644 level of mossy fiber excitatory synaptic input’ to the same cell, such that inhibition
645 reliably conveys mossy fibre rate information (43). Golgi cell ‘firing frequency increases
646 linearly with the amplitude of depolarising current’ (46 p.845). ‘Golgi cells can follow
647 peripheral signals in a continuous fashion, modulating their frequency with the intensity
648 of the stimulus [citing (50, 51)]’ (55). As a result, during movement, the ‘output of Golgi
649 cells varies smoothly as a function of input.…They are capable of modulating their output
650 to reflect their input with little noise or signal distortion’ (56).

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

654 There is a further stage in the pathway to inhibition of granule cells. Golgi cells provide
655 the only inhibitory input to granule cells. As noted, inhibition is almost exclusively by
656 spillover (39-41). Each glomerulus receives innervation by an independent random
657 sample of Golgi cells afferent to the field. Should spillover provide the physiological
658 equivalent of averaging afferent rates by diffusion in the intraglomerular space, as we
659 claim, linearity is preserved.

660 **3.5.3 Time**

661 We simulate a moment in time, a snapshot of concurrent activity. Our parameters are
662 linearly conserved rate information and anatomy. That is not a simplification – we argue

663 that the physiological computation is controlled for other variables, which are noise. It is
664 intuitive to think of signals as discrete episodes with a duration. However, that would
665 mean the combined effect of a set of signals would still include time-variable elements,
666 even in a snapshot, because integration would partly depend on their relative timing and
667 duration. We propose that the cerebellum side-steps that problem by primarily using
668 continuous variables for internal signalling, and that Golgi cells and glomeruli are in a
669 sustained state during periods of activity.

670 During behaviour in freely-moving animals, mossy fibre activity is dense: a high fraction
671 of mossy fibres are active (57). Both mossy fibres (30, 58, 59) and Golgi cells (50) have
672 been reported to fire with a sustained, time-varying signature, though it can depend on
673 modality and the manner of stimulation (60). This would suggest – and we propose – that
674 in the behaving animal Golgi cells receive time-varying but sustained input signals. The
675 large Golgi cell soma, where dendritic signals are integrated, is in this contention in a
676 state of continuously variable change. We note that the pause in Golgi cell firing
677 following a discrete stimulus under anaesthesia (43, 60, 61) disappears from Golgi cell
678 recordings during locomotion (50). Golgi cells fire spontaneously (49), at a rate that is
679 modulated by somatic charge, so that downstream, a glomerulus receives convergent
680 input at continuously variable rates. As a result, there is a sustained build-up of
681 glomerular GABA during behaviour (62) at an adjustable concentration controlled by the
682 mean of afferent Golgi cell firing rates (63).

683 On the face of it, glomeruli could cause a timing issue. Signalling by spillover is
684 sometimes assumed to be ambient and slow. However, the action of glomerular GABA
685 spillover has a fast phasic component – not as fast as synaptic transmission (~1 ms) but
686 with a rise time of only a few milliseconds (63). The response time to a changing rate of
687 multiple active inputs should be faster than the experimentally isolated response to a
688 discrete signal. Unlike the spiky appearance of synaptically-induced inhibitory
689 postsynaptic currents, spillover (43) generates a sustained outward current.

690 Our interpretation is that glomeruli do not cause a special timing issue. A sustained state
691 of biophysical variables may provide a physiological solution to problems (extra variables)
692 that would be caused by coding data in discrete signals.

693

694 # 4.GRANULE CELLS

695 ## 4.1 Contents

696 We next incorporated the Golgi cell ensemble conversion into a model that recodes
697 mossy fibre signals granule cell signals. It was first proposed over fifty years ago that
698 inhibition might be used to select (and change) which permutation of mossy fibre inputs
699 to a granule cell makes it fire (64) – so the same cell can code information combined in
700 different ways – or to adjust the number of inputs needed for a granule cell to fire (65),

701 as part of a mechanism that uses parallel fibre feedback to regulate the percentage that
702 are active.

703 We propose something functionally and mechanistically different. The function is to
704 maintain at a fixed number the minimum number of mossy fibre signals needed for a
705 single granule cell to fire, by auto-adjusting inhibition in step with the sampled
706 distribution. The sampled distribution is the population of mossy fibre rates received by a
707 whole rank. The function is to pixelate information encoded in granule cell signals. Pixels
708 are rank sized, long thin strips a hundred times longer than they are wide (or more). A
709 pixel is the smallest unit of granule cell-coded information, but not the smallest unit of
710 code. It is not necessary to receive the whole output of a rank to receive the information.
711 The same information is coded in any random sample of signals. Moreover, it can be
712 coded in a small fraction of the granule cell population, allowing the code to be 'sparse',
713 although we think not for the traditionally expected reasons. Primary auto-regulation
714 does not use feedback.

715 We describe the physiological mechanism and test it *in silico*. We take the same
716 approach as section 3, describing the physiology to explain the derivation of
717 computationally relevant variables and the values they take. We generate mossy fibre
718 rates as before and again start with a single field, and then simulate recoding by 100
719 fields to test whether they all return the same output, i.e., ranks recode as a unit.

720 Note: the idea that there may be an all-or-nothing result of the intraglomerular 'balance'
721 of excitation and inhibition (section 4.3) has appeared previously (66). The idea that
722 active granule cells may accordingly receive input rates with a narrowed bandwidth at
723 the higher end of the range also appears there.

724 **4.2 Anatomy**

725 The Golgi cell ensemble conversion forms part of the mechanism of granular layer
726 recoding. This section explains the derivation of additional parameters.

727 Granule cells receive input to each dendrite from a single mossy fibre. Granule cells have
728 3–5 dendrites, averaging 4 (the simulation assumes 4, but the result is the same with 3–
729 5). Contact is made in a glomerulus, as noted. Each dendrite of a single granule cell
730 extends into a different glomerulus. Estimates vary of the number of granule cells that
731 receive contact from a single terminal. 50 is mid-range (26, 27), although estimates vary
732 (25). 50 is consistent with our earlier estimates of the number of terminals and granule
733 cells in a field. A granule cell receives contact from a random sample of mossy fibres
734 afferent to a field, and therefore receives a random sample of mossy fibre rates. The
735 sampling unit is the whole cell since it receives only a single input rate to each dendrite.
736 Sample size is therefore 4, and the number of samples per field is the number of granule
737 cells, which we estimate above at 8,750.

738

739 ## 4.3 The granule cell computation

740 A granule cell needs a minimum number of active mossy fibre inputs to fire – thought to
741 be three (29, 67). Certainly, one is too few (39, 68, 69). We take it as three. Three signals
742 is not in itself sufficient, as each of those signals must also prevail over spillover-
743 mediated Golgi cell inhibition. Inhibition is provided by the Golgi cell ensemble
744 conversion, so it is synchronised in a rank. It is not exactly synchronous. A single granule
745 cell receives inhibition of each dendrite from an independent random sample of Golgi
746 cells afferent to the field. Unitary recordings confirm that a granule cell receives
747 inhibition from multiple Golgi cells (43). Nonetheless, it is in a focussed range so that all
748 mossy fibre signals compete against inhibition in a narrow range.

749 In each glomerulus, excitation and inhibition compete for influence. Activation of
750 presynaptic $GABA_B$ receptors on mossy fibres inhibits glutamate release (70), while
751 activation of presynaptic mGluR2 receptors on Golgi cells inhibits GABA release (71). The
752 outcome of the competition is independent of other mossy fibre terminals, even other
753 terminals that provide input to the same granule cell. It is unclear if there is always an
754 emphatic winner, but it is possible, because a swing in the dominant direction may be
755 amplified by positive feedback. So even if all granule cells receive four active mossy fibre
756 inputs, which we assume they do, only a subset fire, those that receive at least three
757 mossy fibre winners.

758 There is a single round of sampling, by granule cells. We term the mean of Golgi cell rates
759 received by a single glomerulus the 'inhibitory threshold' and take it that a mossy fibre
760 signal must exceed the inhibitory threshold to count towards the minimum number
761 needed for a granule cell to fire. The simulation compares mossy fibre signal frequency
762 with the inhibitory threshold at each of 4 dendrites of each of 8,750 granule cells in each
763 of 100 fields. As discussed, there are 700 outputs of the ensemble conversion per field,
764 one for each glomerulus. Excitation and inhibition are paired at random. The comparison
765 is used to identify the subset of granule cells that fire (cells that receive at least 3 mossy
766 fibre signals which exceed the inhibitory threshold). The simulation takes the mean of
767 input rates received by each of those cells. Output of the model is accordingly the
768 number of granule cells that fire in each field, the mean of input rates they each receive,
769 and the average for each field ('GC field mean'). The results are shown in Figures 5 and 6.

