A HYBRID MODEL OF THE CEREBELLUM: AN OVERVIEW

Mike Gilbert,[1]* R. Chris Miall[1]

[1]*University of Birmingham, School of Psychology, Birmingham, UK*

**Correspondence: mhuwg@hotmail.com*

ABSTRACT

Learning models of the cerebellum propose that the cerebellum implements an algorithm which makes iterative adjustments to synaptic transmission strength that collectively determine the response to input in learned patterns following training. We propose instead that pattern recognition and control of firing by output cells are separately handled functions which process independently variable data. This can account for the evidence without a machine learning algorithm. The model is a hybrid of physiological arguments and computational methods used to test and quantify the ideas. We argue *inter alia* that learning operates at the level of functional groups of Purkinje cells defined by their shared climbing fibre input; Golgi cells have several functions and regulation of parallel fibre activity by Golgi cells is not in the expected way; recoding of input to the cerebellum received in the granular layer converts the number of input variables (variables expressed in mossy fibre input to the system) into a much reduced number of functional variables expressed by internal signals traffic; circuits simultaneously execute separate but integrated functions of pattern memory and output coding; they are able to operate separately because the expression of data used in each varies independently of the other; output rates are not learned but controlled by recent relevant input signals in a window opened by pattern memory; the moment-to-moment probability that a Purkinje cell spikes is synchronised across a microzone; a principal function of functional organisation of Purkinje cells into microzones is to increase resolution of rate coded information received by the output cells of the circuit, and to do so in a very short integration window, so that circuit architecture can be explained partly as a device with this function.

Keywords: cerebellum, network, model, circuit, theory

1. INTRODUCTION

The cerebellum is widely thought to implement a supervised learning algorithm which uses iterative adjustment of parallel fibre synaptic weights either to generate binary output (Albus 1971, Brunel, Hakim et al. 2004) or to reduce deviation from a desired outcome (Fujita 1982, Dean, Porrill et al. 2010). Supervision is typically provided by error signals provided by climbing fibres. Circuit function on these principles assumes that output is driven by remembered patterns of active parallel fibres. The unit of learning is a Purkinje cell, storage and expression of pattern memory are independent of other Purkinje cells, and Purkinje cell rates are treated as output. Output is learned, that is, training displaces the naïve response to input rates.

We are in conflict with this view on all points. Instead, we suggest, the cerebellum is designed to separate the functions of pattern detection (which cues output) and control of firing of output cells (which codes output), so that pattern memory has a gating but otherwise permissive role and output rates are controlled separately. Separation of the functions of pattern learning/detection and control of output rates requires that the data they each use are represented as different and independent variables, so that there is not mutual interference of the execution of either function with the other. It is partly for this reason that mossy fibre input to the cerebellum is recoded in the granular layer: to cause the number of variables coded in the binary pattern and permutations of rates received as input into a dramatically reduced number expressed by internal signals activity, in parallel

fibre input to modular circuits. This is also to prevent interference by redundant external variables, variables for which the cerebellum has no use.

This is only half of the picture. The cerebellum also has the function of turning linear rate coded data (spikes/*t*, so that the speed of data transfer is limited by time *t*) into a form that can be transmitted much faster, so that output rates can be under finely graduated control in a short nuclear integration window. It does this by converting input rate variables into parallel fibre signals with a fixed bandwidth in an adjustable range that follows the mean of mossy fibre input rates, and by coordinating Purkinje cell firing on a very short timescale across the population of a microzone. This is a reason for the functional organisation of Purkinje cells into microzones, and of nuclear projection neurons in groups. Coordination of Purkinje cells is a functional imperative and accounts for duplication of mossy fibre signals in sagittal strips (mossy fibres have a sagittally-aligned terminal branching pattern), and multiple random samplings of input rates by granule cells.

Memory and control of output are not at the level of single Purkinje cells but microzones. There is no requirement (or norm) that climbing fibre instruction signals are driven by errors (also unnecessary to explain the evidence), and – importantly – output rates are not learned. Synaptic weights are functionally binary, and input rates control output rates ad hoc – indeed graded weights would interfere with this function.

The proposed model is a hybrid, a detailed physiological model with support from computational modelling. The ideas presented here are a summary and not a full quantitative description of the proposed mechanism of cerebellar function (which will be published separately). The idea that function of the cerebellum will have a computational explanation has displaced other approaches. In that approach, the cerebellum is the

physiological implementation of a machine learning answer to the problem of turning input into whatever output the model says it has – an abstract solution to a hypothetical problem. We offer instead a synthesis of a physiological and mathematical model – a detailed mosaic of the evidence tested quantitatively by simulation. Explained in this way, there is no need for an algorithm to account for the evidence.

FIGURE 1

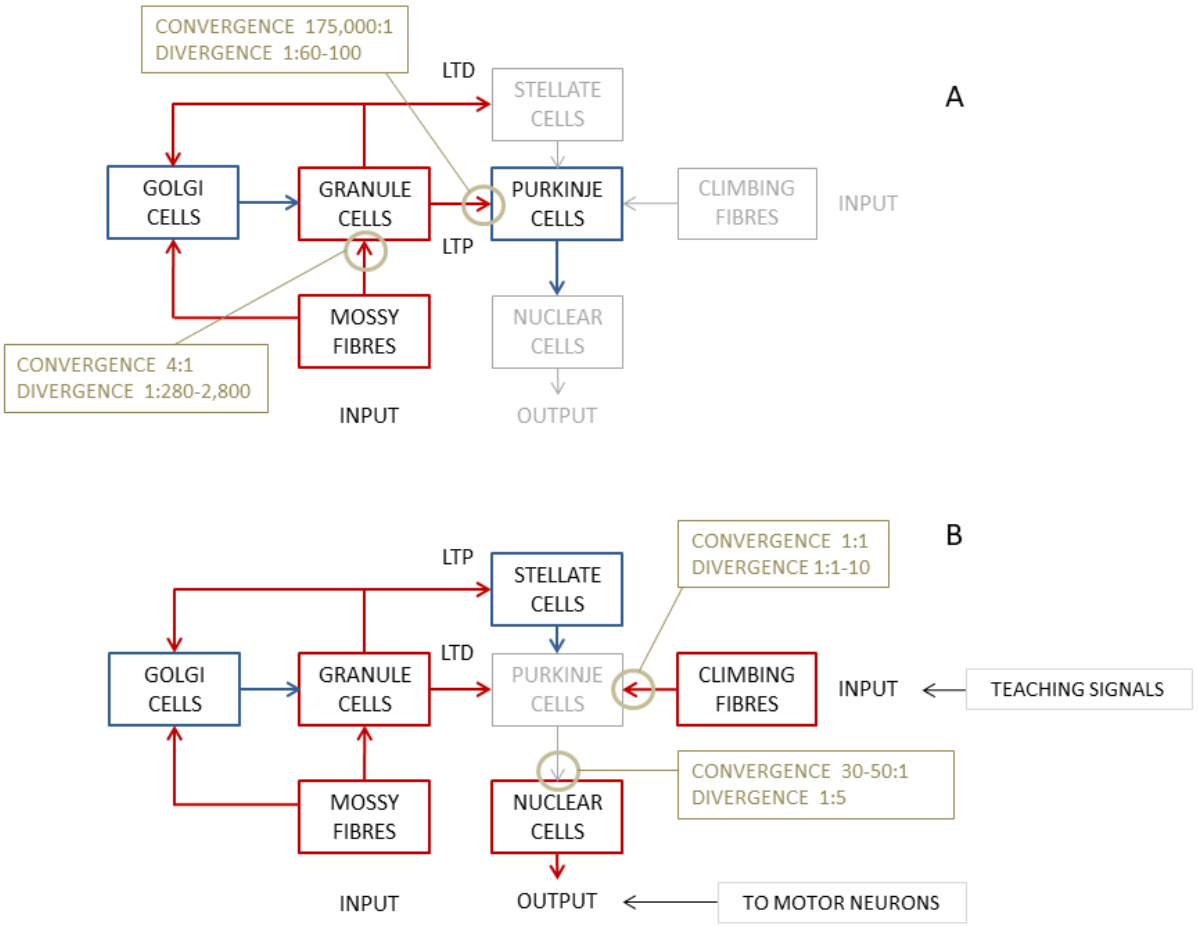


FIGURE 1

Schematic of cerebellar circuitry. A: Prior to training, strong Purkinje cell firing inhibits nuclear cells, the output cells of the circuit. Red boxes and arrows: active glutamate neurons; blue: active GABA neurons; grey: silent neurons. **B:** Following training with a conditioning protocol or direct stimulation that mimics a conditioning protocol, firing of Purkinje cells is weakened or suspended in the conditioned response partly as a result of long-term depression (LTD) of the parallel fibre-Purkinje cell synapse and long-term potentiation (LTP) of the parallel fibre-stellate cell synapse, causing a phasic reduction of inhibition by Purkinje cells of nuclear cells.

2. RECODING

2.1 Background

Most glutamatergic input to the cerebellum is carried by mossy fibres which terminate on granule cells in the inner layer of the cerebellar cortex, the granular layer. Granule cell axons rise into the outer layer of the cortex, the molecular layer, where they divide to form parallel fibres, so called because they run parallel to the surface of the cerebellum and to each other, which make contact in passing on Purkinje cells and inhibitory interneurons including Golgi cells. Golgi cells in turn inhibit granule cells. Purkinje cells are organised functionally into long thin groups of several hundred cells – microzones (Oscarsson 1979, Ozden, Sullivan et al. 2009, Bengtsson, Ekerot et al. 2011, Ramirez and Stell 2016). A microzone occupies a thin sagittal slice of the cerebellar cortex perpendicular to the surface.

The number of mossy fibre signals that drive input to a microzone is large. The area of the cerebellar cortex that provides parallel fibre input to a microzone is measures around 6 x 15 mm, the range of parallel fibres (Brand, Dahl et al. 1976, Mugnaini 1983)[1] and the length of a microzone (an estimate for the C3 region: Dean, Porrill et al. 2010) respectively. Scaled up, that's around the size of four-and-a-half tennis courts side by side supplying input to a net (the microzone) 50 m long and 0.5 m thick. Granule cell bodies would be ¼ the diameter of a tennis ball and 12-14 million of them would be packed under the court to an average (but variable) depth of 70 cm.

It has long been hypothesised that recoding of mossy fibre signals into granule cell signals sparsens and decorrelates the binary pattern of parallel fibre activity (which parallel fibres are active and which silent) (Marr 1969, Albus 1971), in order to increase storage capacity. The mechanism includes Golgi cells in the role of adjusting the granule cell 'input threshold' – the number of mossy fibre inputs needed to make a granule cell fire.[2] A higher threshold reduces the number that receive enough input to fire, and vice versa. By adjusting the threshold Golgi cells dampen the effect of variation of the number of mossy fibre inputs to the system. Golgi cells thus (in this contention) provide homeostatic control of parallel fibre activity, which is maintained at a low level (Marr 1969, Albus 1971, Billings, Piasini et al. 2014, Cayco-Gajic, Clopath et al. 2017, Cayco-Gajic and Silver 2019). So in this model regulation of parallel fibre activity is by Golgi cell adjustments of the granule cell input

[1] Reported for cats, chickens and monkeys; 5 mm is reported for rats Harvey, R. J. and R. M. Napper (1988). "Quantitative study of granule and Purkinje cells in the cerebellar cortex of the rat." J Comp Neurol **274**(2): 151-157, Harvey, R. J. and R. M. Napper (1991). "Quantitative studies on the mammalian cerebellum." Prog Neurobiol **36**(6): 437-463.

[2] 'Input' (here and generally) means contact by an active cell as opposed to mere innervation.

threshold, output is sparse, and the function is to increase storage capacity (by recoding mossy fibre signals into a larger state space).

Absent from the physiological support for this arrangement are (1) how the role of Golgi cells is implemented on a quantified physiological level that can be simulated; (2) how recoding affects granule cell firing rates; (3) whether and how control of parallel fibre activity (numbers active) and the rates they fire at are kept separate so that they vary independently; (4) how external variables (variables contained in mossy fibre input to four-and-a-half nominal tennis courts) are selected for an effect (on either active numbers or rates) or prevented from having one; and (5) what other functions parallel fibre regulation has.

We claim to answer these questions. The discussion is in two parts. In addition to recoding variables contained in the binary pattern of input to the granular layer, the mossy fibre-granule cell relay also has the function of recoding variables contained in the permutation of rates they each fire at. These are considered in turn.

2.2 Recoding variables contained in the binary pattern of input

It is part of the present proposal that the function of Golgi cells is not to adjust the input threshold of granule cells in response to input variables (and in particular in response to a changing volume of mossy fibre input) but to make it *independent* of input variables – not only the number of mossy fibre inputs to the system but mossy fibre firing rates – so that it is constant in all conditions. Part of the argument is a simulation of regulation of parallel fibre activity that shows it is confined to a fixed and low level with a fixed threshold.

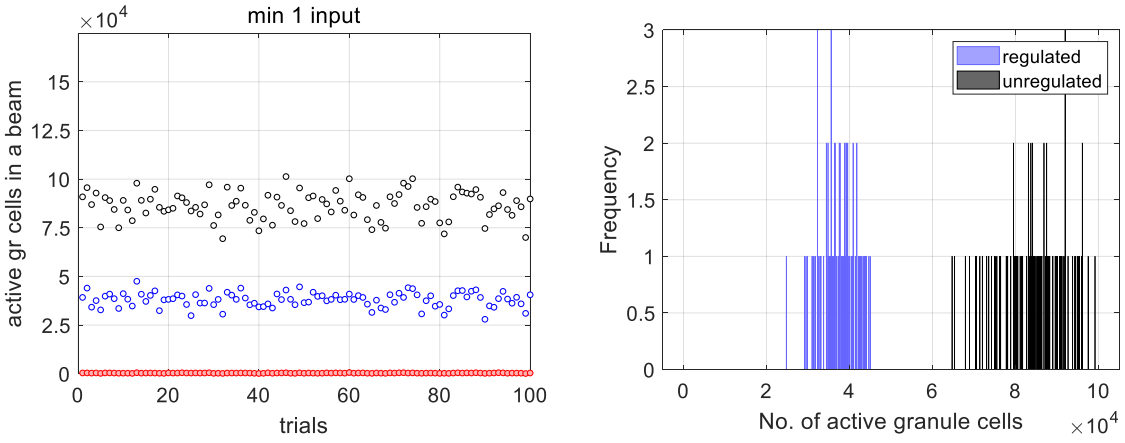
Assuming a fixed input threshold, then, the probability that a granule cell will meet threshold can be derived with a binomial from the proportion of active mossy fibres out of the total that supply that location. This gives the expected number that fire because the law of large numbers holds that at large numbers of independent trials, the ratio of the outcomes will converge towards the proportions predicted by their probability.

Granule cells have 3-5 dendrites (average 4) each of which receives contact from a single mossy fibre. Assuming contact is at random (i.e. mossy fibres do not select which granule cells they contact) it is very likely that a granule cell receives input to each of its 4 dendrites from 4 different mossy fibres. Mossy fibres branch terminally. Each branch ends in a cluster of terminals (average 7-8 terminals per cluster) (Wu, Sugihara et al. 1999, Sultan and Heck 2003, Shinoda and Sugihara 2013). An estimated 100 mossy fibres terminate in a region the size of cluster field so that a cluster field contains an average of around 700-800 terminals (Sultan and Heck 2003). Each terminal receives a single dendrite from each of what may be around 50 granule cells (Jakab and Hámori 1988, Gao, Proietti-Onori et al. 2016) though estimates vary (Ritzau-Jost, Delvendahl et al. 2014).

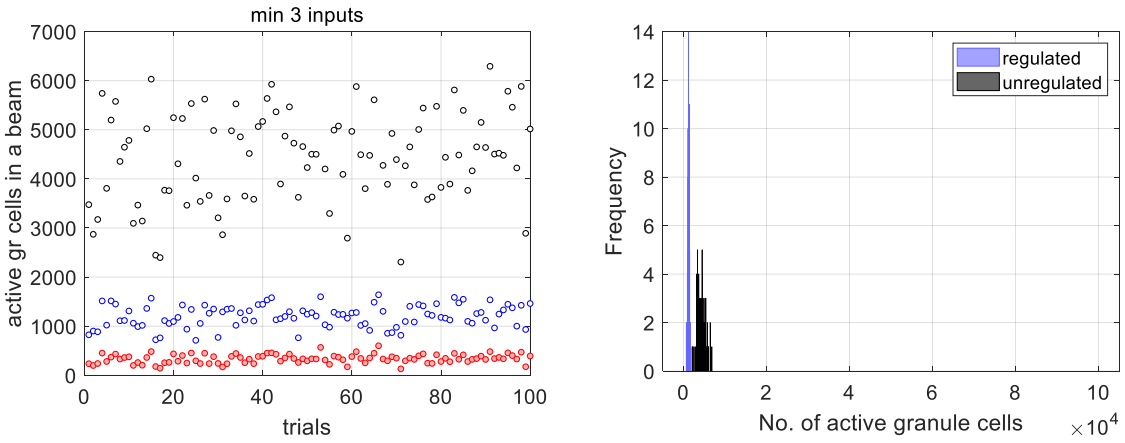
Golgi cells inhibit the mossy fibre-granule cell relay in an arrangement termed a glomerulus. The glomerulus is ensheathed by a semi-permeable membrane which restricts neurotransmitter diffusion. Inhibition has a strong tonic component (Duguid, Branco et al. 2012). Activation of presynaptic $GABA_B$ receptors on mossy fibres inhibits glutamate release (Mitchell and Silver 2000a), and activation of presynaptic mGluR2 receptors on Golgi cells inhibits GABA release (Mitchell and Silver 2000b). The result is competitive mutual suppression whose outcome depends on phasic adjustments of the neurotransmitter balance (Kanichay and Silver 2008, Cesana, Pietrajtis et al. 2013).

