Supporting Information

Walter and Khodakhah 10.1073/pnas.0812348106

SI Text

Artificial Neural Network. To estimate the pattern recognition capacity of Purkinje cells, we developed an artificial neural network representing a Purkinje cell in the cerebellar cortex (P-ANN). The P-ANN incorporated our empirically determined response variability and pattern size from experiments done in acutely prepared rat cerebellar slices as well as the linear algorithm. The connectivity and learning rule of the P-ANN used here were modeled after a commonly used Purkinje cell artificial neural network (1, 2) and simulated a Purkinje cell receiving 150,000 independent parallel fiber inputs.

Design of the P-ANN. The P-ANN was made of 2 components. The first was an ANN that grouped parallel fibers into patterns and determined the consequence of learning on the strength of each pattern. This aspect of the ANN was virtually identical to that of Steuber et al. (2), but our ANN incorporated the linear firing rate algorithm instead of pauses. The second component incorporated the response variability of Purkinje cells to the repeated activation of different parallel fiber input patterns of comparable strength which was experimentally determined in acutely prepared rat cerebellar slices (Fig. 1*D* of the main text). In essence, in our P-ANN the parameters obtained experimentally from real Purkinje cells in slices replaced the parameters obtained from the in silico biophysical model of a Purkinje cell in the simulations of Steuber et al. (2).

The output of the ANN used here was a linear function of the strength of its input (3). With asynchronous parallel fiber input and intact inhibition, 650 novel inputs increased the firing rate by 200 spikes per second. Under these conditions, learned patterns were generated by randomly selecting 650 inputs from the entire pool of 150,000 parallel fibers and decreasing the strength of all inputs constituting the pattern by 50% to mimic long-term depression (LTD) of the parallel fiber-to-Purkinje cell synapse (4–6). Furthermore, the novel patterns were also randomly selected from the entire pool of 150,000 parallel inputs and could thus include inputs that as a consequence of learning had undergone LTD.

Estimating Pattern Recognition Capacity. We used this P-ANN to evaluate the capacity of a Purkinje cell in pattern recognition by altering the number of patterns it had to learn and quantifying its ability to distinguish between learned and a similar number of novel patterns. We assessed the resulting pattern recognition capacity by calculating the signal-to-noise ratio (s/n) of the maximum firing rate of a Purkinje cell in response to learned patterns as compared to novel ones using:

$$\text{s/n} = \frac{2(\mu_l - \mu_n)^2}{\sigma_l^2 + \sigma_n^2}$$

where μ_l and μ_n and are the means and σ_l and σ_n are the standard deviations of the learned and novel maximum firing rate distributions (2, 7, 8).

Whereas in the absence of learning, each pattern increased the firing of a Purkinje cell by 200 spikes per second, given the learning rule described above, after learning, the strengths of all learned patterns were reduced to that which increased the firing rate by 100 spikes per second. If one assumes that the spontaneous baseline firing rate of a Purkinje cell before presentation of a pattern is 50 spikes per second, then presentation of every learned pattern should increase the firing rate of a Purkinje cell to 150 spikes per second (Fig. S1 *A* and *C*). However, because of the inherent variability in the response of a Purkinje cell to the repeated presentation of the same stimulus strength, the actual distribution of Purkinje cell firing rates in response to learned patterns was a Gaussian with a mean, μ_l, of 150 spikes per second and as determined from the experiments shown in Fig. 1*D* of the main text an experimentally determined standard deviation, σ_l, of ≈17 spikes per second (see Fig. S1 *B* and *D*, blue).

Before learning, the Purkinje cell maximum firing rate distribution resulting from the activation of novel patterns was also Gaussian. However, in contrast to the distribution of learned patterns, the mean and the standard deviation of the distribution of the response of Purkinje cells to novel patterns were both a function of the number of patterns to be learned (Fig. S1 *A* and *C*, red). For instance, the mean maximum firing rate of the novel distribution, μ_n, was inversely correlated to the number of patterns learned. This is because the larger the number of learned patterns, the higher the probability that some of the depressed inputs from learned patterns will also be included in novel patterns. Thus, the larger the number of patterns to be learned, the smaller the differences between the means of the novel and learned distributions (Fig. S1 *A* and *C*).

As shown in Fig. S1 *B* and *D* (purple), in the P-ANN the standard deviation of the distribution of the response of Purkinje cells to novel patterns, σ_n, was the algebraic sum of a term that represented the impact of the inclusion of depressed inputs in novel patterns from the ANN (Fig. S1 *A* and *C*, red) and the experimentally determined inherent response variability of ≈26 spikes per second for Purkinje cells at 250 spikes per second. This first term was obtained from the ANN simulation and increased with increasing the number of patterns to be learned. As the number of patterns to be learned was increased, a decrease in the mean and a concurrent increase in the standard deviation of the novel distribution maximum firing rates were combined to decrease the s/n (Fig. S1 *B* and *D*).