770 **4.4 Results**

771 We find:

772 i) The frequency distribution of the means of mossy fibre rates received by each granule
773 cell is narrower than the sampled distribution, normal, and has the same mean. The
774 distribution of the means of active granule cells is still narrower because they receive
775 mossy fibre rates at the higher end of the range (Fig 5 column 2). This typically holds

776 even in a single field when only a very small percentage of granule cells fire (~1% active)
777 (Fig 5 column 2 row 6, Fig 6 column 2).[3]

778 ii) A sagittal row of fields which all sample the same population of mossy fibre rates, all
779 return the same output (Fig 5 column 3, Fig 6 column 3). The same output means, the
780 granule cell frequency distribution has the same mean, spread and shape. Accordingly,
781 any random sample of the entire output of a rank has the same statistics.

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

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

788 Our interpretation is that recoding pixelates multidimensional mossy fibre input,
789 converting it into a low-resolution internal code. Pixels are rank sized – there is no finer
790 grade of resolution. Any random sample of signals inside pixel boundaries (over a modest
791 minimum) codes the same information. There is a statistically conserved linear and
792 precise relationship between the mean of the sampled distribution and the mean of
793 firing rates of randomly sampled granule cells. Auto-modulation of Golgi cell inhibition of

[3] The focus of the granule cell range depends on the range of the sampled distribution (Fig 6 column 2 rows 4–6).

794 granule cells means the number of granule cells that fire is independent of mossy fibre
795 rates. It also regulates the number of co-active granule cells without feedback.

796 *Figures 5 and 6 about here*

797 4.5 Functional anatomy

798 The same comments apply here as the commentary on Golgi cells in section 3.5. The
799 treatment here is briefer because some of it is already covered.

800 If granule cells collectively code mossy fibre rate information in the form we propose,
801 weighted mossy fibre to granule cell transmission would interfere, so we would expect
802 transmission to be adapted to conserve rate information. The short and equal length of
803 granule cell dendrites would suggest light and equal filtering. The mossy fibre-granule
804 cell connection has multiple adaptations to transmit high frequency signals (72) across a
805 wide bandwidth. Faithful transmission of rate information is reported (25, 73). Vesicle
806 release and replenishment are fast (74) and postsynaptic AMPA receptors operate in
807 their linear range (75), where they are resistant to desensitisation (54). Conversion of
808 depolarising somatic charge into granule cell firing is a simple function such that granule
809 cells 'have a relatively linear and uncomplicated conversion of depolarisation level to
810 spike rate' (69 p.2393, citing Jorntell and Ekerot 2006 and D'Angelo et al 1998).

811 However, linear transmission is contentious because it has been reported that
812 transmission strength may be variable. There is variation between granule cells (and

813 trials) of EPSCs following mossy fibre stimulation *in vitro* (75), leading to the suggestion
814 that fast vesicular release may allow precise signalling with heterogeneous synaptic
815 weights. However, in this study stimulation was typically not above 2 Hz to avoid short-
816 term depression observed at higher frequencies, and GABA was blocked.

817 An alternative interpretation is that during behaviour, at behavioural mossy fibre rates,
818 there is glomerular build-up of intraglomerular glutamate, and signalling by spillover
819 becomes important, as it is for inhibitory signals (39, 63). Spillover from multiple release
820 sites (200-400 at a single mossy fibre terminal) is plausibly sufficient to mitigate
821 variability of vesicular size, release probability, and the number of release points at
822 synapses made on any individual cell. Indeed, short-term synaptic plasticity may have the
823 function of increasing the relative influence of spillover.

824 Still, there *are* weights. Different strength and short-term dynamics of the mossy fibre to
825 granule cell connection have been reported in vestibular slices, with $GABA_A$ receptors
826 blocked. The amplitude of the response, measured at the soma, is reported to be
827 statistically correlated to the source of the mossy fibre that is stimulated (76), suggesting
828 that the strength of the connection depends on the source.

829 However, the presence of weights does not necessarily mean that the function is to
830 modulate firing rates. Indeed, the authors of the study don't suggest that, but propose
831 that it may temporally code the source of inputs because different combinations have a
832 different response onset time. An alternative, which they mention, is that some inputs

833 may be in a supporting role, reminiscent of ‘driver’ and ‘modulatory’ signals in the
834 thalamus and cortex (77). Feasibly, signals in a supporting role must be present for the
835 postsynaptic granule cell to fire, but do not affect the granule cell firing rate.

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

837 4.6 Physiology implements a binomial function

838 While mossy fibres are densely active, as noted, they may not all be active, and the
839 permutation of active and inactive cells may change with time. Indeed, doubtless during
840 behaviour it is in a state of flux. This does not impact function because the effect of silent
841 cells on the outcome of recoding is simply the same as the effect of weak signals.

842 Regulation by inhibition of the number of granule cells that fire has the same
843 computational result as silencing most mossy fibres, leaving a minority active. The ‘silent’
844 cells are the ones that fire at rates below the inhibitory threshold – the losers in the
845 glomerular competition. The probability that a given cell fires depends on the percentage
846 of winners. Inhibition controls the probability by controlling the percentage. Because the
847 number of granule cells in a field is large, the number of active and inactive granule cells
848 is very near the expected proportions.

849 The proportions are given by a binomial function. In that sense, physiology implements a
850 binomial distribution. The expected number of granule cells that fire in a field is given by

851

852 $$E_{fld} = T \sum_{x=j}^{n} \left(\binom{n}{x} [P(e > i)]^{x} [1 - P(e > i)]^{n-x} \right)$$

853

854 where T is the total number of granule cells per field, n is the number of dendrites per
855 granule cell, and j is the minimum number that must receive excitation that prevails over
856 inhibition, given by $e > i$.

857 The simulation does not include a binomial function, it derives output from input by
858 comparing excitatory to inhibitory input to each glomerulus, glomerulus by glomerulus,
859 the way it is derived physiologically. The reason we mention the binomial formula is that
860 it is part of the functional context of synchronised inhibition that tracks mossy fibre rates.
861 As noted, there is a long-standing proposal that Golgi cell control of the number of
862 effective inputs needed to make a granule cell fire might be used to regulate the density
863 of parallel fibre signals activity (65). We counter propose that the opposite is necessary –
864 the minimum number of effective inputs must be unvarying for control of the number
865 that fire by probability. The primary controller is feedforward inhibition of granule cells
866 (and not parallel fibre feedback) for that reason: so that inhibition tracks mossy fibre
867 rates.

868 It is worth stating that a minimum number of effective inputs is also minimum sample
869 size, which is low anyway. For that reason, it is also safeguards the statistically reliable
870 encoding of mossy fibre rate information in granule cell firing.

5. DISCUSSION

5.1 Reprise

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 functional units – ensembles. The function of ensembles is to convert mossy fibre input to a 3-field strip into synchronous inhibition of granule cells in the middle field. The function of inhibition is to participate in recoding which converts mossy fibre signals into granule cell signals. 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 is a linear function of the mean of the sampled population (i.e., of mossy fibre input rates to the whole rank).

The ensemble conversion forms part of a model of recoding that converts mossy fibre rate information into a well-predicted granule cell code. Since at any moment all fields randomly sample the same population of mossy fibre rates, and receive synchronised inhibition, they all return the same output code. That is, recoding is unitised. The output of a rank is the smallest unit of granule cell encoded information. Resolution is low: a rank is a long thin strip of the granular layer. The mean firing rate of granule cells in each field is a linear function of the mean of the sampled distribution. As all fields return the same code, and information is collective coded in statistics, the same information is coded in firing of *any* random sample of granule cells in a rank.

5.2 Pixels offer a substrate of control of Purkinje cells

We model a moment in time, a snapshot of a state of flux, in a single rank. The cerebellum receives a torrent of data during movement. The region that provides parallel fibre input to a microzone contains around 40 ranks, 20 each side, giving rise to parallel fibres that intermingle in the molecular layer. We do not include other ranks.

With that caveat, 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 (14, 15). 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, even if only a small fraction of parallel fibres are active, many make contact. If 0.5% are active, a single Purkinje cell receives ~875 signals, enough to code information as statistics. Since the output of recoding is pixelated, and the same information is contained in any random sample of granule cell signals inside (nominal) pixel boundaries, and Purkinje cells randomly sample granule cell signals, all Purkinje cells receive the same code, with the same timing.

5.3 A sparse granule cell code is economic and quiet

The recoding mechanism also provides regulation of the fraction of active granule cells. Auto-modulation of inhibition decouples the fraction from mossy fibre rates. Without adjustable inhibition of granule cells by Golgi cells, stronger mossy fibre rates would mean that fewer mossy fibre signals would be needed to make a granule cell fire. So, if rates went up, more would fire. 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 we have no use for the traditional reason. There are other reasons for a low and stable number, however.

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

929 energy saving. Granule cells fire in short bursts. A single burst is at a single firing rate.
930 Modulation of firing of parallel fibre targets is achieved by the turnover of parallel fibre
931 rates. A low fraction of granule cells firing in short bursts is a large saving.

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

937 Stability may be enhanced by granule cell ascending axon contact on Golgi cell basal
938 dendrites (45). A Golgi cell receives ~400 functional synaptic contacts from this source,
939 around 25% of the granule cells that a Golgi cell receives contact from in total. 'During
940 sustained mf [mossy fibre] discharge, synapses from local GrCs [granule cells] contribute
941 equivalent charge to GoCs [Golgi cells] as mf synapses' and 'display similarly fast kinetics'
942 (78), providing the substrate for potent local feedback that should further narrow the
943 range of the glomerular means both within and between fields.