FIGURE 2

A



B



C

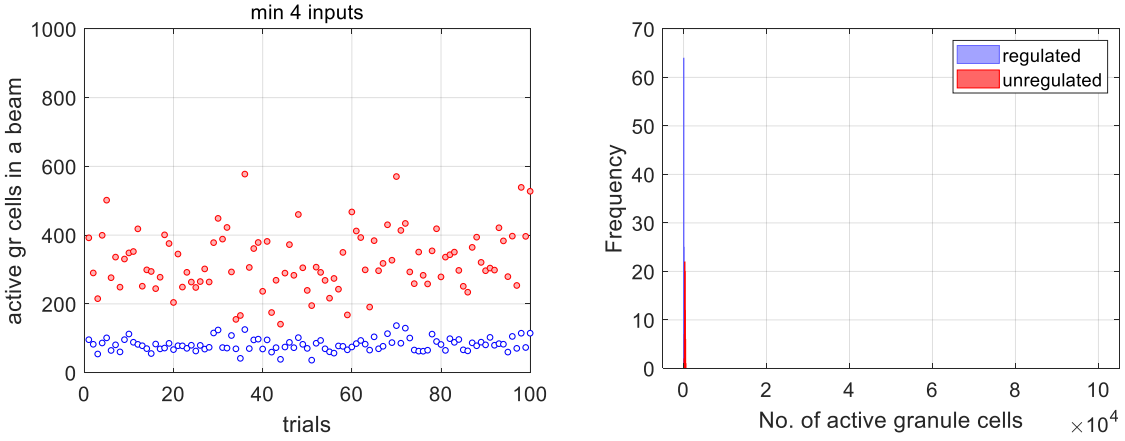


FIGURE 2

Homeostatic regulation of parallel fibre traffic by Golgi cell regulation of granule cell firing. The model estimates the total number of excited granule cells in a mediolaterally-aligned row of 20 mossy fibre terminal cluster fields (together measuring 3,000 x 200 µm), trial by trial, in each of 100 trials. Each field receives input from a random number of mossy fibres in the range 3-30 (out of 100) and contains 8,750 granule cells. The minimum number of mossy fibres needed to make a granule cell fire is 1 in panel A, 3 in panel B (the probable physiological input threshold) and 4 in panel C. Black dots are the estimated number of granule cells meeting the input threshold, disregarding Golgi cell regulation; the number with Golgi cell regulation are in blue, and the number which receive 4 mossy fibre inputs is red, identical in all graphs, for comparison. Note y axes are scaled to the data. The histograms are frequency distributions that show the same black and blue data, and in C the blue and red data. With an input threshold of 3 the average regulated number of active granule cells in a beam is approximately 1,200 out of perhaps 250,000. Of those only a few tens receive 4 inputs – the blue data in panel C. The very large majority all receive 3. This forms part of the control of variables that (would otherwise) affect translation of mossy fibre to granule cell rates.

So, meeting the input threshold is not alone sufficient to make a granule cell fire. Inputs must each be strong enough to be competitive – or, put another way, to prevail over the inhibitory veto – with the combined effect, integrated at the soma, of depolarising the target to firing threshold.

Assuming inhibition vetoes an effect of excitatory input to a glomerulus with a probability $P(v)$, and x out of y mossy fibres are active, the probability that a particular dendrite receives excitatory input and no veto is $(1 - P(v))(x/y)$. Given n = 4 dendrites per granule cell and an input threshold m, the expected number of granule cells in a cluster field (say f_1), out of a total t, that fire is

$$E(f_1) = t\left[\sum_{k=m}^{n} \frac{n!}{k!\,(n-k)!} * \left((1 - P(v)) * \frac{x}{y}\right)^k * \left[1 - \left((1 - P(v)) * \frac{x}{y}\right)\right]^{n-k}\right] \quad (1)$$

where y is a physiological constant (100) and x is in a range 3-30 generated randomly for each field in each trial. t (also a constant) is 8,750 (calculated from convergence of 4:1 and divergence of 1:50).

Golgi cells receive excitatory input from local mossy fibres and from parallel fibres. There is as a result a mutual effect mediated by Golgi cells on fields aligned mediolaterally, joined by parallel fibres, and in a row of fields there is an effect of all fields on all others. $P(v)$ is derived from the density of parallel fibre activity. As a Golgi cell receives only a small sample of passing parallel fibre activity, the number of active parallel fibres that make contact is determined by chance, with a distributed probability. Even Golgi cells on the path of the same parallel fibre traffic are unlikely to receive the same number of inputs and very unlikely to receive contact from the exact subset (or in fact any) of the same active cells. The

probable number that make contact[3] is accordingly a measure of, and quantifies, the effect of raising or lowering density of parallel fibre activity.

The Figure 2 data are derived by summing activity in a row and using it to calculate the influence of fields on each other, then rerunning the calculation for each field to get a new total, and repeating this procedure to home in on a stable value. Other parameters included in the calculation – different for each field and with an effect recalculated in each iteration – include incursion of Golgi cells from neighbouring fields, direct input to Golgi cells from mossy fibres, and the higher probability that local granule cells make contact on a Golgi cell than distant granule cells (that make only parallel fibre contact).

The result – the blue data in Figure 2 – is that parallel fibre activity is confined to a narrow and stable range regardless of the (fixed) granule cell input threshold. There is no need to adjust the input threshold to keep output stable, or to keep it low. So: Golgi regulation works in the reverse of the expected way (in this contention). It does not have the function of adjusting the input threshold (to maintain self-rectifying parallel fibre levels), but of keeping it the same.[4]

Regulated parallel fibre activity is low regardless of the shifting underlying number and binary pattern of mossy fibre inputs, as previously predicted (Marr 1969, Albus 1971, Billings, Piasini et al. 2014, Cayco-Gajic, Clopath et al. 2017, Cayco-Gajic and Silver 2019), but the mechanism of regulation is different. This is not merely a detail of implementation. It makes an important difference to the function it contributes to recoding, and in turn to

[3] More accurately, the probabilities associated with each possible number.
[4] To achieve stable density of parallel fibre activity, it is not, at least proportionately, reflected in the active fraction of the local population, but derived at the scale of the whole strip that provides parallel fibres that pass overhead.

what we can infer about the function that recoding contributes to circuit operation. It was previously thought the function of recoding (in models that need the number of active parallel fibres to be a low fraction) was to increase storage capacity by sparsening and separating recoded patterns of activity. We agree it increases capacity, but not on the way that regulation works or on how pattern memory is stored and expressed. Nor is it the only function recoding has, or correct to think it has more important status than other functions.

2.3 Recoding variables contained in the permutation of input firing rates

The probable physiological granule cell input threshold is 3 (Jörntell and Ekerot 2006, Billings, Piasini et al. 2014). Tonic Golgi cell inhibition of glomeruli means that only stronger signals are transmitted. Assuming contact by mossy fibres on granule cells is at random, granule cells randomly sample mossy fibre frequencies. For those that meet the input threshold the sample size is 3. The number of granule cells that receive contact from 3 active mossy fibres[5] is large, so that if the mossy fibre rates in each sample are averaged (giving the average rate received by that granule cell), the frequency distribution of the sample means is (near) normal by the central limit theorem, even if mossy fibre input to the population is modest, and regardless of the physical distribution of active mossy fibres and range and frequency distribution of the rates they fire at.

The distribution is only approximately normal because the number of mossy fibre rates in a sample is small. This raises the question: why don't granule cells take a bigger sample? A possible reason is that a low number of inputs to granule cells is necessary for optimised

[5] In a mutually regulating strip connected overhead by parallel fibres.

pattern separation and decorrelation (Cayco-Gajic, Clopath et al. 2017). However, a sample size of 3-5 is sufficient to transform the shape of the original distribution (the frequency distribution of mossy fibre rates) into a recognisably (if untidy) bell-like distribution of the sample means.[6] The number of samples in Figure 3 is 4,500, an approximate estimate (derived from Figure 2) of the number of granule cells that receive 3 or 4 inputs.

Only samples that contain strong signals cause a granule cell to fire, because excitatory inputs must individually be strong enough to win the glomerular competition (and collectively they must be strong enough to depolarise the cell). These form the 'top slice' of the distribution of the sample means. Thus the mean of input rates to the subset of granule cells that fire has a normalised frequency distribution (Figure 3): meaning, the bandwidth and shape of the distribution are fixed regardless of the range and frequency distribution of mossy fibre rates, and of the number and binary pattern of active mossy fibres (providing inputs to the system).

Also, the bandwidth is much narrower than the range of mossy fibre rates. Because the mean of the sample means is equal to the mean of the sampled distribution (that is, of mossy fibre rates), again by the central limit theorem, and the distribution of the sample means is normal, granule cell rates vary with the mean of mossy fibre rates (because a normal distribution follows the mean).

A key function of Golgi cells (in this contention) is to make regulation of the density of active parallel fibres (discussed in the last subsection) independent of mossy fibre rates as well as of the number and binary pattern of active mossy fibres providing input to the system.

[6] Another possible reason may be that a larger number of short dendrites would increase the probability that the same mossy fibre would be resampled.

Generally stronger (say) mossy fibre rates might be expected to mean fewer inputs to a granule cell are needed to make it fire, so that more should fire (because it is more probable), other things being equal. Instead (we claim that) an increase in mossy fibre frequency, to the extent it makes more granule cells fire, has a self-rectifying effect on the granule cell input threshold. That effect is mediated by a shift in the probability distribution that predicts the number of active parallel fibres which contact a Golgi cell, increasing the probability of stronger inhibition of granule cells, with proportionate downward pressure on the number that receive 3 competitive excitatory inputs, and therefore on the number that fire.

This creates a loop of mutually regulating probabilities which fix the input threshold as well as regulating (density of) parallel fibre activity, and can be quantified and modelled to give the regulated fraction of active parallel fibres. Because the self-regulated density of parallel fibre activity is independent of firing rates, output rates vary faithfully with input rates (where input and output here mean input and output of recoding).

It is a further function of Golgi cells is that this adds a second layer of randomisation to decorrelation (the first layer is provided by random contact by mossy fibres on granule cells). As a result, patterns are uniformly dense, because active parallel fibres are randomly distributed in the sagittal plane. Moreover, because the regulated level of parallel fibre activity is standard (because it is independent of input variables), density is ubiquitous and time-invariant. Fixed density[7] has important functional roles. Some of these are discussed after the next section.

[7] This does not mean there is a uniform distribution or fixed density of active granule cells in the underlying granular layer, or even that a fixed fraction is active of the local number that meet the input threshold.

FIGURE 3

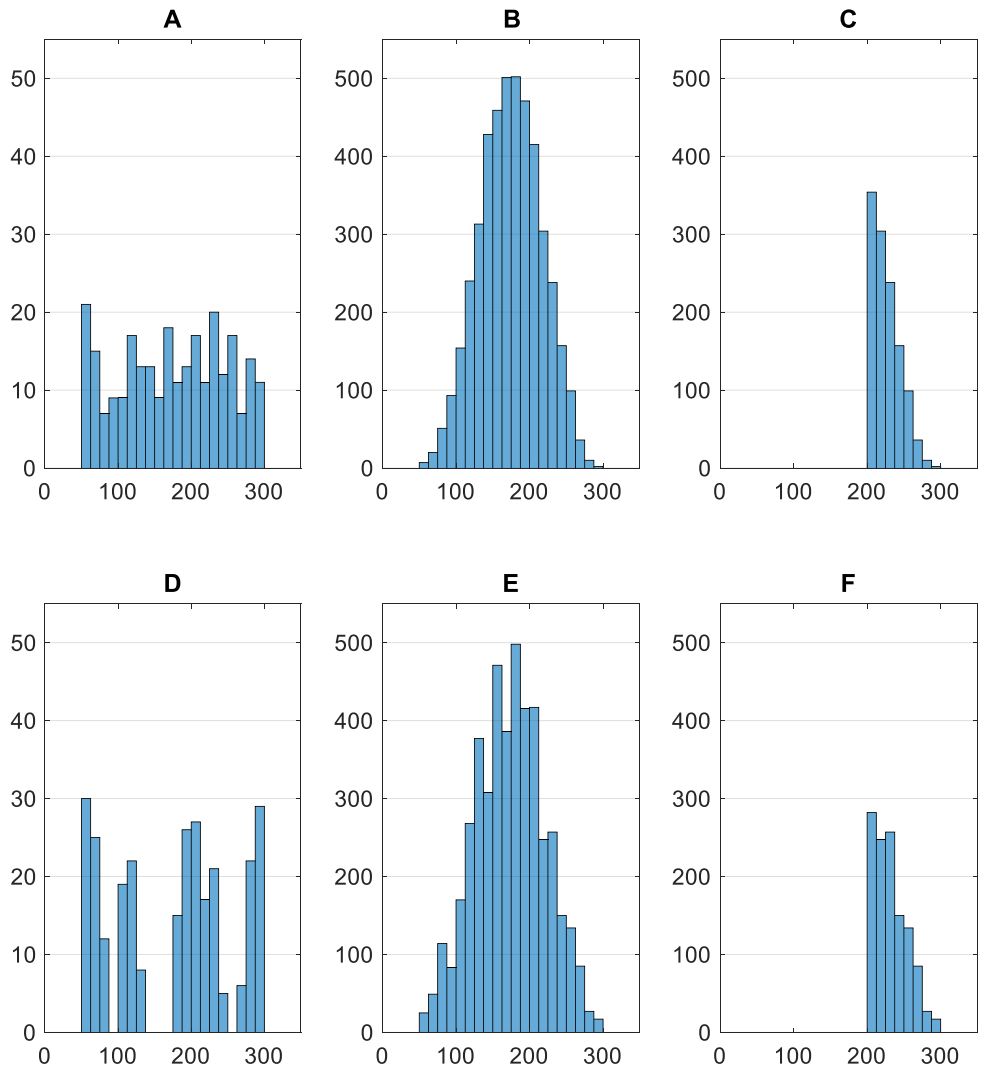


FIGURE 3

Random sampling by granule cells of mossy fibre rates. Panels **A** and **D** show a simulation of randomly distributed firing rates of a population of 300 mossy fibres, in the range 50-300 Hz, thought to be the range for activation under behaviour. The probability of firing is manipulated in B to create a discontinuous distribution. 300 is the average number of mossy fibres that are active in Figure 2 out of the population of 2,000 that innervate a mediolateral row of 20 mossy fibre terminal cluster fields (a 200 µm x 3 mm strip: a 'beam').

In panels **B** and **E** the A and D data, respectively, are each randomly sampled 4,500 times, sample size 3. Each sample represents the input to a granule cell. A minimum of 3 is needed for a granule cell to fire (but is not enough on its own). The number of samples (4,500) is our previous estimate of the number of granule cells in a beam that receive contact from either 3 or 4 active mossy fibres (the large majority 3, hence the sample size). The sample means (the mean of the rates received by each granule cell) are plotted as a frequency distribution. The arbitrary distribution of mossy fibre rates in A and the discontinuous distribution in D are both converted to an approximately normal distribution by the central limit theorem.

A subset of this group – which receive high mean rates of excitatory input – fire. The distribution of the sample means for this group is shown in panels **C** and **F**. Each top slice contains 1,200 granule cells, the average parallel fibre activity estimated in Figure 2. The bandwidth of this group is narrower than the mossy fibre range and the shape of the distribution is independent of the mossy fibre distribution. The range of the top slice, but not the bandwidth, varies with the mean of mossy fibre rates (because the mean of the sample means approximates the mean of the sampled population).

2.4 Linear translation of mossy fibre rates into granule cell rates

A final function of Golgi cells – by preserving a stable input threshold – is that it means granule cell firing rates do not receive a variable effect from the number of excitatory inputs they receive. To repeat: three inputs – the probable input threshold (Jörntell and Ekerot 2006, Bengtsson and Jörntell 2009, Billings, Piasini et al. 2014) – are not alone sufficient to make a granule cell fire. Only a subset fire, made up, broadly speaking, of those receiving

the strongest aggregate convergent input, where 'strong' refers to the sum of input rates. The average of input rates is a reliable measure of the sum because almost all (granule cells that fire) receive 3 inputs, because the input threshold is fixed and 4 inputs is improbable (and the number of granule cells that receive 4 inputs that are also competitive is very low indeed: Figure 2C).

The mossy fibre-granule cell relay is adapted to transmit short duration mossy fibre action potentials received at high frequency, and impressively adapted (in a number of ways) for high frequency transmission (Ritzau-Jost, Delvendahl et al. 2014 p.152, Delvendahl and Hallermann 2016). Also, the glomerular competition means that in theory the postsynaptic effect of a mossy fibre 'winner' is undiluted, or more weakly diluted, so that a granule cell dendrite is activated at a level proportional to the input firing rate. Moreover, other variables are controlled – each granule cell dendrite receives a single input equidistant from the soma. Granule cells 'have a relatively linear and uncomplicated conversion of depolarisation level to spike rate' (Bengtsson and Jörntell 2009 p.2393, citing Jorntell and Ekerot 2006 and D'Angelo et al 1998), such that the granule cell firing rate is reported to vary faithfully with input signal frequency (Rancz, Ishikawa et al. 2007). Thus, for a number of reasons including the fixed granule cell input threshold, granule cell rates reliably reflect (selected, by recoding) mossy fibre rates.

2.5 Isolation of operational variables

Because the density of parallel fibre activity is independent of mossy fibres rates, and translation to granule cell rates is faithful, density of the binary pattern of active parallel

fibres is independent of the rates they fire at. As a result, Purkinje cells receive no different effect from the number or binary pattern of inputs they receive, because these are functionally invariant, the first because it is constant and the second because it is random, and because parallel fibre contact is exclusively on spines on distalmost branchlets of the Purkinje cell arbour. The only functional difference (between binary patterns) is that each is specific to a different number and pattern of mossy fibre inputs (and permutation of mossy fibre rates, since that too is coded in the binary pattern of parallel fibre activity). Similarly, functional variation of the range, bandwidth, permutation and frequency distribution of mossy fibre rates are removed from recoded internal – that is, parallel fibre – activity at the scale of input to a Purkinje cell.

Thus in addition to sparseness of recoded parallel fibre activity, the granular layer causes the number of functional variables expressed in parallel fibre activity to be dramatically cut down to (1) input specificity of recoded patterns; and (2) the adjustable range (as opposed to bandwidth) of the top slice, expressed as granule cell rates. It also procures that these vary independently. This permits those variables to be used in different functions without interference by the execution of either function with the operation of the other, despite being closely functionally integrated. This suggests a challenge to the idea that synaptic weights are a filter that controls output rates, because it obviates the assumption (by learning models) that rates are controlled by learned patterns.

The isolation of rates – by removing an effect of other variables on transmission – is consistent with the reported linear rate coding in the cerebellum (for example Rancz, Ishikawa et al. 2007, Jelitai, Puggioni et al. 2016), and the consistently-reported linear relationship of firing rates and task-related parameters. ‘The firing rate of many cerebellar

neurons is a linear function of task related parameters…[and this] has been found at all levels of the cerebellar circuit' (Raymond and Medina 2018 p.239, who provide references). We note in this connection that temporal coding is reported to be absent at the parallel fibre-Purkinje cell synapse (van Beugen, Gao et al. 2013), seemingly leaving rates to control the simple spike rate, and through the simple spike rate the firing rate of nuclear projection neurons that carry the output of the circuit.

2.6 Pattern memory is stored at circuit level

Assuming regulation is ubiquitous, so that the pattern of parallel fibre activity is randomly and therefore uniformly distributed in the sagittal plane, input to a microzone is evenly distributed along its full length. As a result, a climbing fibre volley (microzones are defined by their climbing fibre input, so that climbing fibre input is received as a broadside by the entire Purkinje cell population by definition) is inevitably paired with convergent and uniformly dense parallel fibre input all the way down. Pattern memory trained with a conditioning protocol is therefore stored across a whole microzone – memory is stored and presumably therefore expressed at circuit level rather than at (what learning models propose is) the level of single Purkinje cells.

Finally, 'density' of active parallel fibres is shorthand for the proportion that are active, rather than number per unit area. We mention this distinction because a stable number per unit area would mean the variable size of the Purkinje cell arbour (which varies substantially between peaks and furrows of the folded cerebellar cortex) (Eccles, Ito et al. 1967) would affect the number of active cells that make contact – so that the number would be a

variable. A stable proportion does not have this effect because the size and shape of the Purkinje cell arbour and the thickness of the underlying granular layer co-vary (Eccles, Ito et al. 1967) – the granular layer is thinner in furrows and thicker in peaks, commensurate with the larger Purkinje cell arbour size in peaks.