The P-ANN simulation was computed for different numbers of parallel fibers forming a pattern. Each pattern size corresponded to its specific set of experimental conditions.

Adaptation of the ANN to Estimate Pattern Recognition Capacity When Information Is Encoded as Pauses. We also examined the pattern recognition capacity of a Purkinje cell, assuming that it uses pauses to encode information instead of firing rate. We used the same ANN and approach of incorporating experimental variability in determining the s/n. We restricted a thorough analysis to synchronous inputs because, with asynchronous inputs, it was not possible to get a burst–pause response with inhibition intact. After block of inhibitory synaptic transmission it was possible to get burst–pause responses, although under these conditions the pauses were so short and had such a relatively large coefficient of variation that when their variability was incorporated into our P-ANN, the pattern recognition capacity was marginal.

To adapt our P-ANN to estimate pattern recognition capacity of Purkinje cells if they encode information as pauses, we first estimated the number of inputs that form a pattern. A burst–pause response in Purkinje cells is generated when the strength of the parallel fiber input is beyond that which increases the maximum firing rate to ≈250 spikes per second. With synchronous inputs and with inhibition blocked, we estimate that release of neurotransmitter from a minimum of 100 parallel fibers is needed to produce a burst–pause response. However, because

100 parallel fibers represent the minimum synaptic strength that produces a burst–pause response, they will no longer be strong enough to generate a pause if they are weakened as a consequence of learning. To accommodate this, we designated the number of parallel fibers that form a pattern as 200, such that after learning they had sufficient strength to produce a burst–pause response. In doing so, we assumed that doubling the strength of the input doubled the pause length. This is somewhat of a generous assumption because, although the pause duration is a linear function of the EPSP and the charge injected by the inputs, the slope of the correlation is <1 (unpublished observations and ref. 2). The consequence of this assumption is that the ANN would overestimate the pattern recognition capacity when information is encoded as pauses.

In acutely prepared rat cerebellar slices and with synaptic inhibition blocked, strong parallel fiber synaptic inputs generate pauses that had an average duration of 82 ± 53 ms (mean ± SD.) and that approximately halved in duration (43 ± 22 ms) after a standard LTD protocol (2). Based on these observations, the ANN was reconfigured such that before learning, an input pattern generated a 82-ms-long pause, and after learning, the duration of the pause was halved to 41 ms. Similar to the assumption made for our ANN that was based on the linear firing rate algorithm, here too, we assumed that the duration of the pause was a linear function of the strength of parallel fiber inputs. To incorporate the variability of real Purkinje cells into the P-ANN (as done for the linear firing rate algorithm), based on the experiments on Purkinje cells in acutely prepared rat cerebellar slices reported above, we used an average pause SD of 17 ms. We then varied the number of patterns that a Purkinje cells had to learn and calculated the corresponding s/n exactly as it was done for estimating pattern recognition capacity when Purkinje cells encoded information in their firing rate. Based on this P-ANN, the estimated s/n for discriminating between pauses before and after learning was ≈6, a value that is in remarkable agreement with that of 5.6 ± 1.3 found experimentally (2).

Estimating the Number of Parallel Fiber Inputs That Form a Pattern. An estimate of the number of parallel fiber inputs constituting a pattern was required to determine the pattern recognition capacity of Purkinje cells. We assumed that Purkinje cells use their full linear dynamic range and that activation of an unlearned pattern increases the firing rate of a Purkinje cell from a spontaneous rate of 50 to ≈250 spikes per second. It has recently been demonstrated that the maximum firing rate of Purkinje cells after stimulus is linearly related to the charge injected by the corresponding EPSC and that, assuming a membrane potential of −60 mV, the asynchronous injection of ≈30 pC of charge is needed to drive the firing of Purkinje cells to ≈250 spikes per second when inhibitory synaptic transmission is pharmacologically blocked (3) (see also Fig. S2*A*). The charge needed to increase the firing rate to ≈250 spikes per second is larger in the presence of intact inhibition, and it is on the order of ≈78 pC (3).

If each input releases neurotransmitter only once, then the number of inputs in a pattern is simply obtained by dividing the above charge estimates by the average charge injected by a single parallel fiber. We thus experimentally determined the charge injected by single parallel fiber inputs by electrically stimulating the granule cells and measuring resulting EPSCs in Purkinje cells voltage clamped at −60 mV. We found that estimates of the charge injected by a single granule cell EPSC were not very reliable because of the small amplitude and the relatively large noise associated with them. Thus, to more accurately estimate the charge injected by a parallel fiber input the stimulus intensity was increased to activate several granule cells (Fig. S2*B*) and the relationship between the total charge injected and peak EPSC amplitude established. This was used in combination with a previously obtained estimate of a single parallel fiber EPSC amplitude (≈11.7 pA) to estimate that a single parallel fiber input injects ≈0.12 pC of charge. Thus, with inhibition blocked, and assuming that each input releases neurotransmitter only once, each pattern is made of 250 parallel fiber inputs (30 pC/0.12 pC; Fig. S2*C*, black symbols), whereas with intact inhibition, the pattern size is 650 (78 pC/0.12 pC).