944 5.4 Do Golgi cells receive a significant influence of excitatory or inhibitory 945 input from other sources?

946 Further regulation may be provided by feedback via parallel fibres received by dendrites
947 which some Golgi cells extend into the molecular layer. It is a long-standing idea that
948 feedback may contribute to self-regulation of the fraction of co-active parallel fibres (64,

949 65). Our expectation is that the strength of an effect by this pathway is modest (because
950 fine adjustments are sufficient for its function), consistent with a leading role in
951 regulation of feed-forward inhibition. Modulation of Golgi cell rates by single parallel
952 fibre signals is weak (46), although sub-linear integration is mitigated by gap junctions
953 (79). Also, only a modest number of parallel fibre signals are received at any time. A small
954 fraction of parallel fibres which pass through a Golgi cell territory make contact (~1,200
955 out of a number in the hundreds of thousands (78)). Of those, only some are active at
956 any time, and if the parallel fibre code is sparse, it is a small fraction. For example, if 1%
957 of parallel fibres are active, it is 12. This would suggest the effect of parallel fibre inputs
958 on Golgi cell firing is relatively weak, and control by direct input from mossy fibres is the
959 main form of control. Possibly, small modulations of Golgi cell somatic charge are
960 sufficient to make small adjustments that regulate a sparse code – that is, the number of
961 granule cells that need to be affected is relatively modest, so fine margins are sufficient.
962 As noted, not all Golgi cells have apical dendrites, so the collective effect of parallel fibre
963 feedback is presumably diluted because a significant population of Golgi cells don't
964 receive it.

965 Do Golgi cells receive a significant influence of input from other sources? The evidence is
966 incomplete but currently suggests there may not be any, or any that are significant.
967 Contrary to early reports, neither Purkinje cells nor climbing fibres contact Golgi cells (35,
968 who give references), and only a modest minority of inhibitory inputs to Golgi cells (if
969 any) are from molecular layer interneurons (80), which generate weak synaptic currents

970 (81), consistent with either extremely weak or wholly absent innervation (82). There is
971 conflicting evidence whether Golgi cells make inhibitory synaptic contact on each other
972 (48, 82). The simulation does not include an effect on Golgi cell firing by other sources of
973 input.

974 5.5 High-level problem solving: a comparison of the present proposals and 975 the supervised learning model

976 Physiological neural networks face some high-level problems which therefore neural
977 network models share. The concluding part of the Discussion compares our proposed
978 solution of those problems with the supervised learning model solution, the name of the
979 traditionally dominant class of cerebellar models.

980 1) Networks must distil output from input values. The traditional cerebellar model, and
981 many generic artificial neural networks, hold that computational derivation of output
982 from input is by modification of synaptic weights under climbing fibre tuition. However,
983 in the cerebellum, it has ultimately emerged that the evidence is unconvincing for that
984 and good for linear transmission. We propose, in conflict with the traditional model, that
985 the computation is mediated by cell type morphologies and network architecture, a
986 dynamic result of passive anatomy. As anatomy is known in detail, there is rich evidence
987 to provide parameters for a simulation, down to the level of biological messiness at a
988 local scale. Anatomy is underrated in theory. It tends to be thought of as merely a vehicle

989 to move data around, and therefore not evidence of anything. We propose that anatomy
990 has a computational function and therefore that evidence of anatomy is evidence of the
991 mechanism.

992 2) Signals must be correctly routed – not straightforward because in each layer of cells
993 they pass through, each signal is received by multiple targets and each of those targets
994 receives multiple signals, without an obvious key to the organisation of individual
995 connections, if any. The traditional theoretical solution is external supervision (as in
996 'supervised learning'). This is provided by climbing fibres which signal errors and teach
997 change at synapses which are active in a learning window. The model assumes that the
998 result of training is corrected motor output. It therefore does not matter which particular
999 pattern of parallel fibres is active. It also assumes that Purkinje cells individually receive
1000 correctly routed signals, without explaining why they are organised into microzones.

1001 We propose instead that internally encoded information has large spatial dimensions.
1002 Unitised coding in topographically spatially-defined blocks, and in statistics, means there
1003 is no need to route signals at a smaller scale. This is consistent with the organisation
1004 reported in mapping studies and is a plausible substrate to drive coherent firing of
1005 functionally-grouped Purkinje cells. It does not need supervision. In fact, we believe that
1006 error signals are a good candidate to explain the vestibulo-ocular reflex but probably not
1007 function of non-floccular circuits.

1008 3) Insentient physiological networks must compute input firing rates without needing to
1009 discriminate between signals on the basis of information they don't have – this can be
1010 likened to receiving input values measured on different scales without knowing what the
1011 scales are. Input to the cerebellum is eclectic and in multiple modalities. In some cases, a
1012 single granule cell can receive multi-modal input. In the traditional model, supervision
1013 solves this problem. It is unnecessary for the cerebellum itself to discriminate because
1014 input patterns are selected for training by the timing of instruction signals.

1015 Our alternative solution is that resolution of high-dimensional input data is sacrificed at a
1016 surprisingly profligate level. The cerebellum disregards source and modality data below a
1017 physical scale which mirrors topography. It is therefore unnecessary to discriminate
1018 below that scale. Mapping studies show that mossy fibre signals (evoked by cutaneous
1019 stimulation) share the topographical organisation of climbing fibres (22-24). Input data
1020 are not normalised, as that would mean by definition that it was meaningfully converted
1021 to a common internal scale. Rather, at sub-topographical dimensions, signal source and
1022 type are simply discounted – they have no effect on the cerebellar computation or
1023 therefore on output. This has the flexible advantage that networks do not need to be
1024 adapted to the source or type of input they receive – anatomically, networks are
1025 interchangeable, so network design can be modular.

1026

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.

REFERENCES

1. Raymond JL, Medina JF. Computational Principles of Supervised Learning in the Cerebellum. Annual review of neuroscience. 2018;41:233-53.
2. Albus JS, editor The Marr and Albus theories of the cerebellum-two early models of associative memory. Digest of Papers COMPCON Spring 89 Thirty-Fourth IEEE Computer Society International Conference: Intellectual Leverage; 1989 Feb. 27 1989-March 3 1989.
3. Brunel N, Hakim V, Isope P, Nadal JP, Barbour B. Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell. Neuron. 2004;43(5):745-57.
4. Fujita M. Adaptive filter model of the cerebellum. Biological cybernetics. 1982;45(3):195-206.
5. Dean P, Porrill J, Ekerot CF, Jörntell H. The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. Nature reviews Neuroscience. 2010;11(1):30-43.
6. Welsh JP, Yamaguchi H, Zeng XH, Kojo M, Nakada Y, Takagi A, et al. Normal motor learning during pharmacological prevention of Purkinje cell long-term depression. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(47):17166-71.
7. Schonewille M, Gao Z, Boele HJ, Veloz MF, Amerika WE, Simek AA, et al. Reevaluating the role of LTD in cerebellar motor learning. Neuron. 2011;70(1):43-50.
8. ten Brinke MM, Boele HJ, Spanke JK, Potters JW, Kornysheva K, Wulff P, et al. Evolving Models of Pavlovian Conditioning: Cerebellar Cortical Dynamics in Awake Behaving Mice. Cell reports. 2015;13(9):1977-88.
9. Isope P, Barbour B. Properties of unitary granule cell-->Purkinje cell synapses in adult rat cerebellar slices. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2002;22(22):9668-78.
10. Ekerot CF, Jörntell H. Parallel fibre receptive fields of Purkinje cells and interneurons are climbing fibre-specific. The European journal of neuroscience. 2001;13(7):1303-10.
11. Oscarsson O. Functional units of the cerebellum - sagittal zones and microzones. Trends in Neurosciences. 1979;2:143-5.
12. Garwicz M, Ekerot CF, Jorntell H. Organizational principles of cerebellar neuronal circuitry. News in Physiological Sciences. 1998;13:28-34.
13. Ozden I, Sullivan MR, Lee HM, Wang SS. Reliable coding emerges from coactivation of climbing fibers in microbands of cerebellar Purkinje neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2009;29(34):10463-73.
14. Harvey RJ, Napper RM. Quantitative study of granule and Purkinje cells in the cerebellar cortex of the rat. The Journal of comparative neurology. 1988;274(2):151-7.
15. Harvey RJ, Napper RM. Quantitative studies on the mammalian cerebellum. Progress in neurobiology. 1991;36(6):437-63.
16. Brand S, Dahl AL, Mugnaini E. The length of parallel fibers in the cat cerebellar cortex. An experimental light and electron microscopic study. Experimental brain research. 1976;26(1):39-58.
17. Mugnaini E. The length of cerebellar parallel fibers in chicken and rhesus monkey. The Journal of comparative neurology. 1983;220(1):7-15.