3. BINARY MEMORY OF CEREBELLAR CIRCUITS

3.1 Background

Near coincidence of climbing fibre and parallel fibre stimulation (therefore 'paired' input) induces long-term depression of the parallel fibre-Purkinje cell synapse (Hansel, Linden et al. 2001, Ito 2001, Qiu and Knopfel 2009): the synapse is 'trained'. Learning theories of the cerebellum propose that, following training, Purkinje cells acquire a learned response to input in a known pattern as a result of incremental synaptic weight adjustments trained under climbing fibre tuition (Albus 1971, Fujita 1982, Brunel, Hakim et al. 2004, Dean, Porrill et al. 2010). In this way learning displaces the naïve response to input rates, contributing to a learned effect on motor output. Modified synaptic weights determine the response to a particular set of input signals, which in that sense the system remembers. Adjustments are made by an algorithm. Pattern memory and control of output coding are inseparable aspects of the learned response, both driven by a single binary pattern of active input cells (and permutation of input values).

We propose instead that pattern matching is at circuit level, rather than the level of individual Purkinje cells. There is no memory of individual patterns. Discrimination is at the level of, and between, the whole class of known patterns, and the residual class of all other

patterns. Memories are not stored as graded synaptic weights. The response does not discriminate between patterns within a class. There is no graded or intermediate response to a partial match. A fraction of the Purkinje cells which innervate a nuclear group is sufficient to veto output of the circuit – if any part of a known pattern (received all along a microzone) is a mismatch, it blocks the response of the whole circuit. This is not to suggest that the output of the circuit is binary, but that the mechanisms of pattern detection and control of nuclear rates are separate. The function of a determination – match or not – is to select which circuits have output and when, but does not control the nuclear rate. The response to a match is permissive – permitting but not coding output. Overlap of stored patterns is well tolerated and useful.

3.2 A small number of Purkinje cells is sufficient for strong contact on a whole nuclear group

The substantial majority of the output of the cerebellar cortex which converges on a functional group of nuclear cells is from the same microzone or the same functional but dispersed group of microzones which form part of a multizonal circuit (Pantò, Zappalà et al. 2001, Apps and Garwicz 2005). Purkinje cells fire spontaneously at robust rates (Häusser and Clark 1997, Raman and Bean 1999, Cerminara and Rawson 2004, Zhou, Lin et al. 2014), and individually make powerful contact on each of their nuclear targets via 24-36 boutons, each containing multiple synaptic densities (Telgkamp, Padgett et al. 2004, Person and Raman 2012a). 70 out of 86 boutons examined with electron micrographs of mouse medial and lateral cerebellar nuclei had multiple synaptic densities (Telgkamp, Padgett et al. 2004). Out of 10 boutons reconstructed from a single slice all 10 had multiple synapses with an average

of 9.2 ± 1.3 densities per bouton. Purkinje cells outnumber nuclear cells by over 10 to 1 (11:1 in mice, for example) and each Purkinje cell makes contact on 4 or 5 nuclear cells, with convergence of 30-50:1 (in rats, both referenced in Person and Raman 2012a).

The strong firing of Purkinje cells and their individually strong contact on nuclear cells mean that a single Purkinje cell may significantly impact on firing of its targets (Pedroarena and Schwarz 2003). Modulation of nuclear cell firing requires the 'substantial co-modulation of a large proportion of the PCs [Purkinje cells] that innervate the cell' (Bengtsson, Ekerot et al. 2011, abstract). We assume that output of a microzone is not topographically organised – so that a single Purkinje cell may contact at random any 4 or 5 nuclear cells (for convenience we use 5) in the nuclear target group. 'To date, there is no evidence to support [the idea] that different PCs [Purkinje cells] of the microzone control specific CN [cerebellar nuclei] cells within the micro-group [associated group of nuclear cells]' (Bengtsson and Jorntell in Apps, Hawkes et al. 2018 p. 663).

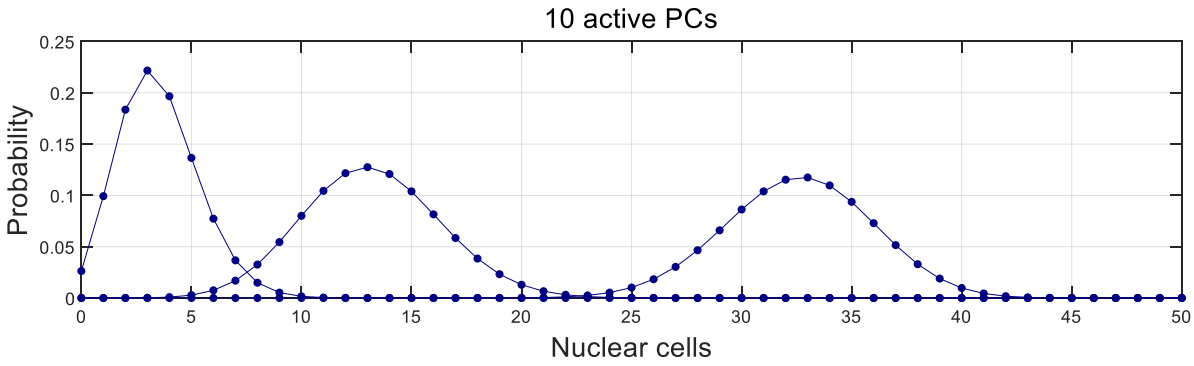
Assuming 9 synapses on a nuclear cell per Purkinje cell bouton, and contact by a Purkinje cell on 5 nuclear cells, 5 spontaneously active Purkinje cells may inhibit at high frequency as much as half a nuclear group, making around 216-324 synapses on each nuclear cell. The activation of many synapses across several boutons would represent substantial inhibitory drive, from only 2.5% of the Purkinje cells in the microzone. Adding 5 more active Purkinje cells would result in inhibition by 10 Purkinje cells of 5-50 nuclear neurons representing 10%-100% of the target nuclear group, at 216-324 synapses per cell if input is distributed to 100% of the group, or up to 2,160-3,240 synapses in the improbable event that input converges on just 10% of the nuclear group, and so on. These examples are intended to illustrate that a small fraction (2.5-5%) of the population of Purkinje cells in a microzone

may powerfully inhibit nuclear firing across a substantial proportion of the target group.

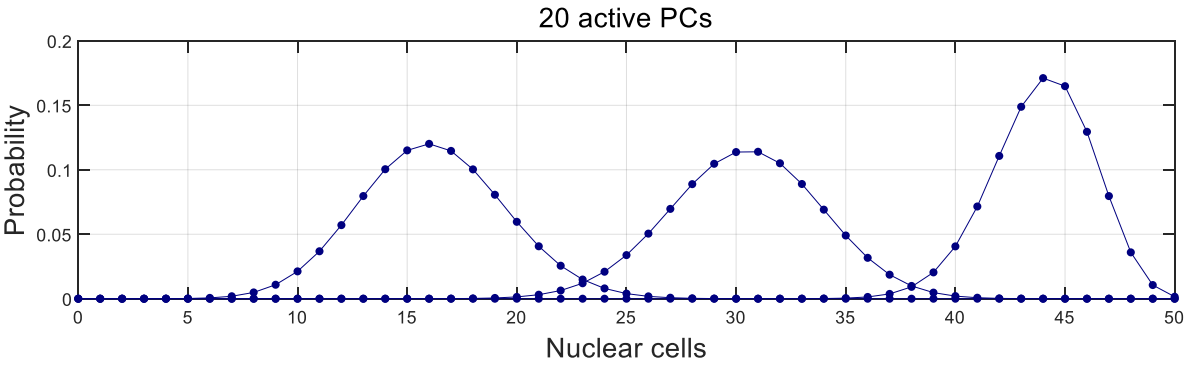
How substantial is illustrated in Figure 4.

FIGURE 4

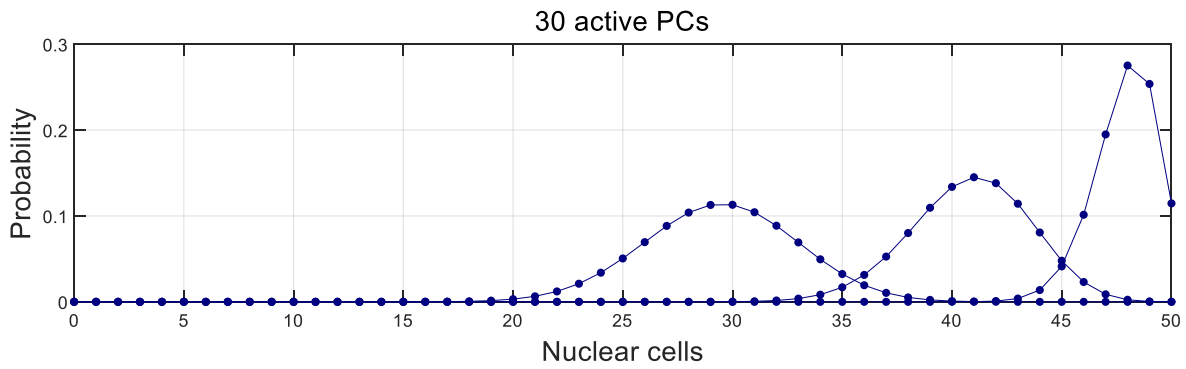
Ai



Aii



Aiii



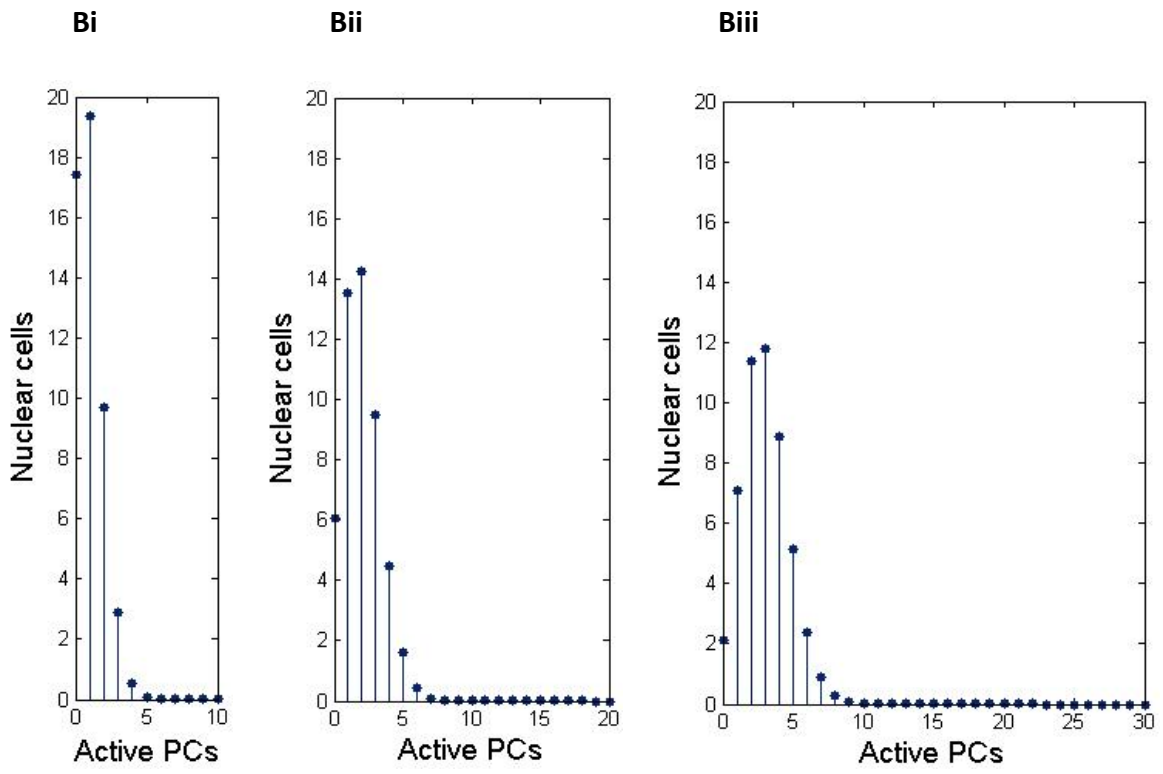


FIGURE 4

A fraction of Purkinje cells afferent to a nuclear group is sufficient to make strong contact on the whole of the group. Ai – iii: The probability that the number of nuclear cells on the x axis (in the range 1-50) receives contact from 1 or more (right peak), 2 or more (centre), and 3 or more (left) out of 10, 20, or 30 (Ai - iii respectively) active Purkinje cells. (y axes are scaled to the data.) The significant feature is how near the right the peaks are, and the narrowness of the majority of the area under the curve (indicating a narrow range of the most likely number of nuclear cells). In Aii (20 active Purkinje cells), for example, it is likely that all but a small number of nuclear cells receive contact from at least 1 Purkinje cell and the majority receive input from at least 2. **Bi – iii:** Stem diagrams showing the expected number of nuclear cells receiving contact from 0 active Purkinje cells, from 1, from 2 and so on, where the total number of active Purkinje cells is 10 (Bi), 20 (Bii) and 30 (Biii). With 10 active Purkinje cells, a

probable ~35% of nuclear cells receive no contact at all, suggesting that 10 Purkinje cells may not be enough to veto output of a nuclear group (without help from nuclear interneurons). With 30 active Purkinje cells only 2 nuclear cells receive no contact from active cells and over 80% receive contact from 2 or more.

3.3 Functional significance

The anatomy of contact by Purkinje cells on the output cells of the cerebellar circuit allows us to draw inferences about its function, assuming random distribution of contact by a Purkinje cell within a nuclear group.

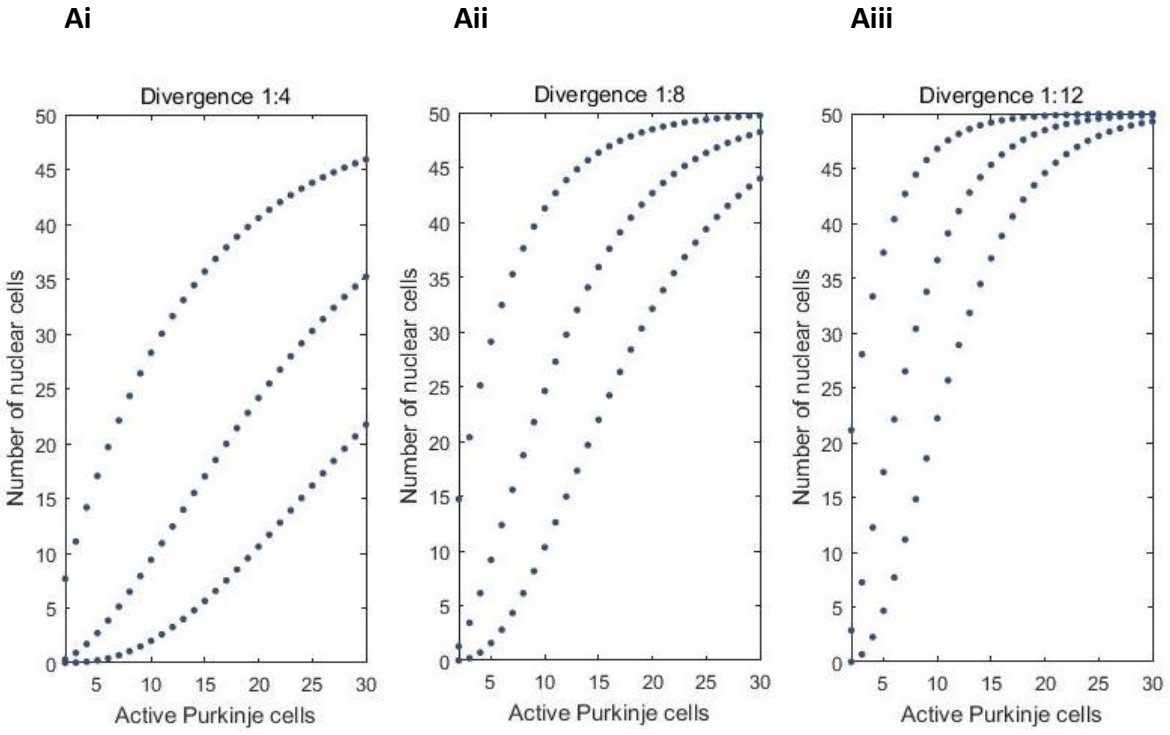
1) The inhibition of a nuclear group reaches functional saturation efficiently, that is, inhibition by Purkinje cells of nuclear cells is broadly equally distributed even at low numbers of active Purkinje cells.
2) A modest number of spontaneously active Purkinje cells is enough to strongly inhibit the whole of a nuclear group. The Purkinje cells do not have to be clustered together but can be any group of the necessary minimum number. However, it is likely that a (mediolateral) row of Purkinje cells that spans a microzone behaves as a group because they sample the same parallel fibre activity.
3) The coordinated suppression of Purkinje cell firing across the whole population of a microzone is necessary in order to disinhibit a nuclear cell group, because less than full coordination means that each nuclear cell is at high risk of receiving strong inhibition.

4) Most of the inhibition of a nuclear cell is functionally supernumerary for most of the time. This is necessary in order that it can be any handful of Purkinje cells which blocks output of the circuit if Purkinje cell rates are not co-modulated across the whole group.

3.4 Adaptations that abet a strong effect by a handful of Purkinje cells

Efficiency of the inhibition of a nuclear group by a low number of Purkinje cells may be increased by nuclear interneurons (see Figure 5), and by the distribution of inhibitory and excitatory contact on nuclear projection neurons.

FIGURE 5



Bi - iv

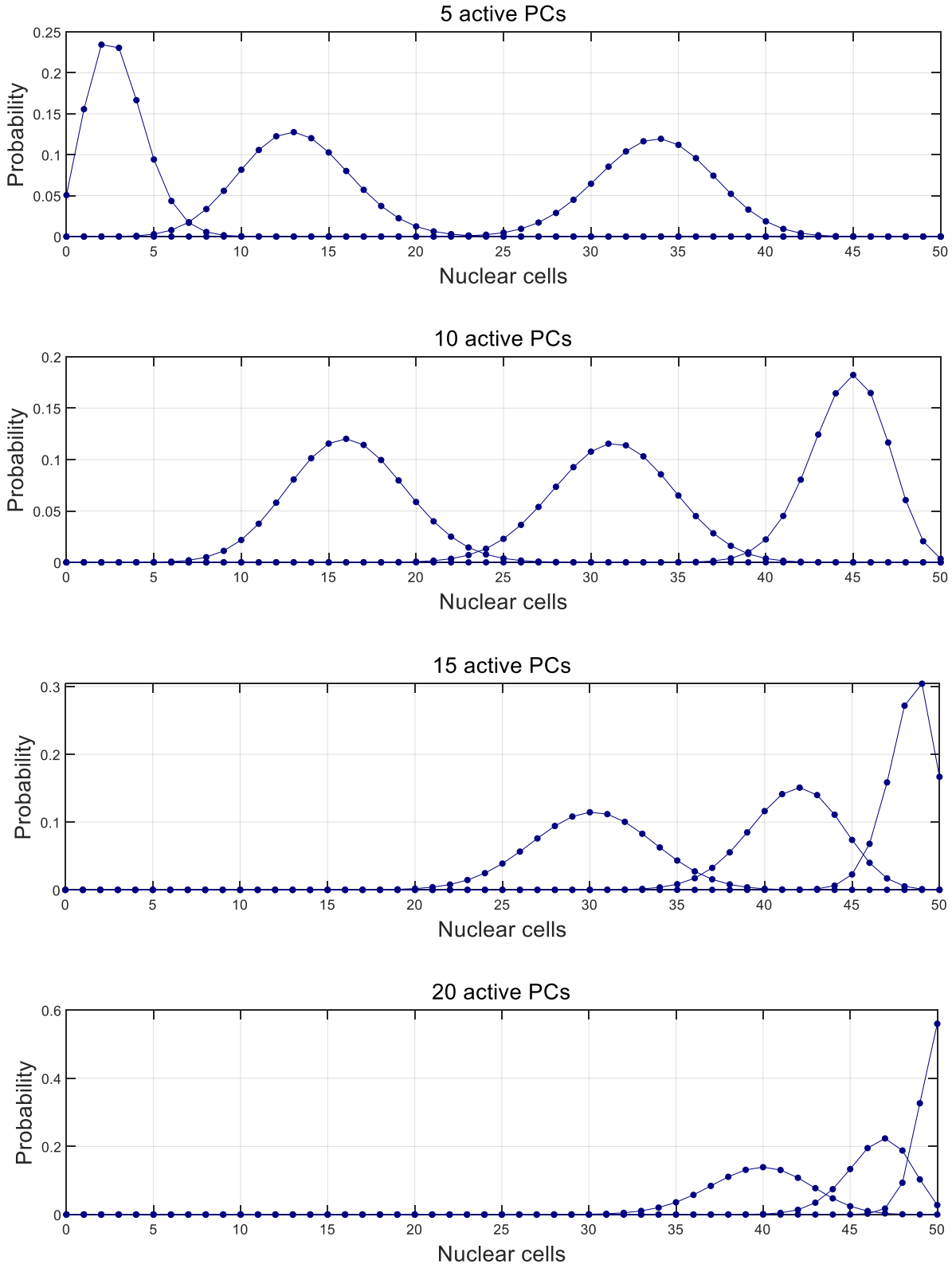


FIGURE 5

The effect of increasing divergence of Purkinje cells onto nuclear cells. Deep nuclei contain excitatory (presumed glutamatergic) interneurons that fire spontaneously and which are normally inhibited by Purkinje cells (Uusisaari and Knopfel 2008) but released from inhibition in the conditioned response. This may effectively increase divergence of Purkinje cells onto nuclear cells because a Purkinje cell has an effect on more than only the nuclear cells it contacts directly. **Ai – iii:** The expected number of nuclear cells, out of a group of 50, which receive contact from at least 1 Purkinje cell (top curve), at least 2 (middle curve) and at least 3 (bottom curve), out of a number x which are active. A steeper curve indicates that fewer active Purkinje cells are needed to inhibit a nuclear group. With divergence of 1:4 (first panel) around 20 active Purkinje cells are needed for an expected 40 out of 50 nuclear cells to receive contact from at least 1 and only around half a nuclear group receives contact from at least 2, whereas with divergence of 1:12 (third panel) and 20 active Purkinje cells, all nuclear cells receive contact from at least one active Purkinje cell and almost all from at least 2. **Bi – iv:** The probability that a number of nuclear cells, x, receives contact from 1 or more (right peak), 2 or more (centre), and 3 or more (left) out of 5, 10, 15 or 20 (Bi - iv respectively) active Purkinje cells, with divergence of 1:10. Compare Ai and Aii to Bii and Biv respectively.

Contact by Purkinje cells on nuclear cells 'is characterised by preferential targeting of cell somata rather than dendrites' (Uusisaari and De Schutter 2011, p.3443), while the majority (75%) of excitatory inputs are distal (de Zeeuw and Berrebi 1996). Purkinje cell synapses are

therefore positioned to block an effect of excitatory input to nuclear cells. This powerful inhibitory veto is strengthened by inhibition of excitatory interneurons (Uusisaari and Knopfel 2008) because it weakens or silences intrinsic firing of interneurons, so that the tonic effect of intrinsic Purkinje cell activity is both to directly inhibit nuclear cells and to block tonic excitation. The significance of a strong inhibitory bottleneck is that it increases the potency of inhibition by a Purkinje cell of its nuclear targets, so that fewer are need to be potent.

3.5 The form of pattern memory

Our previous estimate (in section 2) of the density of parallel fibre activity (number active per unit area in the sagittal plane) would mean that a Purkinje cell receives contact from several hundreds at a time (n = ~900, assuming a Purkinje cell span of 300 µm). Because activity is evenly and randomly distributed at the scale of input to a Purkinje cell (section 2), training teaches an evenly and randomly distributed pattern of synapses. Pattern density is regulated, uniform and ubiquitous. This consistency is undisturbed by folding of the cerebellar cortex because the size and shape of the Purkinje cell arbour and the thickness of the granular layer are adjusted so that the number of parallel fibres that contact a Purkinje cell is unaffected (to reason from the unaffected number that intersect the Purkinje cell dendritic field) (Eccles, Ito et al. 1967).

The narrow range of the regulated density of active parallel fibres causes a predictable result of overlap of stored patterns. The amount of overlap is predicted by the number stored, and the same for all patterns. The proportion of synapses which also belong to 1

other pattern, and to 2, and 3 and so on is also predictable (Figure 6) and also the same for all stored patterns. As more patterns are stored the ratio of total trained to total untrained (with paired input) synapses[8] shifts. Moreover the split is the same for all Purkinje cells trained to the same number of patterns and therefore all Purkinje cells in the same microzone (and all microzones in the same circuit), because climbing fibre signals are received in a volley across the whole population, and are always paired.

An estimated 80-85% of parallel fibre-Purkinje cell synapses are strongly long-term depressed, to the extent that there is 'no detectable somatic response' to granule cell stimulation (Isope and Barbour 2002 p.9676). This is consistent with a high estimate of 'electrically silent' synapses made by parallel fibres activated by cutaneous stimulation (Ekerot and Jörntell 2001). The Isope and Barbour detection threshold could in theory leave room for compound responses (Boris Barbour, private correspondence dated 7 December 2018) i.e. it does not conclusively rule this out. However, if so, this is likely to be weak and (not less important) it would not be pattern-specific but generic, that is, the same for all trained patterns, because random overlap of uniformly dense patterns means they all overlap in the same predictable proportions.

Because trained synapses are functionally inert there is no effect of unrelated memories on the response to other patterns. Trained synapses are (functionally) the 'same' weight whether they are part of one pattern or several. The function of parallel fibre-Purkinje cell synaptic depression is to render transmission functionally negligible (in this view). Graded weights are not necessary for this task – in fact graded transmission would impair function.

[8] 'Untrained' is defined by default as not trained with paired input.

FIGURE 6

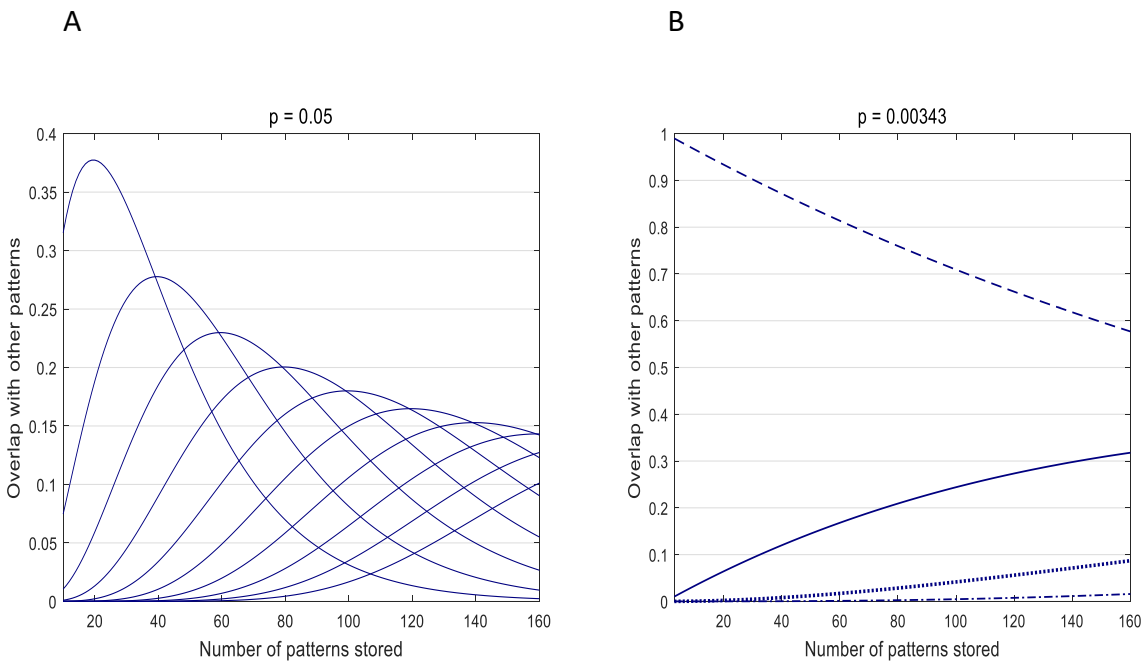


FIGURE 6

The changing relative proportions of a pattern which overlap with 1 other pattern, and with 2, and 3 and so on as more patterns are stored, is given by

$$y = \frac{x!}{k!\,(x-k)!} * p^k * (1-p)^{x-k}$$

where y is the proportion of each pattern (the same for all of them) that overlaps with k other patterns, $x + 1$ is the total number of patterns stored, and p is the the fraction of parallel fibres that are active (in reality a constant, because the level of parallel fibre activity is regulated). In (A) p = 5% of the total number of parallel fibres. The left peak is the proportion of synapses which also belong to one other pattern, the next one to the right is the proportion that also belong to 2 other patterns, the next is the proportion that belong to 3 other patterns,

and so on, up to 10 other patterns. The x axis starts at 10 for convenience in running the calculations. (B) assumes 1,200 out of 350,000 parallel fibres are active, the estimated (in section 2) number at internally regulated levels of activity, so p = 0.00343. The pattern is greatly stretched as if pulled from the right. (The purpose of A, with a higher, non-physiological number active, is to show the relationship of the relative proportions of overlap not evident only from the range in B.) The dashed line is the fraction of a pattern that does not overlap with others. The solid line is the proportion of synapses which also belong to one other pattern. The dotted line is the proportion that also belong to 2 other patterns, and the dots and dashes are the proportion that overlap with 3 other patterns. Overlap with 4 and higher numbers is either extremely low or absent in the displayed range.

3.6 Discussion

A nuclear group is held under tonic inhibition unless there is a coordinated reduction in the firing rate of the entire population of Purkinje cells that innervate it. Put another way, a microzone must receive a match along its full length to modulate nuclear firing. A pattern that does not meet this condition does not lift the nuclear inhibitory blockade.

A failure to meet that threshold does not trigger a graded response but no response at all, regardless of how good or bad the match is overall, or which parts are good and which parts are bad. There is no proportionate (or any other) response to a partial match. This is important because the response to a partial match would be arbitrary. In this sense, counter to traditional expectation, pattern matching controls the response to input that a circuit does *not* recognise, and not patterns it does. That response is the ongoing and functionally

unweakened inhibition of nuclear cells that carry the output of the circuit. This is the default state.

The supernumerary inhibition by Purkinje cells of nuclear cells and strong individual contact, necessary to permit a small handful of (any) Purkinje cells to block output, creates a need (and so can account) for specialist adaptations of Purkinje cell-nuclear cell contact which permit nuclear cells to maintain a baseline rate of firing even under heavy inhibitory bombardment, which remains very sensitive to dynamics of the simple spike rate and can be adjusted in either direction (Telgkamp and Raman 2002, Pedroarena and Schwarz 2003, Telgkamp, Padgett et al. 2004, Turecek, Jackman et al. 2017).

We stress that we do not propose an all-or-nothing response of individual Purkinje cells or suggest that Purkinje cells must be silenced to modulate firing of nuclear cells. Merely if any (more than a low minimum) are not in step with the modulation of the rest, they are sufficient to block an effect on (the whole of) a nuclear group. If only a small part of a pattern is a mismatch (and regardless of which part), it is sufficient to block output.

Contrary to traditional learning models, the effect of training is to *eliminate* variation of the post-synaptic effect of known patterns, because input in a known pattern is received exclusively at severely depressed synapses, and this is true for all known patterns. The pattern that cues output does not also code it.

All unknown patterns are equal in having no effect through 85% of synapses and otherwise addressing a random sample of operational synapses. That is, pattern memory is binary – the response indicates only if a pattern is known or not, and does not discriminate individually between known patterns, or between unknown patterns. There is no need for

graded synaptic weights, because learned patterns do not control the output firing rate (we claim). The role of pattern matching in the conditioned response is instead permissive – it selects which circuits have output and when, but does not code output.

'Binary' here means something very different from the perceptron meaning. There, a trained Purkinje cell is limited to two response, say 1 or 0 (Brunel, Hakim et al. 2004). Both responses are learned i.e. the result of iterative adjustment of synaptic weights. All patterns have an effect on the postsynaptic Purkinje cell which either falls above or below a threshold. The size of both classes is limited by storage capacity. Learning uses an algorithm to adjust synaptic weights. The correct response following training requires a repeat of a learned permutation of input rates. Our meaning is *none* of these things. The response does not control Purkinje cell rates (output rates are controlled separately), there is no algorithm, pair-trained synaptic weights are invariant (all silent), and unknown patterns do not need learning, so the size of the class of unknown patterns is unlimited. Also, the response (mediated by the pattern matching function) of the postsynaptic cell is triggered by the binary pattern of inputs (some on, some off) only. The permutation of rates does not affect the response.

4. STELLATE CELL NETWORKS

4.1 Background

As noted, learning models propose that the cerebellum implements a supervised learning algorithm which uses iterative adjustment of parallel fibre synaptic weights. Purkinje cells in this way acquire a learned response driven by input in a remembered pattern, displacing the

naïve response to input rates (Albus 1971, Fujita 1982, Brunel, Hakim et al. 2004, Dean, Porrill et al. 2010).

Following training with a conditioning protocol (or direct stimulation which mimics a conditioning protocol) Purkinje cells respond to the conditioned stimulus with a reduction or pause in firing (Jirenhed, Bengtsson et al. 2007, Rasmussen, Jirenhed et al. 2008). Purkinje cells are interleaved with and receive contact from inhibitory interneurons (Palay and Chan-Palay 1974, Paula-Barbosa, Tavares et al. 1983, Sultan and Bower 1998) which in turn receive contact from parallel fibres. It is thought the conditioned response of Purkinje cells is abetted by feed-forward inhibition of Purkinje cells by molecular layer interneurons (MLIs) driven by parallel fibres, following training.

It has been reported that 'MLIs encode locomotion-dependent changes in GC [granule cell] input with linear changes in firing rate' (Jelitai, Puggioni et al. 2016 p.6) (counting spikes in 200 ms bins to calculate frequency). (The method of measurement in this study did not permit a distinction between stellate cells and basket cells.) But what restricts feed-forward inhibition to the learned response, and what chooses the source of input signals to control it?

4.2 Bad memory

Pair training potentiates the parallel fibre-stellate cell synapse (Jörntell and Ekerot 2003, Rancillac and Crépel 2004, Smith and Otis 2005, Jörntell and Ekerot 2011). Potentiation is reversed by unpaired, parallel-fibre-only input. Assuming the high proportion of silent parallel fibre-Purkinje cell synapses (Isope and Barbour 2002) is because they have received

endogenous training with paired input, we would expect a similar proportion of parallel fibre-stellate cell synapses to be potentiated under tuition of the same climbing fibre signals (and we can confirm mathematically that anatomical differences don't affect the proportion).

Because trained synapses are potentiated, the high proportion should mean that – on the face of it – feed-forward inhibition is indiscriminate, that is, driven by known and unknown patterns alike. Put another way, stellate cells are pattern blind. We are able to confirm this by calculating the probability distribution for input to operational synapses with a known and with an unknown pattern of parallel fibre activity (see deeper level section). Learned blindness of interneurons is a problem the cerebellum has therefore been obliged to solve. In fact, the position is different for stellate cells at superficial and deeper levels of the molecular layer because input they receive from parallel fibres is topographically stratified. The main focus of the following discussion is on deeper-stratum stellate cells but superficial level is discussed first.

4.3 Superficial molecular layer interneurons

MLI morphology varies continuously (Palay and Chan-Palay 1974, Paula-Barbosa, Tavares et al. 1983) with the depth of the cell body in the molecular layer. Among other things, deeper cells have longer main axons, more and longer collaterals, and drop more descending collaterals that terminate in the Purkinje cell layer (Paula-Barbosa, Tavares et al. 1983, Sultan and Bower 1998). MLIs at deeper level are networked, so that they receive tonic inhibition – MLIs fire intrinsically (Häusser and Clark 1997, Ruigrok, Hensbroek et al. 2011).

By contrast, connections between superficial stellate cells, whose axons wander around without leaving their dendritic field (Palay and Chan-Palay 1974), are less frequent or absent, so that superficial cells receive weak or no contact from each other.

Pattern blindness raises the question, what is learning for at the parallel fibre-stellate cell synapse? Evidence comes from the C3 region of the cerebellar cortex, which is involved in forelimb movement. Stellate cells in this region respond (by firing) exclusively to stimulation of an associated discrete region of the body surface, or receptive field, (Jörntell and Ekerot 2002, Ekerot and Jörntell 2003, Jörntell and Ekerot 2003, Jörntell and Ekerot 2011), and not to stimulation of other fields. It is thought that this stimulation drives paired input which trains potentiation of the parallel fibre-stellate cell synapse. Input to stellate cells evoked by stimulation of other fields is to untrained synapses. Transmission at an 'untrained' synapse is not weak but nil (Jörntell and Ekerot 2003, even at several 100 Hz: Henrik Jorntell, private correspondence dated 31 March 2017).

The interneurons used in these studies were activated by cutaneous stimulation, effectively making it a selection requirement that they were outer level, although they were not expressly selected for their depth. This is because there is, in the C3 region of the cerebellar cortex in adult cats, 'a specific depth distribution of granule cells depending on the type of input they received' (Jörntell and Ekerot 2006, p.11795, Quy, Fujita et al. 2011). Input triggered by cutaneous stimulation is received by superficial granule cells. Superficial granule cells prevalently bifurcate in the outer level of the molecular layer (Palay and Chan-Palay 1974, Zhang and Linden 2012), where they contact outer level stellate cells.