18. Wu HS, Sugihara I, Shinoda Y. Projection patterns of single mossy fibers originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei. The Journal of comparative neurology. 1999;411(1):97-118.
19. Sultan F, Heck D. Detection of sequences in the cerebellar cortex: numerical estimate of the possible number of tidal-wave inducing sequences represented. Journal of physiology, Paris. 2003;97(4-6):591-600.
20. Shinoda Y, Sugihara I. Axonal Trajectories of Single Climbing and Mossy Fiber Neurons in the Cerebellar Cortex and Nucleus. In: Manto M. SJD, Rossi F., Gruol D.L., Koibuchi N., editor. Handbook of the Cerebellum and Cerebellar Disorders: Springer, Dordrecht; 2013.
21. Shinoda Y, Sugihara I, Wu HS, Sugiuchi Y. The entire trajectory of single climbing and mossy fibers in the cerebellar nuclei and cortex. Progress in brain research. 2000;124:173-86.
22. Apps R, Garwicz M. Anatomical and physiological foundations of cerebellar information processing. Nature reviews Neuroscience. 2005;6(4):297-311.
23. Apps R, Hawkes R. Cerebellar cortical organization: a one-map hypothesis. Nature reviews Neuroscience. 2009;10(9):670-81.
24. Ji Z, Hawkes R. Topography of purkinje cell compartments and mossy fiber terminal fields in lobules ii and iii of the rat cerebellar cortex: Spinocerebellar and cuneocerebellar projections. Neuroscience. 1994;61(4):935-54.
25. Ritzau-Jost A, Delvendahl I, Rings A, Byczkowicz N, Harada H, Shigemoto R, et al. Ultrafast action potentials mediate kilohertz signaling at a central synapse. Neuron. 2014;84(1):152-63.
26. Gao Z, Proietti-Onori M, Lin Z, Ten Brinke MM, Boele HJ, Potters JW, et al. Excitatory Cerebellar Nucleocortical Circuit Provides Internal Amplification during Associative Conditioning. Neuron. 2016;89(3):645-57.
27. Jakab RL, Hámori J. Quantitative morphology and synaptology of cerebellar glomeruli in the rat. Anatomy and embryology. 1988;179(1):81-8.
28. Eccles JC, Ito M, Szentágothai J. The cerebellum as a neuronal machine. Berlin, New York etc.: Springer-Verlag; 1967. 335 p. p.
29. Jörntell H, Ekerot CF. Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2006;26(45):11786-97.
30. van Kan PL, Gibson AR, Houk JC. Movement-related inputs to intermediate cerebellum of the monkey. Journal of neurophysiology. 1993;69(1):74-94.
31. Palay SL, Chan-Palay V. Cerebellar cortex : cytology and organization. Berlin: Springer; 1974.
32. Barmack NH, Yakhnitsa V. Functions of interneurons in mouse cerebellum. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2008;28(5):1140-52.
33. Llinás RR, Negrello MN. Cerebellum 2015 [Available from: http://www.scholarpedia.org/article/Cerebellum.
34. Lange W. Regional differences in the distribution of golgi cells in the cerebellar cortex of man and some other mammals. Cell and tissue research. 1974;153(2):219-26.
35. Pietrajtis K, Dieudonné S. Golgi Neurons. In: Manto M, Schmahmann JD, Rossi F, Gruol DL, Koibuchi N, editors. Handbook of the Cerebellum and Cerebellar Disorders. Dordrecht: Springer Netherlands; 2013. p. 829-52.
36. Eccles JC, Llinás R, Sasaki K. Golgi cell inhibition in the cerebellar cortex. Nature. 1964;204:1265-6.

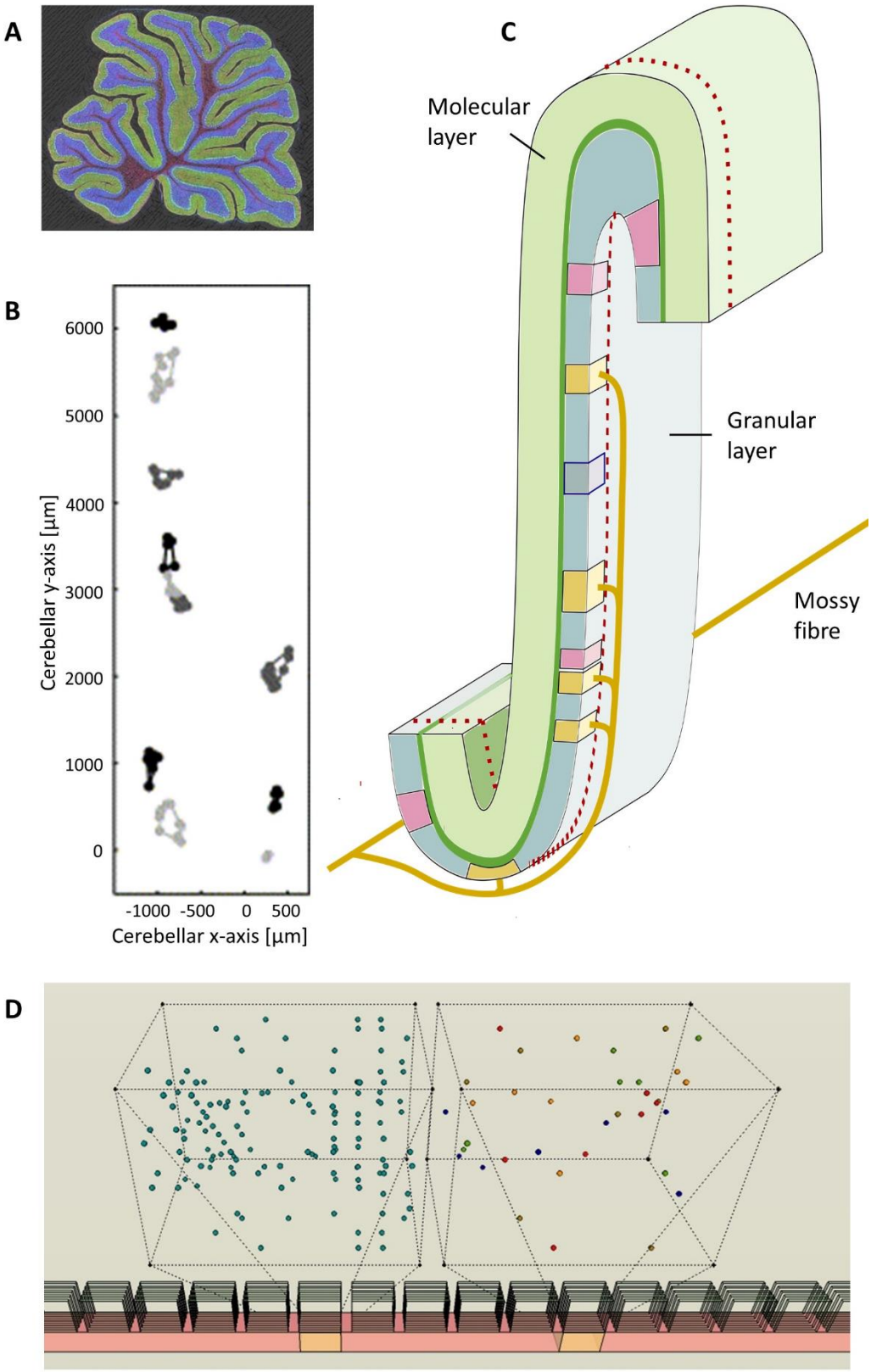
1137 37. Hámori J, Szentágothai J. Participation of Golgi neuron processes in the cerebellar
1138 glomeruli: an electron microscope study. Experimental brain research. 1966;2(1):35-48.
1139 38. Rossi DJ, Hamann M. Spillover-mediated transmission at inhibitory synapses promoted by
1140 high affinity alpha6 subunit GABA(A) receptors and glomerular geometry. Neuron.
1141 1998;20(4):783-95.
1142 39. Duguid I, Branco T, London M, Chadderton P, Hausser M. Tonic inhibition enhances
1143 fidelity of sensory information transmission in the cerebellar cortex. The Journal of neuroscience
1144 : the official journal of the Society for Neuroscience. 2012;32(32):11132-43.
1145 40. Nusser Z, Sieghart W, Somogyi P. Segregation of different GABAA receptors to synaptic
1146 and extrasynaptic membranes of cerebellar granule cells. The Journal of neuroscience : the
1147 official journal of the Society for Neuroscience. 1998;18(5):1693-703.
1148 41. Brickley SG, Cull-Candy SG, Farrant M. Development of a tonic form of synaptic inhibition
1149 in rat cerebellar granule cells resulting from persistent activation of GABAA receptors. The
1150 Journal of physiology. 1996;497 (Pt 3):753-9.
1151 42. D'Angelo E, Solinas S, Mapelli J, Gandolfi D, Mapelli L, Prestori F. The cerebellar Golgi cell
1152 and spatiotemporal organization of granular layer activity. Frontiers in neural circuits. 2013;7:93.
1153 43. Duguid I, Branco T, Chadderton P, Arlt C, Powell K, Hausser M. Control of cerebellar
1154 granule cell output by sensory-evoked Golgi cell inhibition. Proceedings of the National Academy
1155 of Sciences of the United States of America. 2015;112(42):13099-104.
1156 44. Kanichay RT, Silver RA. Synaptic and cellular properties of the feedforward inhibitory
1157 circuit within the input layer of the cerebellar cortex. The Journal of neuroscience : the official
1158 journal of the Society for Neuroscience. 2008;28(36):8955-67.
1159 45. Cesana E, Dieudonné S, Isope P, Bidoret C, D'Angelo E, Forti L, editors. Novel granule cell-
1160 Golgi cell excitatory input in the cerebellar granular layer. FENS Abstr; 2010.
1161 46. Dieudonné S. Submillisecond kinetics and low efficacy of parallel fibre-Golgi cell synaptic
1162 currents in the rat cerebellum. The Journal of physiology. 1998;510 (Pt 3):845-66.
1163 47. Ros H, Sachdev RN, Yu Y, Sestan N, McCormick DA. Neocortical networks entrain
1164 neuronal circuits in cerebellar cortex. The Journal of neuroscience : the official journal of the
1165 Society for Neuroscience. 2009;29(33):10309-20.
1166 48. Vervaeke K, Lorincz A, Gleeson P, Farinella M, Nusser Z, Silver RA. Rapid
1167 desynchronization of an electrically coupled interneuron network with sparse excitatory synaptic
1168 input. Neuron. 2010;67(3):435-51.
1169 49. Ruigrok TJ, Hensbroek RA, Simpson JI. Spontaneous activity signatures of morphologically
1170 identified interneurons in the vestibulocerebellum. The Journal of neuroscience : the official
1171 journal of the Society for Neuroscience. 2011;31(2):712-24.
1172 50. Edgley SA, Lidierth M. The discharges of cerebellar Golgi cells during locomotion in the
1173 cat. The Journal of physiology. 1987;392:315-32.
1174 51. Miles FA, Fuller JH, Braitman DJ, Dow BM. Long-term adaptive changes in primate
1175 vestibuloocular reflex. III. Electrophysiological observations in flocculus of normal monkeys.
1176 Journal of neurophysiology. 1980;43(5):1437-76.
1177 52. Rasmussen A, Zucca R, Jirenhed DA, Johansson F, Ortenblad C, Svensson P, et al. Golgi
1178 cell activity during eyeblink conditioning in decerebrate ferrets. Cerebellum (London, England).
1179 2014;13(1):42-5.
1180 53. DiGregorio DA, Nusser Z, Silver RA. Spillover of glutamate onto synaptic AMPA receptors
1181 enhances fast transmission at a cerebellar synapse. Neuron. 2002;35(3):521-33.