The wiring of input to C3 circuits effectively extends modular cerebellar circuit wiring to the body surface. Stimulation evokes mossy fibre and climbing fibre signals that travel by

different pathways to converge on the same vertical slice of the cerebellar cortex (Garwicz, Ekerot et al. 1998, Garwicz, Jörntell et al. 1998). Climbing fibres terminate on Purkinje cells while mossy fibres terminate on granule cells in the subjacent granular layer. On the whole, there is 'a close correspondence [of terminal fields] between inputs conveyed by climbing fibres to the molecular layer and those conveyed by mossy fibres to the underlying granular layer' (Apps and Hawkes 2009 p.677). Thus learning is induced with a single peripheral stimulus which evokes paired input – not just the same type of stimulus but the same event. This 'is a feature that seems to be observed across species and other parts of the cerebellar cortex (for example, crus II in the rat)' (Apps and Garwicz 2005, p.305).

Thus at outer level paired patterns of parallel fibre activity are not perfectly random, being made up of the subset of parallel fibres that are the axons of local granule cells in the underlying strip of the granular layer. As a result, stellate cells at this level are very good at discriminating between activity of local granule cells and signals originating elsewhere, in other circuits, triggered by stimulation of other fields.

4.4 Deeper molecular layer interneurons

At deeper level, where input is not sourced in the same way, pattern separation in this way is not possible. Deeper-lying stellate cells receive contact from the parallel fibres of deeper-lying granule cells, which in the C3 region receive signals triggered by limb movements (Jörntell and Ekerot 2006), presumably by activation of receptors and nerve endings in muscles and tendons – proprioceptors. A pattern of parallel fibre activity at this level has no special relationship with the source of climbing fibre signals, does not have the same trigger

as climbing fibre signals, and terminates in a disorganised way, mixing signals from different sources.

Because we (claim to) know that the density of active parallel fibres is in a fixed narrow range (derived in section 2), we can calculate the probability distribution for input to a stellate cell – the probability that given any particular pattern of parallel fibre activity a stellate cell receives contact from 0 active cells, from 1, from 2 and so on.

A Purkinje cell receives an estimated 175,000 parallel fibre synapses (Napper and Harvey 1988), one each from around the same number of granule cells (Harvey and Napper 1991), while a roughly equal number of parallel fibres pass through the same space without making contact. A deeper-level stellate cell dendritic field (in adult rats) extends around 110-130 µm in both horizontal and vertical directions (Palay and Chan-Palay 1974, pp.217-221). Assuming a Purkinje cell and stellate cell arbour size of 300 x 300 µm and 120 x 120 µm respectively, some 350,000/~6 or ~60,000 parallel fibres pass through a stellate cell dendritic field, assuming parallel fibres are evenly distributed. If a stellate cell receives contact from 1,000 parallel fibres, around 1 in every 60 parallel fibres which pass through a field makes contact. 1 in ~200 are active (derived from the estimated proportion that are active), so around 300. There is accordingly a probability of 1/60 = ~0.01667 that an active cell makes contact. The probability of contact by k (out of 300) active cells is therefore given by

$$\frac{300!}{k!\,(300-k)!} * 0.01667^{k} * (1-0.01667)^{300-k} \tag{2}$$

This gives the Table A probabilities for the range k = 0 to 10 where k is the number of active fibres that contact a stellate cell and p is the probability of that number given a random pattern of parallel fibre activity. There is accordingly a high probability (p = 0.9074) that the number lies between 2 and 8. This agrees closely with the reported number, indicated by 'two to eight substantial EPSPs [excitatory postsynaptic potentials]' (Jörntell and Ekerot 2003 p.9628).

An unknown pattern of input to a stellate cell may be to all trained synapses, all untrained, or a mixture (with all untrained the least likely). The probability of n inputs at trained synapses is the probability of n inputs to any synapse reduced by the probability that at least one input is to an untrained synapse (so: 1 – *P*(none)) but increased by the sum of the probabilities that a higher total number of inputs, a, is reduced by a number, b, to untrained synapses, such that $a - b = n$. For example, the odds that there are two inputs to trained synapses is increased by the product of the probabilities that there are three inputs in total, and any one is to an untrained synapse. So the probability of n inputs at trained synapses with a random pattern is:

$$P(n) - P(n)(1 - 0.85^n) + \sum_{y=1}^{z} P(n+y)\left(\frac{(n+y)!}{y!\,n!} * 0.15^y * (1-0.15)^n\right) \quad (3)$$

where $P(n)$ is the probability of n inputs (to any synapse, derived in Table A) and $n + z$ is the maximum number of inputs with more than insignificant odds (so for example around 10 in Table B because higher numbers have a very low probability).

Table A

k	0	1	2	3	4	5	6	7	8	9	10
p	.006	.033	.083	.14	.176	.177	.147	.105	.065	.036	.018

Table B

n	0	1	2	3	4	5	6	7	8	9	10
p	.0138	.0596	.128	.1828	.1951	.166	.1172	.0703	.0358	.0144	.0035

FIGURE 7

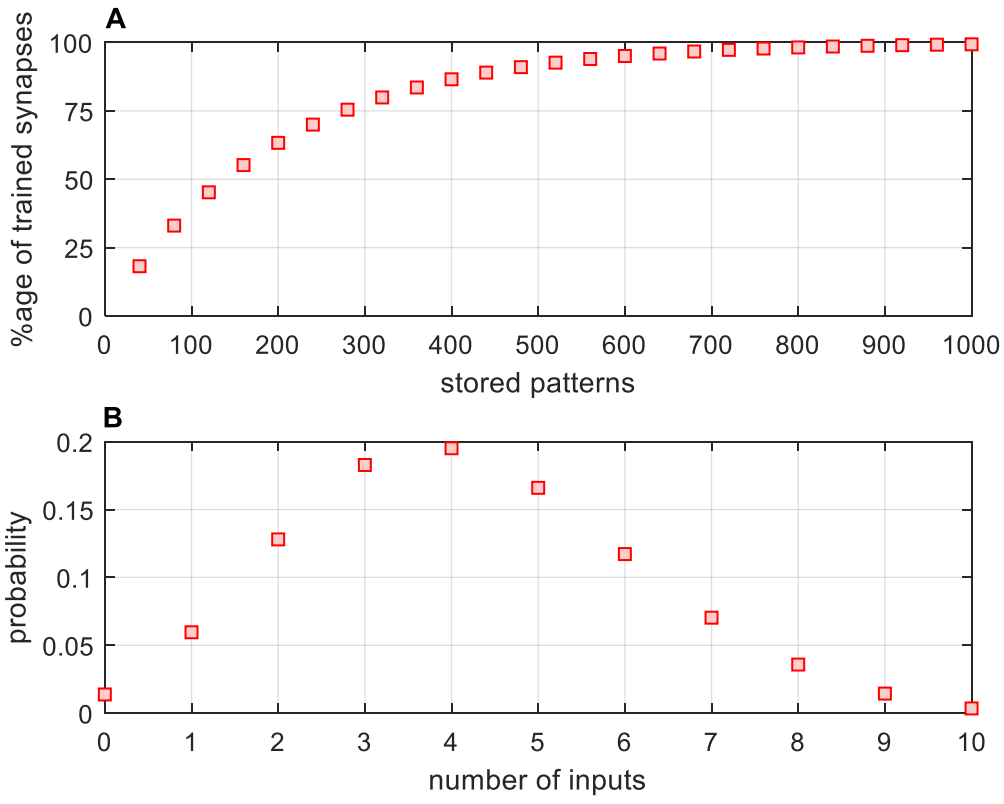


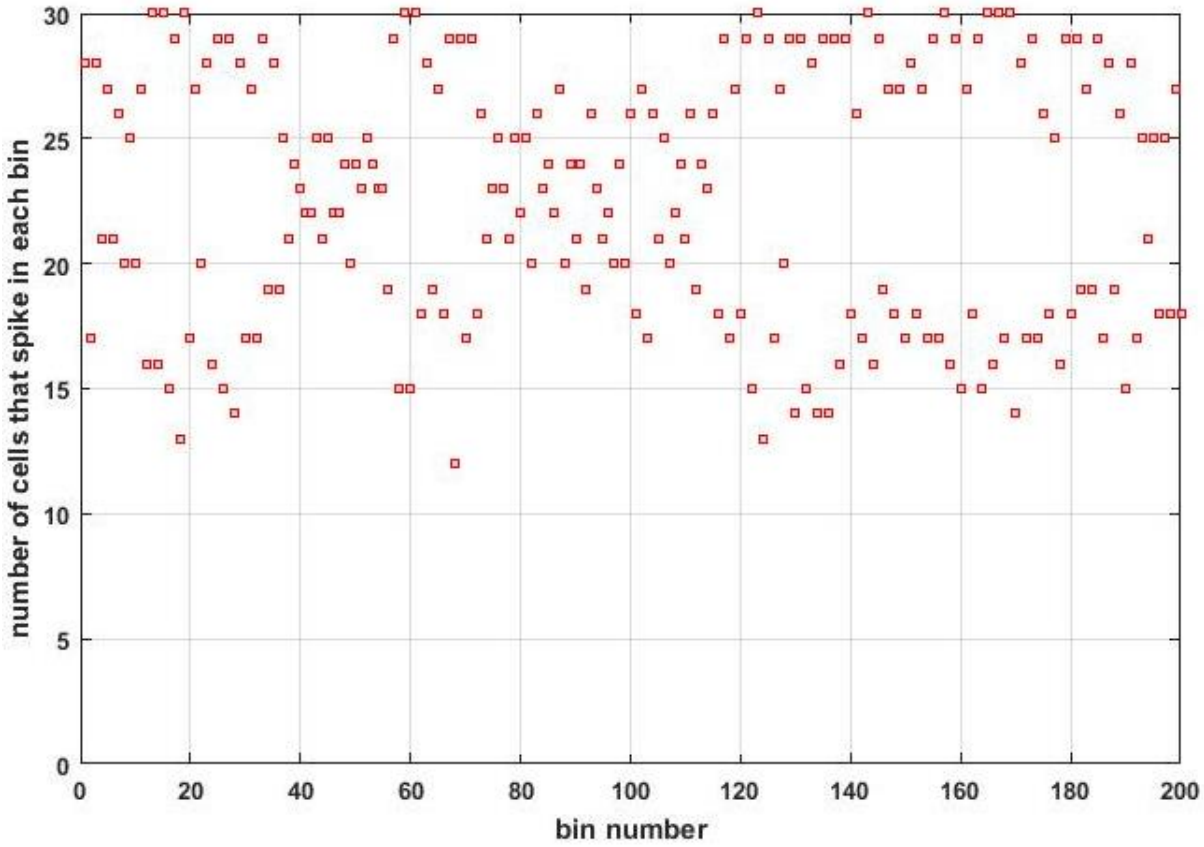
FIGURE 7

A: Estimate of the storage capacity of a stellate cell assuming that a total of 1,000 parallel fibres make contact. Data are thinned for clarity. The number that make contact per pattern is the weighted average of the Table A probability distribution. If 85% of synapses are trained – predicted by the reported proportion for the parallel fibre-Purkinje cell synapse – a stellate cells stores around 370 patterns. **B:** The probability distribution for the number of active parallel fibres that make contact on a stellate cell at trained (therefore operational) synapses, assuming 85% of synapses are trained, with an unknown pattern of input – Table B in graph form. The number is in the range 0 – 10 (a higher number is very unlikely). Each number of inputs has an associated probability.

This argues: 1) The number of inputs received at operational synapses depends much more on chance than on learning. 2) Synaptic weights are not pattern specific, because stored patterns overlap very substantially (and the effect on transmission is not even generic, because pattern size is small). 3) Sub-superficial MLIs are in any case pattern blind, so unable to have an effect that requires them to store and remember patterns. 4) The fixed, regulated level of parallel fibre input to sub-superficial MLIs is insufficient to drive a learned response, predicting a different or additional source. This is consistent with evidence that a number of mechanisms severely weaken an effect of isolated excitatory input to stellate cells on dendritic signalling (Häusser and Clark 1997, Abrahamsson, Cathala et al. 2012, Tran-Van-Minh, Abrahamsson et al. 2016), arguing that a somatic effect is blocked at the low numbers a stellate cell receives (Table B) at regulated levels of parallel fibre activity. It is

also supported by a model of stellate cell network activity that replicates stellate cell spiking under inhibitory input only (Figure 8).

FIGURE 8



17 Hz (top), 24.2 Hz (bottom)

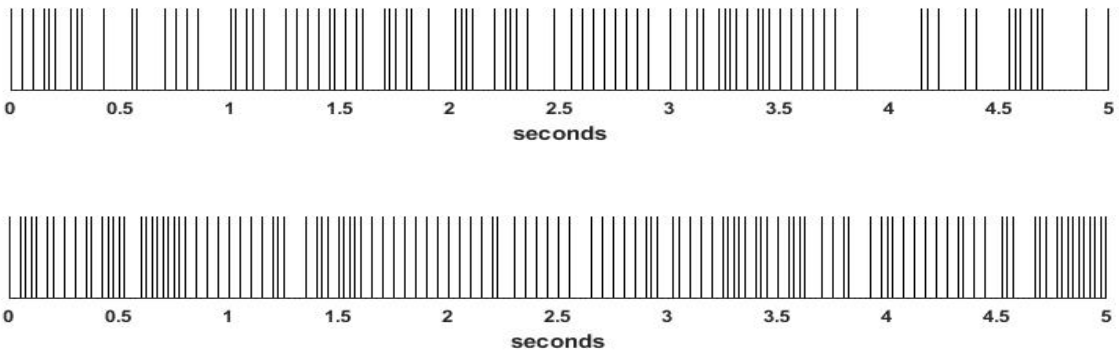


FIGURE 8

Simulation of a stellate cell network. The graph shows the number of deeper-stratum stellate cells (out of 30 which contact a simulated Purkinje cell) that are active in each of two hundred 25 ms bins (so for 5 seconds). Cells were modelled individually, each receiving contact from a number of their neighbours determined by probability that varies with the depth of the cell body in the molecular layer, to reflect depth-dependent anatomy. The un-networked, intrinsic firing rate of a stellate cell is around 40 Hz (hence 25 ms bins). A cell either spikes in a bin or is silent, depending on the number of cells that make contact on it that spike in the bin before. This reflects (in discrete bins for convenience) the unpredictable delay reported to be caused by the erratically-variable time and number of recent input spikes. The firing pattern of individual cells generated by this model (examples are shown under the graph) is a good match with the reported firing pattern (compare Häusser and Clark 1997 figs 1 and 3). The model replicates stellate cell spiking under inhibitory input only, suggesting that an effect of excitatory input is normally blocked at low numbers of inputs, so at homeostatically regulated density of parallel fibre activity.

4.5 Control of stellate cells in the conditioned response

What, then, elevates the stellate cell rate during the conditioned response? A candidate to contribute to this function is nucleocortical feedback (Houck and Person 2015, Gao, Proietti-Onori et al. 2016). Feedback is carried by collaterals of nuclear cell axons that project back to the granular layer where they terminate as mossy fibres, forming a closed circuit (Gao,

Proietti-Onori et al. 2016). Nuclear cells fire intrinsically even under heavy tonic inhibition by Purkinje cells, partly as a result of short-term depression at stable afferent rates (Telgkamp and Raman 2002, Pedroarena and Schwarz 2003, Telgkamp, Padgett et al. 2004). Nuclear cells remain very sensitive, however, to the dynamics of the simple spike rate (Telgkamp and Raman 2002), such that their firing is transiently modulated by a *change* of the rate, which can be in either direction (Pedroarena and Schwarz 2003, Baumel, Jacobson et al. 2009). So that a drop in the Purkinje cell rate, such as the learned pause seen following training with a conditioning protocol, causes them to fire at a transiently elevated rate.

Closed-circuit feedback thus provides a targeted and timely injection of extra mossy fibre input to a microzone in the conditioned response, following training. This does not create a positive feedback loop (we suggest, contrary to the reported interpretation) because the dominant effect is to cause an uplift in the number of local granule cells that meet firing threshold – as opposed to an increase in granule cell firing rates. The reason is that many more cells that received only 2 inputs before – and so did not fire – now receive an extra input and thus reach the input threshold (the minimum necessary to fire, illustrated and discussed in Figure 9), than the number that received three and now receive a supernumerary fourth input. The sum of excitatory inputs received by networked stellate cells during the conditioned response is thus temporarily elevated. Elevation is sufficient to overcome the tonic somatic inhibitory blockade. As a result, deeper-stratum stellate cells become phasically sensitive to excitatory input, and inhibit Purkinje cells at elevated, supra-tonic rates.

Mossy fibre input to the cerebellum is vertically topographically organised (reported for the C3 region: Jörntell and Ekerot 2006, Quy, Fujita et al. 2011). Input triggered by cutaneous

stimulation is received by superficial granule cells, and deeper-lying granule cells receive signals triggered by limb movements (Jörntell and Ekerot 2006), suggesting it is driven by activation of internal receptors in muscles and joints: proprioception. Topography is preserved in the molecular layer, with 'granule cells in the inner granule cell layer giving rise to PFs [parallel fibres] in the inner molecular layer and granule cells in the outer granule cell layer giving rise to PFs in the outer molecular layer' (Zhang and Linden 2012, p.122), as a 'prevalent rule' (Palay and Chan-Palay 1974, p.66).

Nucleocortical feedback terminates mainly in the superficial granular layer (Gao, Proietti-Onori et al. 2016), therefore prevalently exciting granule cells whose axon divides in the superficial molecular level. Ascending axon contact on stellate cells is unconfirmed but would mean contact was confined below superficial level to the same closed circuit. At superficial level, discrimination of stellate cells between parallel fibre signals arising in that circuit, in the subjacent granular layer, and in other circuits, is excellent, as discussed (Garwicz, Jörntell et al. 1998, Jörntell and Ekerot 2002, Jörntell and Ekerot 2003), such that feedback would excite stellate cells in the same circuit and have no effect at all in other circuits.

Because stellate cells at deeper level have poor pattern memory control in the conditioned response, when they are briefly responsive, can be by parallel fibre signals which do not need to be in a known configuration (memory is unnecessary either of the binary pattern of activity or the permutation of rates). Input to a microzone in the control window is distinct from the pattern which is used for pattern matching. Signals used in pattern matching are spread across a longer time window (the learning window for synaptic plasticity), and learned by repetition. Signals that control simple spike rates acting through stellate cells –

and therefore nuclear rates – are in a time-adapted (because the intrinsic Purkinje cell response is time-adapted: Johansson, Jirenhed et al. 2014)[9] and narrower window, and are not constrained to a repeating configuration. Possibly the windows overlap and it is likely that both sets of signals are from substantially the same set of (movement-related) sources. However, signals activity in the control window represents more up-to-date information, transmitted by wide-diameter, myelinated, fast-transmitting neurons (Loeb and Mileusnic 2015). It is not rigidly confined to the same set of parallel fibres every time, allowing control of Purkinje cell rates by conditions on *this* occasion not the last time or previous occasions, when the learned pattern was memorised.[10]

4.6 Linear translation of granule cell rates into Purkinje cell rates

MLIs reflect 'granule cell input with linear changes in firing rate' (Jelitai, Puggioni et al. 2016 p.6). The effect of inhibition by MLIs of Purkinje cells is rate dependent and linear, such 'that locomotion-dependent modulation of the balance between excitation and inhibition [of Purkinje cell dendrites] generates depolarising or hyperpolarising dendritic V_m [dendritic membrane voltage] changes that linearly transform into bidirectional modulation of PC SSp [Purkinje cell simple spike] output' (Jelitai, Puggioni et al. 2016 p.9).