54. DiGregorio DA, Rothman JS, Nielsen TA, Silver RA. Desensitization properties of AMPA receptors at the cerebellar mossy fiber granule cell synapse. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2007;27(31):8344-57.
55. Galliano E, Mazzarello P, D'Angelo E. Discovery and rediscoveries of Golgi cells. The Journal of physiology. 2010;588(Pt 19):3639-55.
56. Moore JW, Blazis DEJ. 11 - Simulation of a Classically Conditioned Response: A Cerebellar Neural Network Implementation of the Sutton–Barto–Desmond Model. In: Byrne JH, Berry WO, editors. Neural Models of Plasticity: Academic Press; 1989. p. 187-207.
57. Hartmann MJ, Bower JM. Tactile responses in the granule cell layer of cerebellar folium crus IIa of freely behaving rats. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2001;21(10):3549-63.
58. Arenz A, Silver RA, Schaefer AT, Margrie TW. The contribution of single synapses to sensory representation in vivo. Science (New York, NY). 2008;321(5891):977-80.
59. Giovannucci A, Badura A, Deverett B, Najafi F, Pereira TD, Gao Z, et al. Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. Nature neuroscience. 2017.
60. Holtzman T, Rajapaksa T, Mostofi A, Edgley SA. Different responses of rat cerebellar Purkinje cells and Golgi cells evoked by widespread convergent sensory inputs. The Journal of physiology. 2006;574(Pt 2):491-507.
61. Vos BP, Volny-Luraghi A, De Schutter E. Cerebellar Golgi cells in the rat: receptive fields and timing of responses to facial stimulation. The European journal of neuroscience. 1999;11(8):2621-34.
62. Rossi DJ, Hamann M, Attwell D. Multiple modes of GABAergic inhibition of rat cerebellar granule cells. The Journal of physiology. 2003;548(Pt 1):97-110.
63. Mapelli L, Solinas S, D'Angelo E. Integration and regulation of glomerular inhibition in the cerebellar granular layer circuit. Frontiers in cellular neuroscience. 2014;8:55.
64. Marr D. A theory of cerebellar cortex. The Journal of physiology. 1969;202(2):437-70.
65. Albus JS. A Theory of Cerebellar Function. Mathematical Biosciences. 1971;10(1-2):25-61.
66. Gilbert M, Miall RC. How and Why the Cerebellum Recodes Input Signals: An Alternative to Machine Learning. Neuroscientist. 2021a;DOI: 10.1177/1073858420986795.
67. Billings G, Piasini E, Lorincz A, Nusser Z, Silver RA. Network structure within the cerebellar input layer enables lossless sparse encoding. Neuron. 2014;83(4):960-74.
68. Chadderton P, Margrie TW, Hausser M. Integration of quanta in cerebellar granule cells during sensory processing. Nature. 2004;428(6985):856-60.
69. Bengtsson F, Jörntell H. Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(7):2389-94.
70. Mitchell SJ, Silver RA. GABA spillover from single inhibitory axons suppresses low-frequency excitatory transmission at the cerebellar glomerulus. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2000a;20(23):8651-8.
71. Mitchell SJ, Silver RA. Glutamate spillover suppresses inhibition by activating presynaptic mGluRs. Nature. 2000b;404(6777):498-502.
72. Delvendahl I, Hallermann S. The Cerebellar Mossy Fiber Synapse as a Model for High-Frequency Transmission in the Mammalian CNS. Trends Neurosci. 2016;39(11):722-37.

1226 73. Rancz EA, Ishikawa T, Duguid I, Chadderton P, Mahon S, Hausser M. High-fidelity
1227 transmission of sensory information by single cerebellar mossy fibre boutons. Nature.
1228 2007;450(7173):1245-8.
1229 74. Saviane C, Silver RA. Fast vesicle reloading and a large pool sustain high bandwidth
1230 transmission at a central synapse. Nature. 2006;439(7079):983-7.
1231 75. Sargent PB, Saviane C, Nielsen TA, DiGregorio DA, Silver RA. Rapid vesicular release,
1232 quantal variability, and spillover contribute to the precision and reliability of transmission at a
1233 glomerular synapse. The Journal of neuroscience : the official journal of the Society for
1234 Neuroscience. 2005;25(36):8173-87.
1235 76. Chabrol FP, Arenz A, Wiechert MT, Margrie TW, DiGregorio DA. Synaptic diversity enables
1236 temporal coding of coincident multisensory inputs in single neurons. Nature neuroscience.
1237 2015;18(5):718-27.
1238 77. Sherman SM, Guillery RW. Distinct functions for direct and transthalamic corticocortical
1239 connections. Journal of neurophysiology. 2011;106(3):1068-77.
1240 78. Cesana E, Pietrajtis K, Bidoret C, Isope P, D'Angelo E, Dieudonne S, et al. Granule cell
1241 ascending axon excitatory synapses onto Golgi cells implement a potent feedback circuit in the
1242 cerebellar granular layer. The Journal of neuroscience : the official journal of the Society for
1243 Neuroscience. 2013;33(30):12430-46.
1244 79. Vervaeke K, Lorincz A, Nusser Z, Silver RA. Gap junctions compensate for sublinear
1245 dendritic integration in an inhibitory network. Science (New York, NY). 2012;335(6076):1624-8.
1246 80. Eyre MD, Nusser Z. Only a Minority of the Inhibitory Inputs to Cerebellar Golgi Cells
1247 Originates from Local GABAergic Cells. eNeuro. 2016;3(2).
1248 81. Dumoulin A, Triller A, Dieudonné S. IPSC kinetics at identified GABAergic and mixed
1249 GABAergic and glycinergic synapses onto cerebellar Golgi cells. The Journal of neuroscience : the
1250 official journal of the Society for Neuroscience. 2001;21(16):6045-57.
1251 82. Hull C, Regehr WG. Identification of an inhibitory circuit that regulates cerebellar Golgi
1252 cell activity. Neuron. 2012;73(1):149-58.