[9] Purkinje cells acquire an intrinsic plastic response to known input, even with MLIs pharmaceutically blocked Johansson, F., D. A. Jirenhed, A. Rasmussen, R. Zucca and G. Hesslow (2014). "Memory trace and timing mechanism localized to cerebellar Purkinje cells." Proc Natl Acad Sci U S A **111**(41): 14930-14934.. The response is an adaptively-timed, mGluR7-dependent Johansson, F., H. A. Carlsson, A. Rasmussen, C. H. Yeo and G. Hesslow (2015). "Activation of a Temporal Memory in Purkinje Cells by the mGluR7 Receptor." Cell Rep **13**(9): 1741-1746. transient fall in their firing rate.

[10] Also, the spatial pattern of parallel fibre activity is sensitive to mossy fibre rates, so the timing of the window opened by pattern matching is rate sensitive.

FIGURE 9

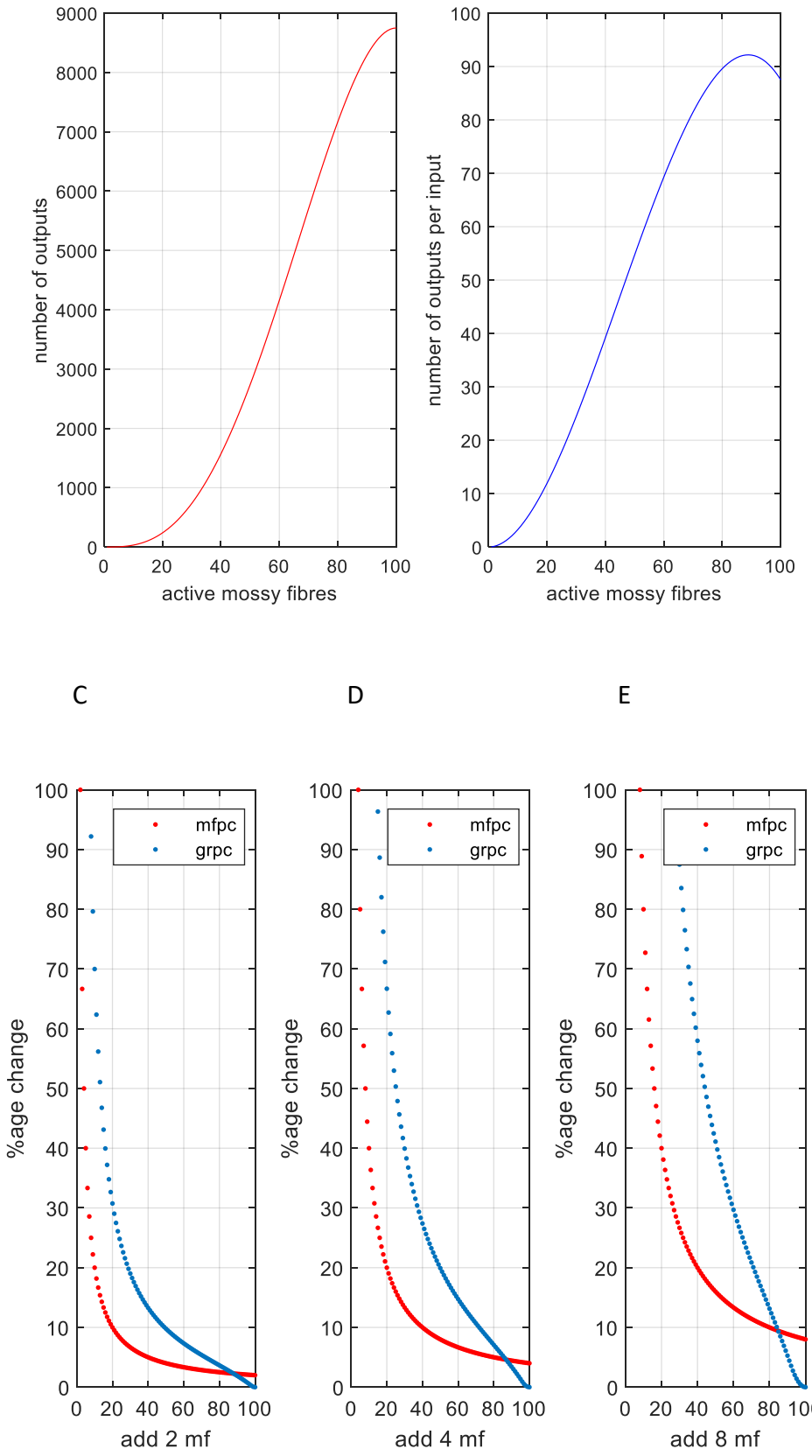


FIGURE 9

Closed circuit feedback in the conditioned response provides extra mossy fibre input to trained circuits. As discussed (in section 2) the number of granule cells that meet the input threshold in a region with the dimensions of a mossy fibre terminal cluster field can be calculated from the number of active mossy fibres. **A:** The number of inputs (active mossy fibres) to a cluster field plotted against the number of outputs (granule cells that meet the input threshold). To fire, a granule cell needs mossy fibre input to either 3 or all 4 dendrites. Modulation by Golgi cells is disregarded. **B:** Nucleocortical feedback adds to the number of mossy fibre inputs received in the superficial granular layer. The number of outputs per input increases as the number of inputs rises. That is, a single mossy fibre (either more or less) has a larger effect if it is one of a larger number of inputs. This relationship breaks down – the curve peaks – at the high end of the range because an extra input is more likely to be to the fourth dendrite of a granule cell which already fires. **C-E:** The percentage change in the number of active granule cells (blue data) caused by the addition of an extra 2, 4 and 8 active mossy fibres (C, D and E respectively) to a number that are already active, x. The red data show the percentage change in the number of mossy fibres. For most of the range a change to the mossy fibre number causes a larger percentage change to the granule cell number. Blue values are lower than red values at very high values of x because at those levels, many granule cells receive 4 inputs, and so still receive enough input to fire if one is removed.

Thus, in the control window, when the balance of excitation and inhibition tips strongly towards inhibition (because direct input to Purkinje cells is received at inert synapses, and an effect of excitatory input to MLIs is unblocked), there is linear translation of granule cell

to simple spike rates. This is in conflict with the perceptron and adaptive filter models where output is controlled by learning and not by input rates.

4.7 Selection of movement-related input signals to control output rates

The longer and more highly collateralised axons of deeper-stratum stellate cells mean that the concentration of synaptic contact by stellate cells on Purkinje cells increases with depth. So assuming the strength of the inhibition of Purkinje cells increases with the concentration of contact, it increases with depth. Therefore, when the sub-superficial block of feed-forward inhibition is suspended in the conditioned response, stellate cells at that level dominate control of the simple spike rate. This has the result that there is depth-dependent selection of parallel fibre signals that control the simple spike rate in the conditioned response, so that control is by mossy fibres received by those granule cells, in the corresponding stratum of the granular layer. In the C3 zone those signals are movement related (Jörntell and Ekerot 2006).

Thus performance-sensitive, phase-locked mossy fibre signals may dominate control of feed-forward inhibition with learned timing provided by training with the endogenous equivalent of a conditioning protocol driven by a movement cycle. Movement-sensitivity may not be limited to signals generated internally. Some skin receptors may provide ‘cutaneous proprioception’ during movement (Jörntell and Ekerot 2006), meaning: distortion of the skin during movement, especially cyclic movement, may cause some mechanoreceptors to generate a time-varying signal.

5. CONTROL OF OUTPUT RATES

5.1 Background

Smooth and precise control of movement at variable speeds, thought to be a function of the cerebellum, is dependent on incremental modulation of motor commands on a short timescale (milliseconds). The large majority of cerebellar circuits (i.e. non-floccular circuits) include a discrete group of deep nuclear cells which contain excitatory projection neurons which carry the output of the circuit.[11] Nuclear cells receive heavy inhibitory input from Purkinje cells, which alone carry the output of the cerebellar cortex. Both Purkinje cells (Häusser and Clark 1997, Raman and Bean 1999, Zhou, Lin et al. 2014) and nuclear cells (Person and Raman 2012b, Mercer, Palarz et al. 2016) fire intrinsically. Following training with a conditioning protocol or afferent stimulation that mimics a conditioning protocol (Jirenhed, Bengtsson et al. 2007, Rasmussen, Jirenhed et al. 2008), Purkinje cells respond with a transient reduction and sometimes a full pause in firing.

The conditioned fall in Purkinje cell rates may in some circuits gate mossy fibre collateral input to nuclear cells. However, not all mossy fibres send collateral projections to deep nuclei. Only 1 out of 15 mossy fibres originating in dorsal column nuclei, for example, had a collateral which terminated in a deep nucleus (Quy, Fujita et al. 2011), and there is generally light collateral input to the dentate nucleus (such that 'suppression of PCs [Purkinje cells]...plays the primary role in generating outputs from DN [the dentate nucleaus]': Ishikawa, Tomatsu et al. 2014). Thus the Purkinje cell rate may in some circuits be the main

[11] Unless otherwise stated, 'nuclear cell/neuron' is used as shorthand for these cells (although deep nuclei also contain other cell types), and circuit as a synonym of microcomplex, including a multizonal microcomplex (a circuit that contains more than one microzone) Apps, R. and M. Garwicz (2005). "Anatomical and physiological foundations of cerebellar information processing." Nat Rev Neurosci **6**(4): 297-311.

and even sole controller of nuclear output cells, while in others it contributes control (Ishikawa, Tomatsu et al. 2014, Jirenhed and Hesslow 2016). Optogenetically controlled pauses in Purkinje cell firing produce robust motor output (Heiney, Kim et al. 2014, Lee, Mathews et al. 2015).

This appears to raise a number of problems. (1) Spike timing of Purkinje cells is individually unpredictably erratic, so on the face of it poorly suited to smoothly graded control of nuclear rates. This is seen, for example, in recordings from single cells across the step cycle of a mouse (Sauerbrei, Lubenov et al. 2015). (2) It is thought that both sensory information and motor commands are predominantly coded as firing rates (Delvendahl and Hallermann 2016). At least two spikes are need to infer a so-called instantaneous rate but a reliable 'reading' requires temporal integration over a longer period. (3) Purkinje cell-mediated inhibitory postsynaptic currents are very fast, with time constants of around 2.5 ms (Person and Raman 2012b, Mercer, Palarz et al. 2016), suggesting a short window. An optimum of 3 ms has been estimated for Purkinje cell targets that control eye movement (Payne, French et al. 2019). How, then, is a smooth response generated at short time scales relevant for behavioural control? (4) Purkinje cells are organised functionally into groups of hundreds of cells called microzones (Oscarsson 1979, Ozden, Sullivan et al. 2009, Ramirez and Stell 2016) whose output is channelled down onto a much smaller number of nuclear cells. There is no evidence of internal organisation (Bengtsson and Jorntell in Apps, Hawkes et al. 2018 p. 663). A nuclear cell therefore receives contact from a random sample of around 30-50 Purkinje cells (Person and Raman 2012a). Moreover each nuclear cell receives contact from a different and variable number of Purkinje cells. A Purkinje cell makes contact on an average of 4-5 nuclear cells, so that each Purkinje cell participates in an average of 4-5

samples. On the face of it, independent Purkinje cell rates would simultaneously drive different nuclear rates with no obvious rationale. (5) What is the purpose of the functional organisation of Purkinje cells and of nuclear cells into groups?

We propose possible answers. This is not a full proposal but is intended to establish the feasibility and need for a viable hypothesis of cerebellar function to include a reason for the interposition of deep nuclei in the translation of input to output.

5.2 Purkinje cell to nuclear cell convergence

Extracellular recordings from Purkinje cells show a phase-dependent increase and decrease in their firing rate during locomotion (in cats: Armstrong and Edgley 1984, Armstrong and Edgley 1988, Edgley and Lidierth 1988). However, individually, Purkinje cells spike erratically during a step cycle (in mice: Sauerbrei, Lubenov et al. 2015). The exact spiking pattern of a single cell can be widely different in each step (Sauerbrei, Lubenov et al. 2015 Figure 2A rasters), but gives a smooth curve averaged across many steps. The given explanation of variation is that this reflects sensitivity to variables that change in each step – exact weight distribution between feet, joint angles, centre of gravity and so on. However, Purkinje cell firing is variable even under tightly controlled conditions – for example when responding to perfectly sinusoidal head rotation in a VOR paradigm (Guo, Ke et al. 2014).

The erratic pattern (but smooth averaged rate) of Purkinje cell spiking can be explained as the expression of a smoothly changing moment-to-moment probability that a Purkinje cell spikes. Convergence of Purkinje cells onto nuclear cells means that a nuclear cell receives at any moment the sum of spikes discharged by 30-50 Purkinje cells (each making heavy

contact, discussed below). If all of these Purkinje cells have the same phase-locked firing probability (but with independent spike times) the sum of spikes behaves like the average, yielding a smoothed curve across a cycle. Synchronisation of the probability across a microzone population of Purkinje cells would mean the nuclear group that receive the output of those Purkinje cells all receive the same smooth averaged rate.

Moreover and importantly, it remains smooth even at low afferent rates. The sometimes lengthy full pause seen in individual Purkinje cells disappears from the collective count, so that Purkinje cell control of nuclear cells can be continuous, permitting smooth control despite long interspike intervals in any one Purkinje cell.

With this arrangement small changes of the summed inhibition are detectable in a short integration window, allowing a sensitive and short latency postsynaptic response. As far as we know coordinated firing probabilities for functionally grouped Purkinje cells has not been reported. But the population rate of irregularly firing floccular Purkinje cells (approximated by binning spikes across the population of recorded cells) has been shown to have a linear, rapidly-translated relationship with eye movement (Payne, French et al. 2019).

5.3 Functional randomised spike timing

This appears superficially similar to a strategy of improving the signal to noise ratio. Generally, neurons respond with different spike timing to time-varying stimulation over multiple trials (Mainen and Sejnowski 1995, de Ruyter van Steveninck, Lewen et al. 1997, Schreiber, Fellous et al. 2004). ‘To what extent this neural variability contributes to

meaningful processing (as opposed to being meaningless noise) is the fundamental question of neuronal coding' (Faisal, Selen et al. 2008 p.293).

The functional and necessary asynchrony of spike timing within a Purkinje cell group is not simply a failure to synchronise, we suggest. Convergence of erratically spiking Purkinje cells is the *solution*, rather than a problem. Less asynchrony generates a less smooth averaged rate: the extreme case is perfect synchrony. It is a feature of this arrangement, then, that irregular firing of Purkinje cells should be brought about by design.

This may be a function of stellate cells. Inhibition by stellate cells causes regular intrinsic firing of Purkinje cells to become erratic (Häusser and Clark 1997, Jelitai, Puggioni et al. 2016), and selective silencing of molecular layer interneurons causes regularity to increase (and leads to locomotor deficits, (Jelitai, Puggioni et al. 2016). Stellate cells themselves fire erratically in vitro (Häusser and Clark 1997, Carter and Regehr 2002, Ruigrok, Hensbroek et al. 2011) and even more so in vivo (Jörntell and Ekerot 2002, Barmack and Yakhnitsa 2008). Stellate cells form planar networks that lie between Purkinje cells (Palay and Chan-Palay 1974, Sultan and Bower 1998), making only a few synaptic contacts on each other (Trigo, Sakaba et al. 2012) at the soma (Lemkey-Johnston and Larramendi 1968).

Purkinje cells also form planar networks. Purkinje cells inhibit each other through recurrent collaterals (Chan-Palay 1971) which extend sagittally, usually in both directions, forming an arbour which is severely flattened in the sagittal plane, in other words aligned along the microzone. Contact on Purkinje cells is typically on near neighbours (Witter, Rudolph et al. 2016). Each Purkinje cell receives contact from an estimated 5 to 10 other Purkinje cells (Witter, Rudolph et al. 2016), mainly at the soma (Palay and Chan-Palay 1974). The similarities with stellate cell networks – inhibitory, planar networks and contact at the soma

from a low number of other cells – suggest that the recurrent inputs may also desynchronise Purkinje cell spike timing.

5.4 Rationale for functional organisation into cell groups

Functionally imperative convergence of Purkinje cells onto nuclear cells provides a reason for Purkinje cells to be the organised into functional groups. Divergence increases effective convergence without needing more Purkinje cells; but as there must be at least enough nuclear cells to diverge onto, it is also a reason for the organisation of nuclear cells into groups. Functional divergence (and therefore minimum nuclear group size) may be increased through excitatory nuclear interneurons. Deep nuclei also contain excitatory (presumed glutamatergic) interneurons that fire spontaneously and – like nuclear projection neurons – are inhibited by Purkinje cells (Uusisaari and Knopfel 2008) and are released from inhibition in the conditioned response, presumably exciting nuclear projection neurons. So each Purkinje cell may contribute to control of more nuclear cells than only the ones it contacts directly.

Why don't Purkinje cells simply diverge onto all nuclear cells in a group? Why interpose interneurons? The individually strong (Person and Raman 2012b) predominantly somatic (Uusisaari and De Schutter 2011) contact of a Purkinje cell on a nuclear cell may impose a limit on the number of synapses for which space is available (at least for significant contact). Each Purkinje cell contributes on average around 30 boutons per nuclear cell (range 24-36) (Person and Raman 2012a), most containing multiple synaptic densities (Telgkamp, Padgett et al. 2004).