1253

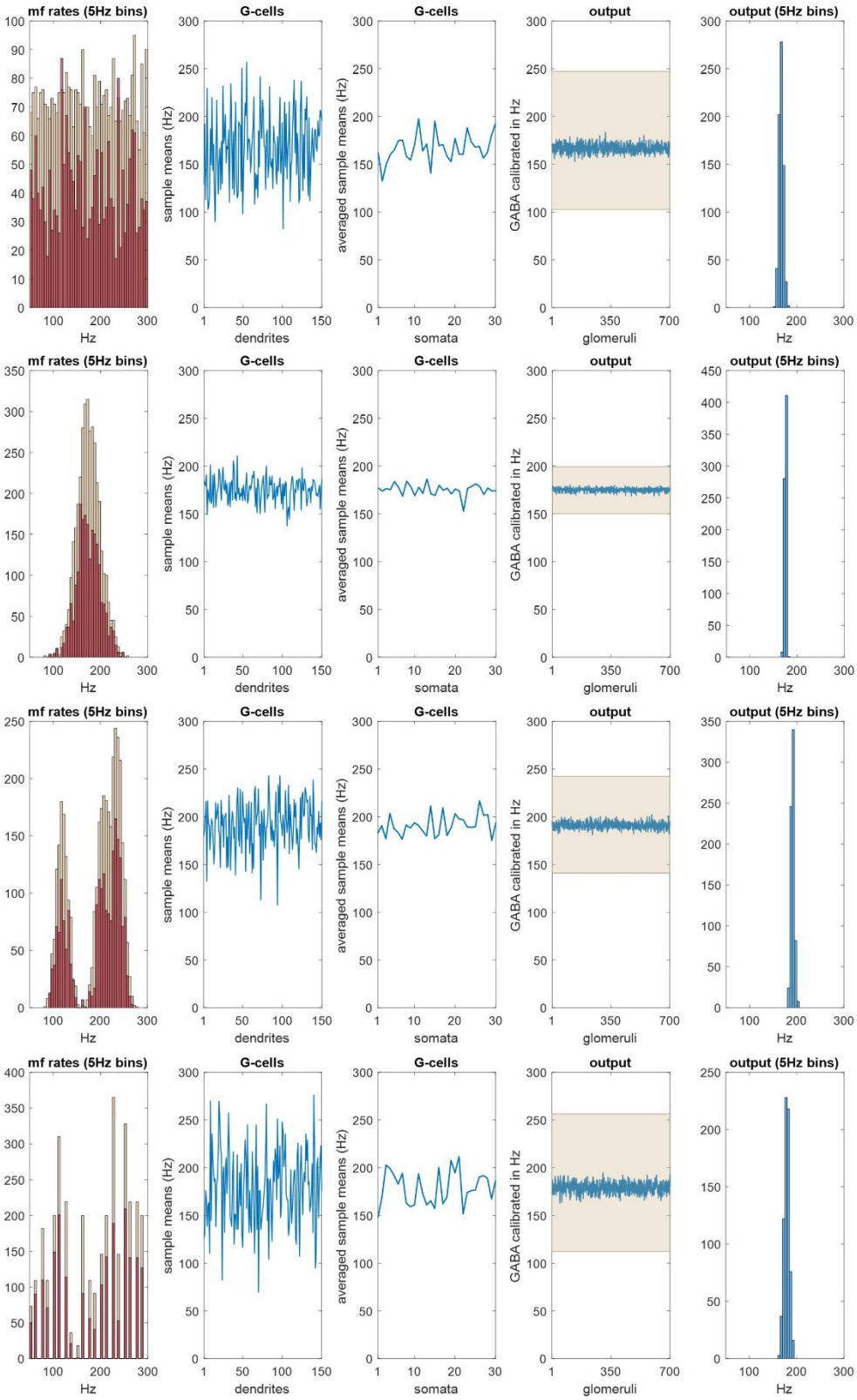
1254


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

1255

1256 **Fig 1. Mossy fibre termination in a rank.** There is random interdigitation of branch
1257 endings of mossy fibres co-afferent to a rank, and at local level random intermingling of
1258 terminals of mossy fibres co-afferent to a field. **(A)** Sagittal section of the rat cerebellum
1259 (obtained by multi-photon microscopy, courtesy Tom Deerinck). The outer layer, the
1260 molecular layer, of the cerebellar cortex is stained green and the inner layer, the granular
1261 layer, is blue. **(B)** Output of a procedure used to group mossy fibre terminal into clusters
1262 (19 fig 6), reproduced with permission. All terminals belong to one mossy fibre, which
1263 terminates in two strips. Clusters are shown in different shades of grey. **(C)** Schematic of
1264 an area of the folded cerebellar cortex. The strip inside the red dotted line is the width
1265 and depth of a microzone but only part of its length. Mossy fibre signals are not received
1266 at a point but by a volume of the granular layer, at several randomly spaced, sagittally-
1267 aligned locations. Yellow zones are terminal locations of a single mossy fibre. A mossy
1268 fibre mediolateral axon may give rise to several sagittal branches, of which here only one
1269 is shown. Size of locations varies depending partly on the number of terminals in each
1270 cluster. Pink zones are terminal locations of a second mossy fibre. In reality, zones
1271 overlap extensively and terminals are intermingled. The zone enclosed by a blue line has
1272 the dimensions of a 'field'. A field is a nominal division of the granular layer with the
1273 average dimensions of a mossy fibre terminal cluster, which we use to quantify granular
1274 layer computations. A sagittal row of 100 fields is a 'rank'. **(D)** Section of a microzone
1275 viewed obliquely from the cerebellar surface, to illustrate the size and random
1276 intermingling of mossy fibre terminals. The molecular layer is represented by squares of

about the dimensions of a Purkinje cell territory. The wire frame boxes are an exploded view of the orange fields. The right box shows randomly distributed terminals of 5 mossy fibres (each a different colour). We estimate that ∼180 mossy fibres contribute terminals to a field, and a field contains a total of around 700 terminals. Proportions of terminal size to field size are preserved. The left box shows terminals of 20 mossy fibres.


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

1283

1284 **Fig 2. Golgi cell ensemble conversion. Column 1** We randomly generated (3,645) mossy
1285 fibre rates with a uniform, normal, twin-peaked or discontinuous distribution
1286 representing input to a row of 100 fields. Rates were randomly sampled to represent
1287 input to individual fields and copies added to reflect variable size of terminal clusters and
1288 that most clusters straddle field boundaries. Red data represent copies-included signals
1289 received by a row of three fields. **Column 2** The first-column red data were randomly
1290 sampled 150 times, sample size 4, representing mossy fibre input to a 3-field population
1291 of 30 Golgi cells, each with 5 basal dendrites (the mean),and each of those receiving
1292 contact from 4 mossy fibres. Data are the sample means, representing dendritic
1293 depolarisation. **Column 3** We next randomly sampled column 2 data without
1294 replacement 30 times, sample size 5, and took the sample means, representing somatic
1295 integration of dendritic signals. **Column 4** We then took 700 random samples of column 3
1296 data, sample size 8–12, reflecting convergent input from 3 fields of Golgi cells onto 700
1297 glomeruli in the middle field (to which a three-field ensemble has output). Data are the
1298 sample means. **Column 5** The column 4 data are presented as a histogram.

1299

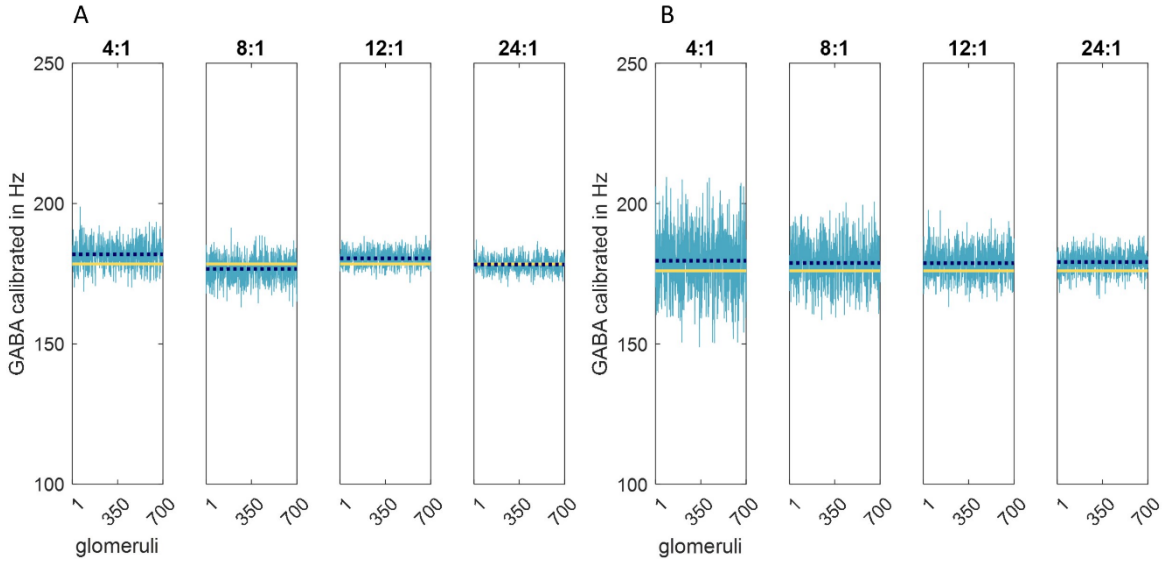
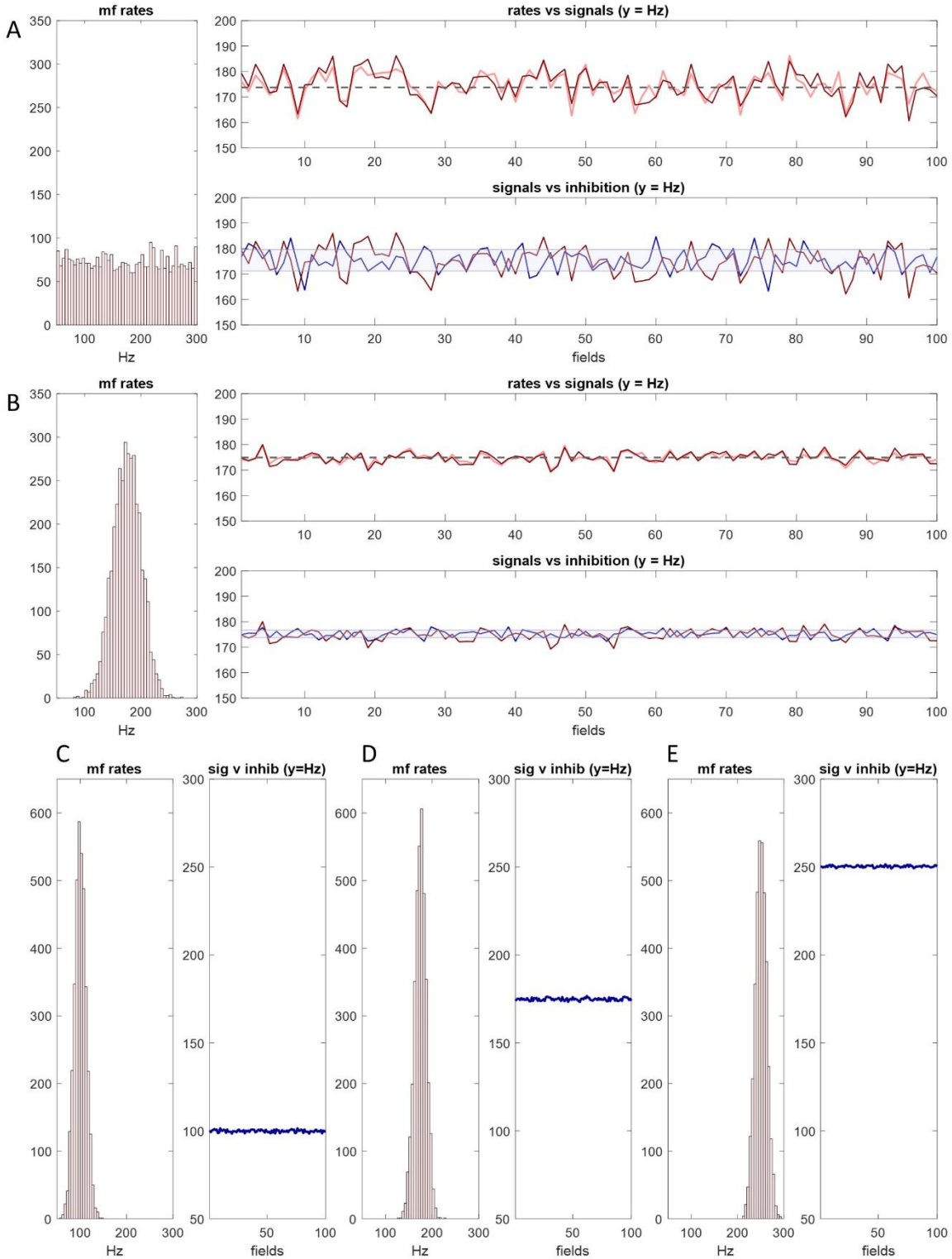


Fig 3. Testing unreported parameters shows that reliability and precision are chiefly impacted at different steps in the computation. All data were obtained in the same way as Fig 2 column 4, but varying ratios for which physiological parameters have not been reported. Data points are accordingly the means of randomly sampled Golgi cell rates received as convergent input to each of 700 glomeruli, the population of the middle field of a Golgi cell ensemble territory. Broken dark blue line: the mean of the sample means. The sampled distribution was generated in the same way as Fig 2 row 4. (The sampled distribution is the population of mossy fibre input rates to a row of 100 fields which includes the ensemble territory.) Yellow line: the mean of the sampled distribution. Varied ratios are the convergence ratio of mossy fibres onto a single Golgi cell basal dendrite (4:1 in Fig 2) and the convergence ratio of Golgi cells onto a single glomerulus is (a random number in the range 8–12:1 in Fig 2). In **(A)** we varied the first and in **(B)** the

1314 second. The ratios are shown above each panel. y axes cover a shorter range than Fig 2.
1315 Blue data – representing inhibition of granule cells – are the output of the ensemble
1316 transformation. A narrower range represents tighter within-field synchrony of inhibition
1317 received by granule cells. Findings were: (1) In both cases, there is a diminishing return of
1318 increasing the ratio. There is a modest return of higher ratios than Fig 2, selected to be
1319 physiologically plausible. (2) The blue mean varies in a small range around the yellow
1320 mean, but within that, varies at random each time the simulation is run. In (A), it varies
1321 from panel to panel because mossy fibre input rates to the simulated ensemble territory
1322 are resampled each time. In (B), that step in the computation is performed only once and
1323 the output used for all panels, so that the effect of changing the second ratio is not
1324 compounded with changing the first. Here, the blue mean is the same in all panels, but
1325 moves randomly when the simulation is rerun, in the same range as (A). So, reliability
1326 (reproducibility of the glomerular field mean) is impacted at the step of sampling of
1327 mossy fibre rates by Golgi cells, but not sampling of Golgi cell rates by glomeruli;
1328 reliability was not obviously affected by varying the first-step ratio (not shown). Precision
1329 (narrowness of the glomerular range) is impacted by both ratios, but chiefly at the
1330 second step.