FIGURE 10

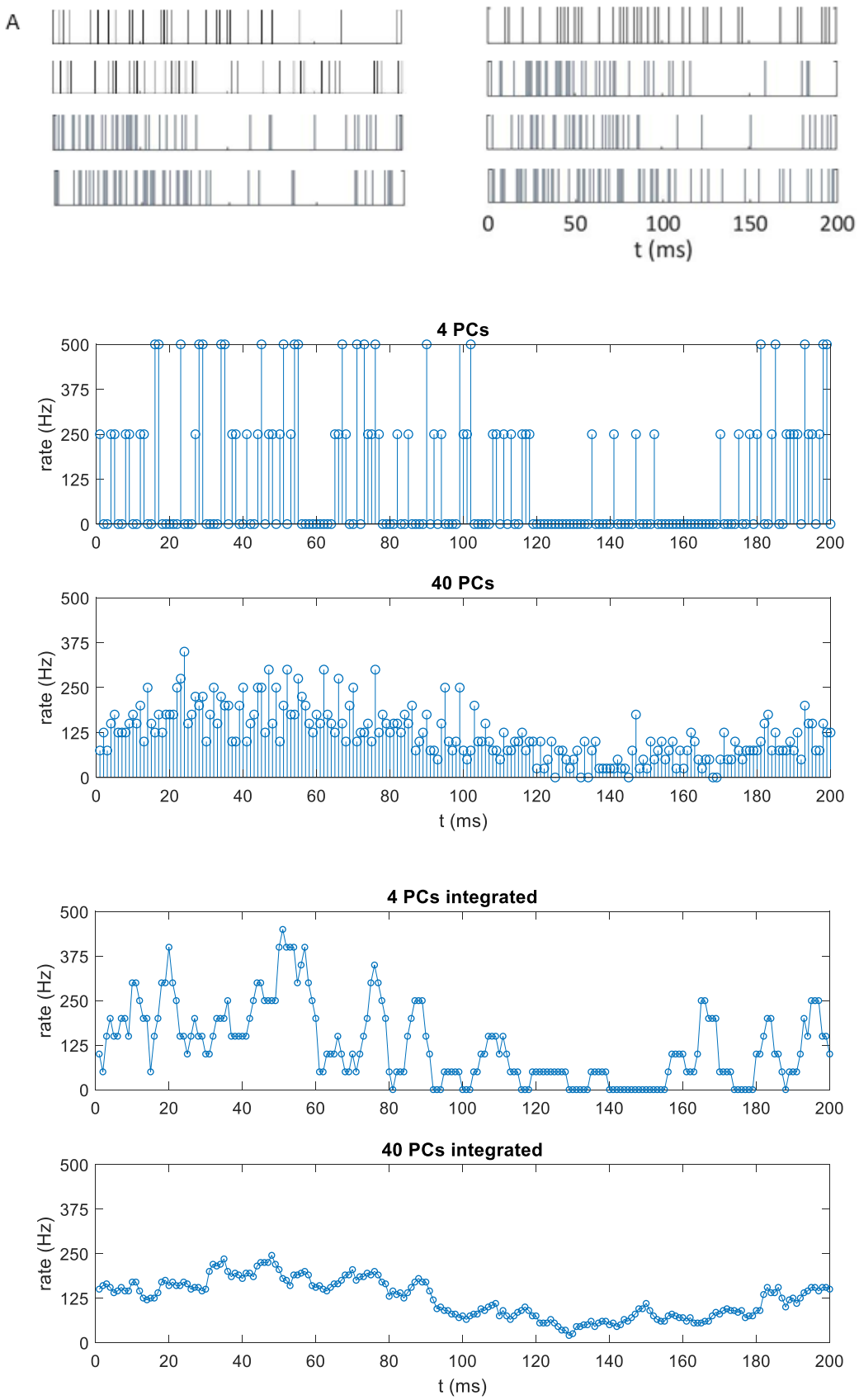


FIGURE 10

Purkinje cell-nuclear cell convergence provides incremental control at short time intervals. **A:** Simulated spiking by 8 Purkinje cells across a mouse step cycle. Each Purkinje cell fires spikes with a changing probability that oscillates sinusoidally between minimum and maximum averaged firing rates of 50-200 Hz, mimicking the range in the mouse. In the landscape panels spikes are generated with the same changing probability with time, and counted in 1 ms bins. Spikes per bin are plotted as firing rates. **1st and 3rd panels:** convergence of 4 Purkinje cells onto a nuclear cell; **2nd and 4th panels:** convergence of 40 Purkinje cells onto a nuclear cell. In the 3rd and 4th panels the firing rate reflects a 5 ms rolling integration window. (The actual duration of the nuclear integration window is unknown). In the 2nd panel (high convergence without integration) the rate still fluctuates rapidly in large steps, even in consecutive bins. The combination of 40:1 convergence plus 5 ms integration in the 4th panel reduces moment-to-moment variation and compresses the range to a relatively smoothly changing rate predicted by the probability of Purkinje cell spiking, in real time. The function is not to mitigate erratic spiking of Purkinje cells – noise (in this sense) is actively generated and functional – but to allow nuclear cells to respond rapidly to coordinated changes in the instantaneous Purkinje cell rate.

It may also be for the functional reason that the signal to noise ratio increases as the square root of the number of averaged data. That is, there is a diminishing reduction of random deviation as convergence increases. Hence there may be little to be gained from higher convergence. However, bisynaptic convergence through interneurons adds a second layer of integration – Purkinje cells onto interneurons and interneurons onto nuclear cells – with

concurrent (as opposed to sequential) integration windows. This both adds smoothness (in simulations) and reduces random variability.

There is a general consensus that the cerebellum has modular functional organisation that may be at the level of microzones (Apps, Hawkes et al. 2018). Synchronisation of spike probability (but not spike timing) among functionally-grouped Purkinje cells would neatly explain random sampling by nuclear cells of Purkinje cell rates. It is functionally immaterial which Purkinje cells contact a particular nuclear cell, and that nuclear cells each receive a different sample.

What is the evidence that the probability of Purkinje cell spiking is coordinated at microzone level? Sets of Purkinje cells located in the same sagittal plane show coordinated simple spike pauses (Ramirez and Stell 2016) which generate motor output (Heiney, Kim et al. 2014, Lee, Mathews et al. 2015). Modulation of nuclear cell firing requires the co-modulation of a large proportion of Purkinje cells that innervate the cell (Bengtsson, Ekerot et al. 2011). That is, a few Purkinje cells whose firing is not modulated can block an effect of the rest, consistent with strong individual inhibition by a single Purkinje cells of each of its targets (Telgkamp, Padgett et al. 2004, Person and Raman 2012a, Person and Raman 2012b). This would suggest that coordination of Purkinje cell firing across the whole population of a microzone is necessary for graded modulation of nuclear firing, because less than full coordination means that each nuclear cell is at high risk of receiving contact from an out-of-step cell.

Could the evidence also be explained by synchronisation of spike timing between Purkinje cells? Synchronisation of spike timing during behaviour is reported to be confined to paired recordings from Purkinje cells that are on-beam, and to be absent in recordings from paired Purkinje cells that are off-beam (Heck, Thach et al. 2007). So spiking is not synchronised

sagittally. Can synchrony in *any* direction be reconciled with stochastic spiking? Yes, if we don't make the assumption that synchronised timing of spikes is a mechanism of motor control. While it feels intuitive it is unreported, and moot, that synchronised cells are in circuits that have output during the synchronised period.[12]

5.5 A possible mechanism of Purkinje cell synchrony

There are several moot but possible mechanisms of coordination. One possibility is to configure circuits so that Purkinje cells in a microzone all receive at any moment what is functionally speaking the same parallel fibre input. We argued earlier that granule cell rates are confined to a narrow range with a fixed bandwidth (and other variables are controlled) such that control of firing of a Purkinje cell is by the only remaining functional variable, the adjustable range. How does the cerebellum contrive that the whole Purkinje cell population of a microzone receives, at any moment, the same range?

The range varies with the mean of mossy fibre rates received by the underlying – therefore mediolateral – strip of the granular layer (argued earlier). Mossy fibre terminal branches end in clusters of terminals (Wu, Sugihara et al. 1999, Sultan and Heck 2003, Shinoda and Sugihara 2013) that are aligned in the sagittal direction (Sultan 2001). Coupled with high divergence onto granule cells, this has the result of randomly intermingling mossy fibre rates in the sagittal direction. If rates are perfectly intermingled, granule cells in a sagittal

[12] Even synchrony time-locked to (forelimb extension during a reaching) movement. The explanation proposed by Heck and colleagues is that synchrony is under control of input from mossy fibres, relayed to Purkinje cells by granule cell ascending axons (and not parallel fibres). There are no paired measurements from the output cells of the circuit. Spike synchrony may be an artefact simply of receiving more excitatory input than normal (generated by the movement), overriding tonic inhibition of Purkinje cells from MLIs and each other. It may even contribute to *block* output, with asynchrony restored in the output phase.

band all randomly sample the same pool of rates. A mediolateral strip of granule cells therefore receives a mixture of rates forming a representative sample of each band it crosses, and two strips that cross the same sagittal bands contain an identical row of samples (if we assume samples are representative of the sampled band). Therefore, if all strips that supply parallel fibres to a microzone cross the same pattern of sagittal bands, they all, at any moment, contain an identical mediolateral row of such samples (like a row of squares on a chess board but without defined or fixed boundaries).

Parallel fibre activity in each strip is therefore driven, moment-by-moment, by functionally identical mossy fibre rates (and the variables that differ between samples – the number and binary pattern of active mossy fibres and permutation of rates they each fire at – are not material *to this function*). In reality, identical is probably an overstatement because it is unlikely that rates in a band are perfectly intermingled, or therefore that an intersection is fully representative of the band as a whole, but it may nonetheless be one of the strategies that contributes to coordination of granule cell rates (and this is a proposal of the model summarised here in outline).

5.6 Translation of Purkinje cell averaged rates into nuclear rates

As noted, the conditioned fall in Purkinje cell rates may in some circuits gate mossy fibre collateral input to nuclear cells, but also exerts control of its own over nuclear rates, which may in some circuits dominate control (Ishikawa, Tomatsu et al. 2014, Jirenhed and Hesslow 2016). Nuclear cells are insensitive to steady afferent rates (see also Turecek, Jackman et al. 2017), but ‘respond to increasing, as well as decreasing, changes in PC [Purkinje cell] firing

rate with immediate modification of their output firing' (Pedroarena and Schwarz 2003 p.713). There are a number of adaptations that make synaptic transmission fast and reliable. These include 'large boutons, glial ensheathment, GABA transporters confined to astrocytes, [and] multiple release sites' (Telgkamp, Padgett et al. 2004 p.123). Spillover confinement to multisynaptic Purkinje cell boutons, with transporters confined to astrocytic processes at the bouton perimeter, procures spillover-mediated transmission reported to have a high response probability of postsynaptic receptors (Telgkamp, Padgett et al. 2004). Thus spillover-mediated transmission mitigates unpredictable single synapse neurotransmitter release, so transmission is reliable and precisely timed as well as fast.

The amount of the increase or decrease in the nuclear rate is likely to be proportionate to the rate of change of the simple spike rate (assuming that firing amongst a population of Purkinje cells that contact a nuclear cell is asynchronous: Indira Raman, personal correspondence dated 4 December 2018).[13] Averaged Purkinje cell rates across a step cycle describe a repeating wave form (Sauerbrei, Lubenov et al. 2015). A downslope, where the averaged rate is falling, elevates nuclear rates, and vice versa. What controls the rate of change of the averaged rate in this phase? (The next two paragraphs recap some work referenced earlier.)

MLIs reflect 'granule cell input with linear changes in firing rate' (Jelitai, Puggioni et al. 2016 p.6). The effect of inhibition by MLIs of Purkinje cell firing is rate dependent and linear, such 'that locomotion-dependent modulation of the balance between excitation and inhibition [of Purkinje cell dendrites] generates depolarising or hyperpolarising dendritic V_m [dendritic

[13] Professor Raman considers firing may not be asynchronous, but that in theory perfectly asynchronous firing would probably have this relationship with nuclear rates.

membrane voltage] changes that linearly transform into bidirectional modulation of PC SSp [Purkinje cell simple spike] output' (Jelitai, Puggioni et al. 2016 p.9). This is a challenge for learning models, which propose that output (following training) is learned, that is, training displaces the naïve response to input rates by making iterative adjustment of parallel fibre synaptic weights (Albus 1971, Fujita 1982, Brunel, Hakim et al. 2004, Dean, Porrill et al. 2010). If instead translation is linear as reported, and control is mediated by MLIs, what controls MLI rates?

The higher concentration of synaptic contact by stellate cells on Purkinje cells at greater depth (to reason from morphology) has the (proposed) result that there is depth-dependent selection of parallel fibre signals that control the simple spike rate in the conditioned response, which deeper-stratum parallel fibres dominate. Mossy fibre input to the cerebellum is vertically topographically organised (reported for the C3 region: Jörntell and Ekerot 2006, Quy, Fujita et al. 2011). Deeper-lying granule cells receive signals triggered by limb movements (Jörntell and Ekerot 2006), suggesting it is driven by activation of internal receptors in muscles and joints: proprioception. Since topography is preserved in the molecular layer (Palay and Chan-Palay 1974, Zhang and Linden 2012), control of MLIs that dominate control of Purkinje cells is by mossy fibre signals from this source.

The control of Purkinje cell firing by proprioception, and linear translation of firing rates, argue that output (of C3 circuits) has a linear, rapidly-translated relationship with movement, where input to a circuit is provided by movement to which the circuit contributes control (provided input is proportionate to movement parameters). Moreover timing is phase-locked, so also effectively movement selected.

To return to the question raised earlier regarding ‘interference’ from lateral inhibition by MLIs, coordination of Purkinje cell firing means it does not matter which stellate cells they receive inhibition from provided the size (area in the sagittal plane) of the afferent network of stellate cells that contact a Purkinje cell is uniform. Indeed an area increased by the range of lateral inhibition should mean the network receives a more representative sample of parallel fibre rates.

6. SOMATOTOPY

What is the evidence that the body location that provides the bulk of the input to a circuit receives the output of the circuit?

A repeated finding of mapping studies is that cerebellar circuits have a complex relationship with the coordinates they represent. Mapping mossy fibre input has long been reported to generate a patchy mosaic i.e. discontinuous representation of body parts (but somatotopy within the shards: the 'fractured somatotopy' of Shambes, Gibson et al. 1978). One reason is that circuits receive (as mossy fibre input) information from more than one source. In the C3 region ‘each microzone receives input from several mossy fibre receptive fields’ (Apps and Garwicz 2005 p.305). Also, it is common for signals originating from the same external source to be sent to more than one circuit (for mossy fibres: Voogd, Pardoe et al. 2003, Pijpers, Apps et al. 2006, for climbing fibres: Fujita and Sugihara 2013).

Mapping by output (stimulating sites in the cerebellum to see which muscles contract) does not yield a more orderly topographical organisation than the fractured input map (stimulating the body surface or muscle and mapping the response in the cerebellum). An

example in humans has been provided by electrical stimulation (60 Hz for 2 seconds) of the posterior cerebellum in patients in surgery for tumours (Mottolese, Richard et al. 2013). The evoked movements are strictly ipsilateral but otherwise 'the same body part can be represented in different sectors, different body parts can be embodied in a single region and the [cerebellar] cortical size of a body part is proportional to its functional importance rather than its physical size' (ibid. p.337).

Nonetheless there is order. The response of climbing fibres to somatosensory stimuli has allowed microzones to be somatotopically mapped i.e. matched with the receptive field (region of the body surface) to which they are responsive (Garwicz, Ekerot et al. 1998, Garwicz, Jörntell et al. 1998). In cats, 'climbing fibres in adjacent microzones are activated from adjacent skin areas, forming a detailed somatotopic map of the ipsilateral forelimb skin, particularly in distal parts' (Ekerot and Jörntell 2003 p.101). Mossy fibre input representing stimulation of the body surface is normally sent to the same circuit as the climbing fibre signal (evoked by the same stimultation) (Garwicz, Jörntell et al. 1998, Voogd, Pardoe et al. 2003, Odeh, Ackerley et al. 2005, Pijpers, Apps et al. 2006, Apps and Hawkes 2009). Evidence is inconclusive but indicates that regions of the cerebellar cortex which map to particular body parts receive input from regions of the pons which in turn receive input from regions of the cerebral cortex which map to the same body parts (Odeh, Ackerley et al. 2005). A study of the relationship of the receptive fields of microzones in the C1, C3 and Y cerebellar cortical zones and movements controlled by anterior interpositus sites to which output of those microzones had been mapped, showed 'a general specificity of the input-output relationship' (Ekerot, Jorntell et al. 1995 p.365).

‘Perhaps nowhere else in the history of ideas has there been a more striking pattern of reliance on metaphors than in the history of reflection about the brain’ (Daugman 1993 p.23). Mechanical analogies for the brain have been with us for centuries. Other analogies include hydraulics, electronics and telegraph metaphors (Hodgkin and Huxley borrowed the mathematics developed for signal propagation in coaxial cables to model the generation of action potentials), and thermodynamics. The computer metaphor, and network models imported from statistics, have been so widely adopted the lines have become blurred between language used by analogy and proposals that are intended to be literal (see Werner 2011 for a review). References to computation, for example, are used in a way that suggests there is thought to be no need for them to be qualified.

The belief that artifical intelligence may provide insight into brain function by analogy is widely accepted. In the cerebellum, it has become a presumption that the cerebellum is the physiological analogue of an algorithm that computes output from input values in a learned pattern of active cells. This may be an entrenched mistake to which we outline here an alternative.

REFERENCES

Abrahamsson, T., L. Cathala, K. Matsui, R. Shigemoto and D. A. Digregorio (2012). "Thin dendrites of cerebellar interneurons confer sublinear synaptic integration and a gradient of short-term plasticity." Neuron **73**(6): 1159-1172.

Albus, J. S. (1971). A Theory of Cerebellar Function.

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

Apps, R. and R. Hawkes (2009). "Cerebellar cortical organization: a one-map hypothesis." Nat Rev Neurosci **10**(9): 670-681.

Apps, R., R. Hawkes, S. Aoki, F. Bengtsson, A. M. Brown, G. Chen, T. J. Ebner, P. Isope, H. Jorntell, E. P. Lackey, C. Lawrenson, B. Lumb, M. Schonewille, R. V. Sillitoe, L. Spaeth, I. Sugihara, A. Valera, J. Voogd, D. R. Wylie and T. J. H. Ruigrok (2018). "Cerebellar Modules and Their Role as Operational Cerebellar Processing Units." Cerebellum.

Armstrong, D. M. and S. A. Edgley (1984). "Discharges of Purkinje cells in the paravermal part of the cerebellar anterior lobe during locomotion in the cat." The Journal of physiology **352**: 403-424.

Armstrong, D. M. and S. A. Edgley (1988). "Discharges of interpositus and Purkinje cells of the cat cerebellum during locomotion under different conditions." J Physiol **400**: 425-445.

Barmack, N. H. and V. Yakhnitsa (2008). "Functions of interneurons in mouse cerebellum." J Neurosci **28**(5): 1140-1152.

Baumel, Y., G. A. Jacobson and D. Cohen (2009). "Implications of functional anatomy on information processing in the deep cerebellar nuclei." Front Cell Neurosci **3**: 14.

Bengtsson, F., C. F. Ekerot and H. Jorntell (2011). "In vivo analysis of inhibitory synaptic inputs and rebounds in deep cerebellar nuclear neurons." PLoS One **6**(4): e18822.

Bengtsson, F. and H. Jörntell (2009). "Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs." Proc Natl Acad Sci U S A **106**(7): 2389-2394.

Billings, G., E. Piasini, A. Lorincz, Z. Nusser and R. A. Silver (2014). "Network structure within the cerebellar input layer enables lossless sparse encoding." Neuron **83**(4): 960-974.

Brand, S., A. L. Dahl and E. Mugnaini (1976). "The length of parallel fibers in the cat cerebellar cortex. An experimental light and electron microscopic study." Exp Brain Res **26**(1): 39-58.

Brunel, N., V. Hakim, P. Isope, J. P. Nadal and B. Barbour (2004). "Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell." Neuron **43**(5): 745-757.

Carter, A. G. and W. G. Regehr (2002). "Quantal events shape cerebellar interneuron firing." Nat Neurosci **5**(12): 1309-1318.

Cayco-Gajic, N. A., C. Clopath and R. A. Silver (2017). "Sparse synaptic connectivity is required for decorrelation and pattern separation in feedforward networks." Nat Commun **8**(1): 1116.

Cayco-Gajic, N. A. and R. A. Silver (2019). "Re-evaluating Circuit Mechanisms Underlying Pattern Separation." Neuron **101**(4): 584-602.

Cerminara, N. L. and J. A. Rawson (2004). "Evidence that climbing fibers control an intrinsic spike generator in cerebellar Purkinje cells." J Neurosci **24**(19): 4510-4517.