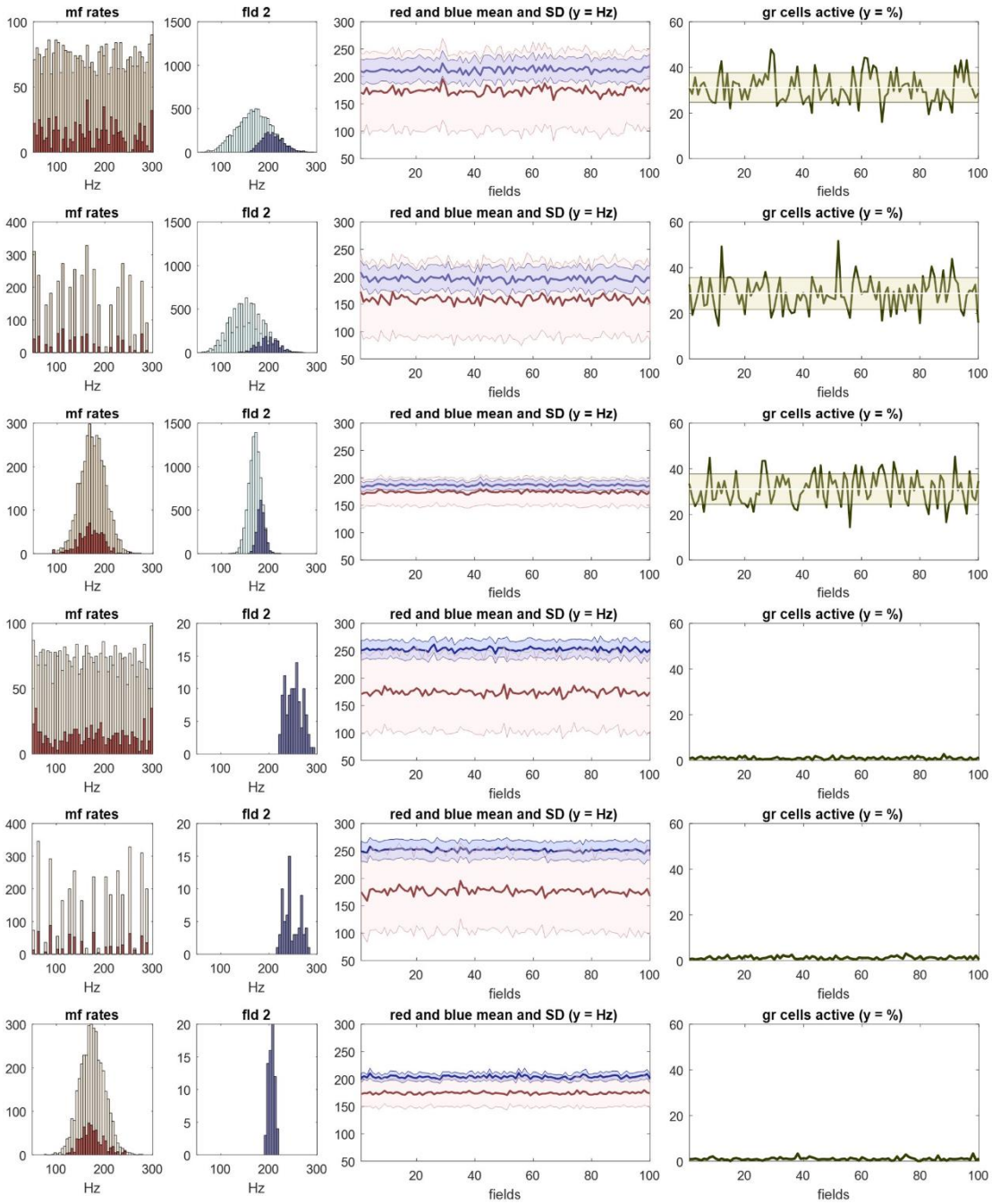
1331


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

1332

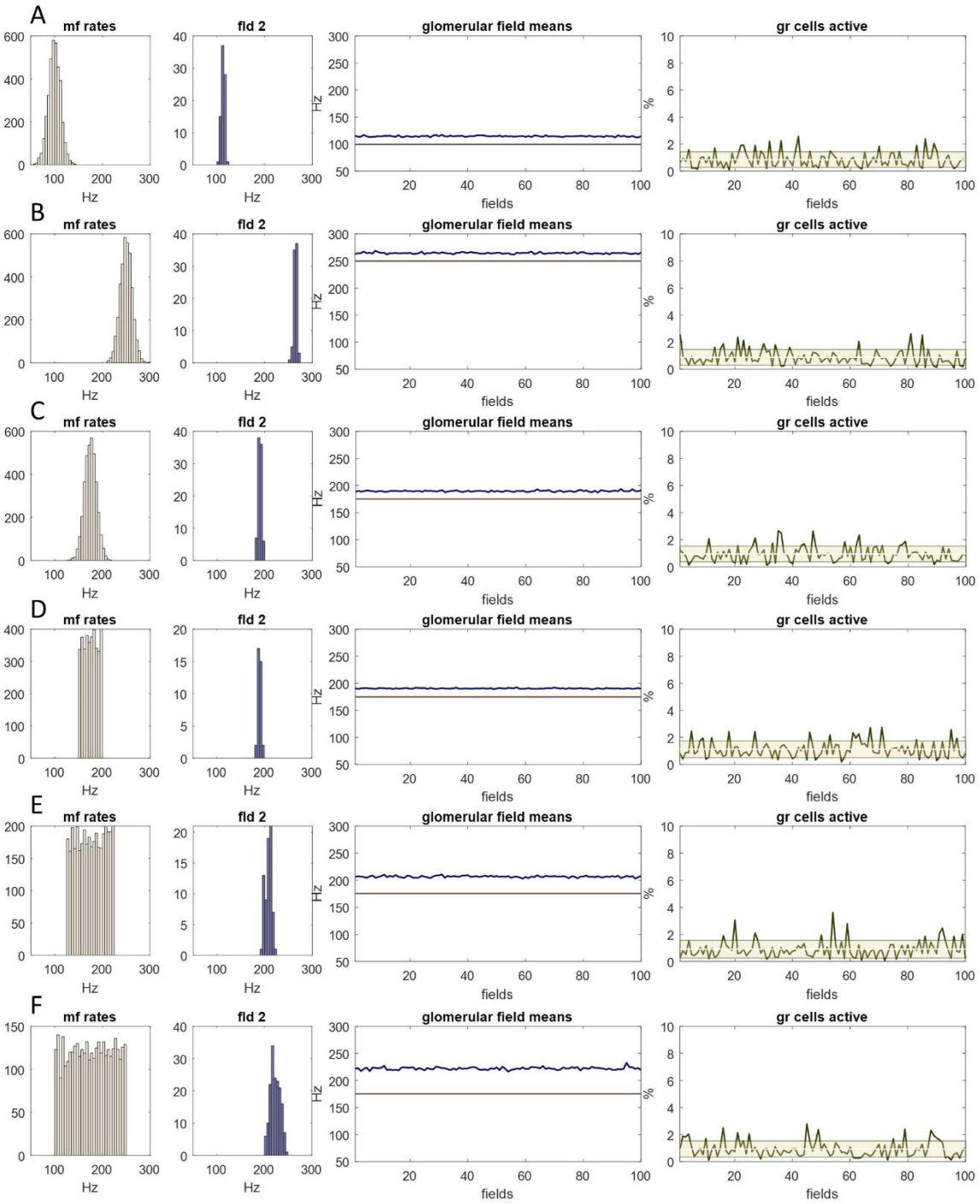
1333 **Fig 4. A whole rank of Golgi cells work as a functional unit. Synchronised inhibition of**
1334 **granule cells is precisely and linearly related to the mean of mossy fibre rates. (A)** and
1335 **(B)** Simulation of conversion by Golgi cells of mossy fire signals into inhibition of a sagittal
1336 row of 100 fields, taking the mossy fibre distribution on the left as input. The top panel
1337 compares the mean of mossy fibre input rates to each field (pink) with the mean with
1338 copy signals added to reflect mossy fibre terminal clustering (red). The dashed grey line is
1339 the mean of the sampled distribution (i.e., the mean of mossy fibre input rates to the
1340 whole rank). The bottom panel compares the copies added data (red) with the
1341 glomerular field mean (blue line). SD of inhibitory thresholds is shown in pale blue. The y
1342 scale of the long panels is Hz. Results are: (1) In all top panels, red data closely track the
1343 pink data, indicating that adding copies does not skew the mean of firing rates received
1344 by a field. (2) Blue data across all fields are in a narrowly circumscribed range. The degree
1345 of precision depends on the mossy fibre range although the conversion works with any
1346 input distribution. A normal mossy fibre distribution with a 2SD range of over 100 Hz is
1347 converted to a blue range of ∼3 Hz. (3) Fields that sample the same input data (that data
1348 being: mossy fibre input rates to a rank) return the same output. This has the
1349 physiological counterpart that inhibition of granule cells is synchronised in a rank. (4) The
1350 blue range is centred on the grey dashed line – the mean of mossy fibre rates received by
1351 the rank as a whole – and varies independently of the red data. That is, field-to-field
1352 variation of mean mossy fibre input rates doesn't flow through to an effect on blue data,
1353 the output of the system. (5) We also tested setting sample size 'optimally' (i.e., higher,

at 24:1) for mossy fibres that contact a single Golgi cell dendrite, and Golgi cells that inhibit a single glomerulus. This only very modestly improved performance: higher ratios do not return a notably narrower blue range (than the range when parameters are set at known or expected physiological values). The results are not shown. **(C)–(E)** The same simulation, sampling the mossy fibre distribution on the left. The blue data are the glomerular field means, i.e., mean inhibition of granule cells in each field (inhibition of granule cells is condensed into a small range by the ensemble conversion, so we use the means to compare fields). y axes of panels on the right of each pair have the same range as x axes of panels on the left. Results: inhibition of granule cells is precisely and linearly coupled to the mean of the sampled distribution.


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

1368 **Fig 5. Incorporating the Golgi cell conversion into a model that recodes input to the**
1369 **cerebellum into internal signals.** We incorporated the Golgi cell ensemble conversion
1370 into a simulation of recoding in the cerebellar granular layer. Recoding converts input to
1371 the cerebellum from mossy fibres into firing of granule cells. **Column 1.** We randomly
1372 generated (3,645) mossy fibre rates with a uniform, discontinuous or normal distribution
1373 representing input to a row of 100 fields. Red data are rates received by a single field
1374 (field 2, selected at random), with copies added, as an example. y axes are scaled to the
1375 data. **Column 2.** Red data for each field were randomly sampled 8,750 times, sample size
1376 4, representing a per field population of 8,750 granule cells, each with 4 dendrites, which
1377 each receive contact from a single mossy fibre terminal. The field 2 sample means are
1378 displayed as a histogram (pale green, top three rows only; there are the same number of
1379 green data in each panel). Blue data represent the subgroup of granule cells that fire (all
1380 rows). In the bottom three rows the subgroup is small (y axes are rescaled) – recoding is
1381 set (by adding a constant to inhibition) where around 1% of granule cells are active.
1382 **Column 3.** The y scale is Hz. The mean (line) and SD (shading) of the red and blue data for
1383 each field in a row of 100 fields. Blue line data are GC field means. **Column 4.** The
1384 percentage of granule cells that fire in each field. The white line and shaded area in the
1385 top three panels are mean and SD respectively. Results are: (1) Blue line data across all
1386 fields are in a narrowly circumscribed range. The degree of precision depends on the
1387 mossy fibre range although the conversion works with any input distribution. (2) Fields
1388 that sample the same input data (that data being: mossy fibre input rates to a whole

rank) return the same granule cell code. That is, the frequency distribution of the mean of input rates to each active granule cell is centred on the narrowly-varying GC field mean and attracted to normal. (3) A whole rank recodes as a single unit. Any random sample of granule cells has the same mean and distribution, if, as we claim, information is collectively encoded in rate statistics of co-active cells. It is not material that fields all return a different permutation of output rates.



1396

1397 **Fig 6. Granule cell rates collectively code mossy fibre rate information. Column 1** We

1398 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, i.e., the 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.