Cesana, E., K. Pietrajtis, C. Bidoret, P. Isope, E. D'Angelo, S. Dieudonne and L. Forti (2013). "Granule cell ascending axon excitatory synapses onto Golgi cells implement a potent feedback circuit in the cerebellar granular layer." J Neurosci **33**(30): 12430-12446.

Chan-Palay, V. (1971). "The recurrent collaterals of Purkinje cell axons: a correlated study of the rat's cerebellar cortex with electron microscopy and the Golgi method." Z Anat Entwicklungsgesch **134**(2): 200-234.

Daugman, J. G. (1993). Brain metaphor and brain theory. Computational Neuroscience. S. E.L., Cambridge, MA: MIT Press.

de Ruyter van Steveninck, R. R., G. D. Lewen, S. P. Strong, R. Koberle and W. Bialek (1997). "Reproducibility and variability in neural spike trains." Science **275**(5307): 1805-1808.

de Zeeuw, C. I. and A. S. Berrebi (1996). "Individual Purkinje cell axons terminate on both inhibitory and excitatory neurons in the cerebellar and vestibular nuclei." Ann N Y Acad Sci **781**: 607-610.

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

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

Duguid, I., T. Branco, M. London, P. Chadderton and M. Hausser (2012). "Tonic inhibition enhances fidelity of sensory information transmission in the cerebellar cortex." J Neurosci **32**(32): 11132-11143.

Eccles, J. C., M. Ito and J. Szentágothai (1967). The cerebellum as a neuronal machine. Berlin, New York etc., Springer-Verlag.

Edgley, S. A. and M. Lidierth (1988). "Step-related discharges of Purkinje cells in the paravermal cortex of the cerebellar anterior lobe in the cat." J Physiol **401**: 399-415.

Ekerot, C. F. and H. Jörntell (2001). "Parallel fibre receptive fields of Purkinje cells and interneurons are climbing fibre-specific." Eur J Neurosci **13**(7): 1303-1310.

Ekerot, C. F. and H. Jörntell (2003). "Parallel fiber receptive fields: a key to understanding cerebellar operation and learning." Cerebellum **2**(2): 101-109.

Ekerot, C. F., H. Jorntell and M. Garwicz (1995). "Functional relation between corticonuclear input and movements evoked on microstimulation in cerebellar nucleus interpositus anterior in the cat." Exp Brain Res **106**(3): 365-376.

Faisal, A. A., L. P. J. Selen and D. M. Wolpert (2008). "Noise in the nervous system." Nature reviews. Neuroscience **9**(4): 292-303.

Fujita, H. and I. Sugihara (2013). "Branching patterns of olivocerebellar axons in relation to the compartmental organization of the cerebellum." Front Neural Circuits **7**: 3.

Fujita, M. (1982). "Adaptive filter model of the cerebellum." Biol Cybern **45**(3): 195-206.

Gao, Z., M. Proietti-Onori, Z. Lin, M. M. Ten Brinke, H. J. Boele, J. W. Potters, T. J. Ruigrok, F. E. Hoebeek and C. I. De Zeeuw (2016). "Excitatory Cerebellar Nucleocortical Circuit Provides Internal Amplification during Associative Conditioning." Neuron **89**(3): 645-657.

Garwicz, M., C. F. Ekerot and H. Jorntell (1998). "Organizational principles of cerebellar neuronal circuitry." News in Physiological Sciences **13**: 28-34.

Garwicz, M., H. Jörntell and C. F. Ekerot (1998). "Cutaneous receptive fields and topography of mossy fibres and climbing fibres projecting to cat cerebellar C3 zone." J Physiol **512 (Pt 1)**: 277-293.

Guo, C. C., M. C. Ke and J. L. Raymond (2014). "Cerebellar encoding of multiple candidate error cues in the service of motor learning." J Neurosci **34**(30): 9880-9890.

Hansel, C., D. J. Linden and E. D'Angelo (2001). "Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum." Nat Neurosci **4**(5): 467-475.

Harvey, R. J. and R. M. Napper (1988). "Quantitative study of granule and Purkinje cells in the cerebellar cortex of the rat." J Comp Neurol **274**(2): 151-157.

Harvey, R. J. and R. M. Napper (1991). "Quantitative studies on the mammalian cerebellum." Prog Neurobiol **36**(6): 437-463.

Häusser, M. and B. A. Clark (1997). "Tonic synaptic inhibition modulates neuronal output pattern and spatiotemporal synaptic integration." Neuron **19**(3): 665-678.

Heck, D. H., W. T. Thach and J. G. Keating (2007). "On-beam synchrony in the cerebellum as the mechanism for the timing and coordination of movement." Proceedings of the National Academy of Sciences **104**(18): 7658.

Heiney, S. A., J. Kim, G. J. Augustine and J. F. Medina (2014). "Precise control of movement kinematics by optogenetic inhibition of Purkinje cell activity." J Neurosci **34**(6): 2321-2330.

Houck, B. D. and A. L. Person (2015). "Cerebellar Premotor Output Neurons Collateralize to Innervate the Cerebellar Cortex." J Comp Neurol **523**(15): 2254-2271.

Ishikawa, T., S. Tomatsu, Y. Tsunoda, J. Lee, D. S. Hoffman and S. Kakei (2014). "Releasing dentate nucleus cells from Purkinje cell inhibition generates output from the cerebrocerebellum." PLoS One **9**(10): e108774.

Isope, P. and B. Barbour (2002). "Properties of unitary granule cell-->Purkinje cell synapses in adult rat cerebellar slices." J Neurosci **22**(22): 9668-9678.

Ito, M. (2001). "Cerebellar long-term depression: characterization, signal transduction, and functional roles." Physiol Rev **81**(3): 1143-1195.

Jakab, R. L. and J. Hámori (1988). "Quantitative morphology and synaptology of cerebellar glomeruli in the rat." Anat Embryol (Berl) **179**(1): 81-88.

Jelitai, M., P. Puggioni, T. Ishikawa, A. Rinaldi and I. Duguid (2016). "Dendritic excitation-inhibition balance shapes cerebellar output during motor behaviour." Nat Commun **7**: 13722.

Jirenhed, D. A., F. Bengtsson and G. Hesslow (2007). "Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace." J Neurosci **27**(10): 2493-2502.

Jirenhed, D. A. and G. Hesslow (2016). "Are Purkinje Cell Pauses Drivers of Classically Conditioned Blink Responses?" The Cerebellum **15**(4): 526-534.

Johansson, F., H. A. Carlsson, A. Rasmussen, C. H. Yeo and G. Hesslow (2015). "Activation of a Temporal Memory in Purkinje Cells by the mGluR7 Receptor." Cell Rep **13**(9): 1741-1746.

Johansson, F., D. A. Jirenhed, A. Rasmussen, R. Zucca and G. Hesslow (2014). "Memory trace and timing mechanism localized to cerebellar Purkinje cells." Proc Natl Acad Sci U S A **111**(41): 14930-14934.

Jörntell, H. and C. F. Ekerot (2002). "Reciprocal bidirectional plasticity of parallel fiber receptive fields in cerebellar Purkinje cells and their afferent interneurons." Neuron **34**(5): 797-806.

Jörntell, H. and C. F. Ekerot (2003). "Receptive field plasticity profoundly alters the cutaneous parallel fiber synaptic input to cerebellar interneurons in vivo." J Neurosci **23**(29): 9620-9631.

Jörntell, H. and C. F. Ekerot (2006). "Properties of somatosensory synaptic integration in cerebellar granule cells in vivo." J Neurosci **26**(45): 11786-11797.

Jörntell, H. and C. F. Ekerot (2011). "Receptive Field Remodeling Induced by Skin Stimulation in Cerebellar Neurons in vivo." Front Neural Circuits **5**: 3.

Kanichay, R. T. and R. A. Silver (2008). "Synaptic and cellular properties of the feedforward inhibitory circuit within the input layer of the cerebellar cortex." J Neurosci **28**(36): 8955-8967.

Lee, K. H., P. J. Mathews, A. M. Reeves, K. Y. Choe, S. A. Jami, R. E. Serrano and T. S. Otis (2015). "Circuit mechanisms underlying motor memory formation in the cerebellum." Neuron **86**(2): 529-540.

Lemkey-Johnston, N. and L. M. Larramendi (1968). "Types and distribution of synapses upon basket and stellate cells of the mouse cerebellum: an electron microscopic study." J Comp Neurol **134**(1): 73-112.

Loeb, G. E. and M. Mileusnic (2015). "Proprioceptors and models of transduction." Scholarpedia **10**: 12390.

Mainen, Z. F. and T. J. Sejnowski (1995). "Reliability of spike timing in neocortical neurons." Science **268**(5216): 1503-1506.

Marr, D. (1969). "A theory of cerebellar cortex." J Physiol **202**(2): 437-470.

Mercer, A. A., K. J. Palarz, N. Tabatadze, C. S. Woolley and I. M. Raman (2016). "Sex differences in cerebellar synaptic transmission and sex-specific responses to autism-linked Gabrb3 mutations in mice." eLife **5**: e07596.

Mitchell, S. J. and R. A. Silver (2000a). "GABA spillover from single inhibitory axons suppresses low-frequency excitatory transmission at the cerebellar glomerulus." J Neurosci **20**(23): 8651-8658.

Mitchell, S. J. and R. A. Silver (2000b). "Glutamate spillover suppresses inhibition by activating presynaptic mGluRs." Nature **404**(6777): 498-502.

Mottolese, C., N. Richard, S. Harquel, A. Szathmari, A. Sirigu and M. Desmurget (2013). "Mapping motor representations in the human cerebellum." Brain **136**(Pt 1): 330-342.

Mugnaini, E. (1983). "The length of cerebellar parallel fibers in chicken and rhesus monkey." J Comp Neurol **220**(1): 7-15.

Napper, R. M. and R. J. Harvey (1988). "Number of parallel fiber synapses on an individual Purkinje cell in the cerebellum of the rat." J Comp Neurol **274**(2): 168-177.

Odeh, F., R. Ackerley, J. G. Bjaalie and R. Apps (2005). "Pontine maps linking somatosensory and cerebellar cortices are in register with climbing fiber somatotopy." J Neurosci **25**(24): 5680-5690.

Oscarsson, O. (1979). "Functional units of the cerebellum - sagittal zones and microzones." Trends in Neurosciences **2**: 143-145.

Ozden, I., M. R. Sullivan, H. M. Lee and S. S. Wang (2009). "Reliable coding emerges from coactivation of climbing fibers in microbands of cerebellar Purkinje neurons." J Neurosci **29**(34): 10463-10473.

Palay, S. L. and V. Chan-Palay (1974). Cerebellar cortex : cytology and organization. Berlin, Springer.

Pantò, M. R., A. Zappalà, R. Parenti, M. F. Serapide and F. Cicirata (2001). "Corticonuclear projections of the cerebellum preserve both anteroposterior and mediolateral pairing patterns." European Journal of Neuroscience **13**(4): 694-708.

Paula-Barbosa, M. M., M. A. Tavares, C. Ruela and H. Barroca (1983). "The distribution of stellate cell descending axons in the rat cerebellum: a Golgi and a combined Golgi-electron microscopical study." J Anat **137 (Pt 4)**: 757-764.

Payne, H. L., R. L. French, C. C. Guo, T. B. Nguyen-Vu, T. Manninen and J. L. Raymond (2019). "Cerebellar Purkinje cells control eye movements with a rapid rate code that is invariant to spike irregularity." Elife **8**.

Pedroarena, C. M. and C. Schwarz (2003). "Efficacy and short-term plasticity at GABAergic synapses between Purkinje and cerebellar nuclei neurons." J Neurophysiol **89**(2): 704-715.

Person, A. L. and I. M. Raman (2012a). "Synchrony and neural coding in cerebellar circuits." Frontiers in Neural Circuits **6**: 97.

Person, A. L. and I. M. Raman (2012b). "Purkinje neuron synchrony elicits time-locked spiking in the cerebellar nuclei." Nature **481**(7382): 502-505.

Pijpers, A., R. Apps, J. Pardoe, J. Voogd and T. J. Ruigrok (2006). "Precise spatial relationships between mossy fibers and climbing fibers in rat cerebellar cortical zones." J Neurosci **26**(46): 12067-12080.

Qiu, D. L. and T. Knopfel (2009). "Presynaptically expressed long-term depression at cerebellar parallel fiber synapses." Pflugers Arch **457**(4): 865-875.

Quy, P. N., H. Fujita, Y. Sakamoto, J. Na and I. Sugihara (2011). "Projection patterns of single mossy fiber axons originating from the dorsal column nuclei mapped on the aldolase C compartments in the rat cerebellar cortex." J Comp Neurol **519**(5): 874-899.

Raman, I. M. and B. P. Bean (1999). "Ionic currents underlying spontaneous action potentials in isolated cerebellar Purkinje neurons." J Neurosci **19**(5): 1663-1674.

Ramirez, J. E. and B. M. Stell (2016). "Calcium Imaging Reveals Coordinated Simple Spike Pauses in Populations of Cerebellar Purkinje Cells." Cell Reports **17**(12): 3125-3132.

Rancillac, A. and F. Crépel (2004). "Synapses between parallel fibres and stellate cells express long-term changes in synaptic efficacy in rat cerebellum." J Physiol **554**(Pt 3): 707-720.

Rancz, E. A., T. Ishikawa, I. Duguid, P. Chadderton, S. Mahon and M. Hausser (2007). "High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons." Nature **450**(7173): 1245-1248.

Rasmussen, A., D. A. Jirenhed and G. Hesslow (2008). "Simple and complex spike firing patterns in Purkinje cells during classical conditioning." Cerebellum **7**(4): 563-566.

Raymond, J. L. and J. F. Medina (2018). "Computational Principles of Supervised Learning in the Cerebellum." Annu Rev Neurosci **41**: 233-253.

Ritzau-Jost, A., I. Delvendahl, A. Rings, N. Byczkowicz, H. Harada, R. Shigemoto, J. Hirrlinger, J. Eilers and S. Hallermann (2014). "Ultrafast action potentials mediate kilohertz signaling at a central synapse." Neuron **84**(1): 152-163.

Ruigrok, T. J., R. A. Hensbroek and J. I. Simpson (2011). "Spontaneous activity signatures of morphologically identified interneurons in the vestibulocerebellum." J Neurosci **31**(2): 712-724.

Sauerbrei, B. A., E. V. Lubenov and A. G. Siapas (2015). "Structured Variability in Purkinje Cell Activity during Locomotion." Neuron **87**(4): 840-852.

Schreiber, S., J. M. Fellous, P. Tiesinga and T. J. Sejnowski (2004). "Influence of ionic conductances on spike timing reliability of cortical neurons for suprathreshold rhythmic inputs." J Neurophysiol **91**(1): 194-205.

Shambes, G. M., J. M. Gibson and W. Welker (1978). "Fractured somatotopy in granule cell tactile areas of rat cerebellar hemispheres revealed by micromapping." Brain Behav Evol **15**(2): 94-140.

Shinoda, Y. and I. Sugihara (2013). Axonal Trajectories of Single Climbing and Mossy Fiber Neurons in the Cerebellar Cortex and Nucleus. Handbook of the Cerebellum and Cerebellar Disorders. M. Manto, J. D. Schmahmann, F. Rossi, D. L. Gruol and N. Koibuchi. Dordrecht, Springer Netherlands**:** 437-467.

Smith, S. L. and T. S. Otis (2005). "Pattern-dependent, simultaneous plasticity differentially transforms the input-output relationship of a feedforward circuit." Proc Natl Acad Sci U S A **102**(41): 14901-14906.

Sultan, F. (2001). "Distribution of mossy fibre rosettes in the cerebellum of cat and mice: evidence for a parasagittal organization at the single fibre level." Eur J Neurosci **13**(11): 2123-2130.

Sultan, F. and J. M. Bower (1998). "Quantitative Golgi study of the rat cerebellar molecular layer interneurons using principal component analysis." J Comp Neurol **393**(3): 353-373.

Sultan, F. and D. Heck (2003). "Detection of sequences in the cerebellar cortex: numerical estimate of the possible number of tidal-wave inducing sequences represented." J Physiol Paris **97**(4-6): 591-600.

Telgkamp, P., D. E. Padgett, V. A. Ledoux, C. S. Woolley and I. M. Raman (2004). "Maintenance of high-frequency transmission at purkinje to cerebellar nuclear synapses by spillover from boutons with multiple release sites." Neuron **41**(1): 113-126.

Telgkamp, P. and I. M. Raman (2002). "Depression of inhibitory synaptic transmission between Purkinje cells and neurons of the cerebellar nuclei." J Neurosci **22**(19): 8447-8457.

Tran-Van-Minh, A., T. Abrahamsson, L. Cathala and D. A. DiGregorio (2016). "Differential Dendritic Integration of Synaptic Potentials and Calcium in Cerebellar Interneurons." Neuron **91**(4): 837-850.

Trigo, F. F., T. Sakaba, D. Ogden and A. Marty (2012). "Readily releasable pool of synaptic vesicles measured at single synaptic contacts." Proc Natl Acad Sci U S A **109**(44): 18138-18143.

Turecek, J., S. L. Jackman and W. G. Regehr (2017). "Synaptotagmin 7 confers frequency invariance onto specialized depressing synapses." Nature **551**: 503.

Uusisaari, M. and E. De Schutter (2011). "The mysterious microcircuitry of the cerebellar nuclei." J Physiol **589**(Pt 14): 3441-3457.

Uusisaari, M. and T. Knopfel (2008). "GABAergic synaptic communication in the GABAergic and non-GABAergic cells in the deep cerebellar nuclei." Neuroscience **156**(3): 537-549.

van Beugen, B. J., Z. Gao, H.-J. Boele, F. E. Hoebeek and C. I. De Zeeuw (2013). "High Frequency Burst Firing of Granule Cells Ensures Transmission at the Parallel Fiber to Purkinje Cell Synapse at the Cost of Temporal Coding." Frontiers in Neural Circuits **7**: 95.

Voogd, J., J. Pardoe, T. J. Ruigrok and R. Apps (2003). "The distribution of climbing and mossy fiber collateral branches from the copula pyramidis and the paramedian lobule: congruence of climbing fiber cortical zones and the pattern of zebrin banding within the rat cerebellum." J Neurosci **23**(11): 4645-4656.

Werner, G. (2011). "Letting the brain speak for itself." Front Physiol **2**: 60.

Witter, L., S. Rudolph, R. T. Pressler, S. I. Lahlaf and W. G. Regehr (2016). "Purkinje Cell Collaterals Enable Output Signals from the Cerebellar Cortex to Feed Back to Purkinje Cells and Interneurons." Neuron **91**(2): 312-319.

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

Zhang, W. and D. J. Linden (2012). "Calcium influx measured at single presynaptic boutons of cerebellar granule cell ascending axons and parallel fibers." Cerebellum **11**(1): 121-131.

Zhou, H., Z. Lin, K. Voges, C. Ju, Z. Gao, L. W. Bosman, T. J. Ruigrok, F. E. Hoebeek, C. I. De Zeeuw and M. Schonewille (2014). "Cerebellar modules operate at different frequencies." Elife **3**: e02536.