In response to a discrete sensory input in vivo (9), and with our method of granule cell stimulation (photorelease of glutamate onto a patch of granule cells), each parallel fiber is likely to release neurotransmitter more than once. Because parallel fibers show a significant paired-pulse facilitation under physiological conditions (10) repeated release of neurotransmitter from the same parallel fiber input will reduce the number of parallel fibers that form a pattern. To adjust the number of inputs constituting a pattern accordingly, we measured the extent of facilitation that is likely to occur under our experimental conditions.

In response to a discrete sensory input in vivo granule cells fire a short burst of ≈3 action potentials at an average rate of 75 Hz (9). When a patch of granule cells was electrically stimulated 3 times at 75 Hz (see Fig. S2*D*), the time course of the first 2 EPSCs resembled that obtained from the asynchronous activation of granule cells by photorelease of glutamate (see Fig. 1 in the main text). Because of the strong paired-pulse facilitation present at this synapse, the second EPSC injected 2.50 ± 0.2 times more charge ($n = 9$) than the first EPSC. Assuming that each input can release neurotransmitter twice, and taking into consideration this paired pulse facilitation, 70 (Fig. S2*C*, red symbols) or 185 parallel fiber inputs are needed to form a pattern in the absence and presence of inhibitory synaptic transmission, respectively.

Versatility of the Linear Algorithm in Pattern Recognition. As discussed in the main text, the utility of the linear computational algorithm in pattern recognition by Purkinje cells was demonstrated under physiological conditions where inhibitory synaptic transmission was intact and parallel fiber inputs were activated asynchronously. Additional experiments were done to evaluate the efficacy of this encoding mechanism under a variety of different conditions. The reproducibility with which the same input was encoded was estimated from the resulting standard deviations. The experimentally determined standard deviations in each condition were incorporated with the artificial neural network to estimate the capacity of Purkinje cells in recognizing patterns.

The data shown in Fig. S3 correspond to a condition where the pattern recognition capacity of Purkinje cells was evaluated for synchronous inputs by electrically activating granule cells in the presence of blockers of inhibitory synaptic transmission.

Averaging and the Implications of Correlated Noise for Pattern Recognition Capacity. As first noted by Eccles (11), the convergence of many Purkinje cells onto neurons of the deep cerebellar nuclei (DCN) is a feature of the cerebellar circuitry that allows, in principle, for the cerebellum to use averaging to reduce noise. However, averaging only reduces noise that is not correlated. The extent to which cerebellar averaging reduces noise and thus increases the s/n of pattern recognition would be less than optimal if part of the noise in the firing rate of individual Purkinje cells is correlated.

Let us consider the situation in which X Purkinje cells have learned the same Y patterns. When a specific learned pattern is presented to these X Purkinje cells there are several sources of noise that will cause variability in their response and thus to their ability to correctly recognize that specific pattern as being learned. These sources are both intrinsic and extrinsic to Purkinje cells. Given that, in the absence of any synaptic input, Purkinje cells fire action potentials at a rate of ≈50 spikes per second, intrinsic sources include factors that govern their pace

making and excitability (such as the state of their numerous voltage-gated ion channels). Extrinsic sources include differences in the timing of inhibitory synaptic inputs from numerous spontaneously active interneurons presynaptic to each Purkinje cell and the fact that the release probability of individual parallel fiber inputs is significantly <1. Given the nature of these noise sources, it is very unlikely that the noise contributed by them to each of the X Purkinje cells is correlated. Thus, averaging would reduce the noise coming from these sources by a factor of $\sqrt{X}$. The consequence of this reduction in noise is that by the virtue of averaging the cerebellum recognizes a specific pattern as learned with a higher signal to noise ratio than an individual Purkinje cell.

It is worthwhile to consider the efficacy of the averaging mechanism described above for recognizing a specific pattern if some of the noise among different Purkinje cells is correlated. Examples of such correlated noise sources might be Purkinje cell–Purkinje cell collaterals, electrical coupling between Purkinje cells or common inputs from interneurons. Although there is currently no evidence to suggest that these sources result in significant correlated noise among Purkinje cells that converge onto a common target neuron, we nonetheless evaluated the consequences of the presence of correlated noise on the efficacy of averaging in improving pattern recognition. Fig. S4 shows how the resulting s/n depend on the extent to which noise is correlated in Purkinje cells (0, 25%, 50%, and 100% noise correlation). As should be immediately obvious, the only instance in which averaging fails to improve the signal to noise ratio is when noise is 100% correlated. It should be noted however, that the noise contributed by the pattern recognition scheme is unique to each neuron and, thus, it follows that under these conditions, noise will never be 100% correlated.

The suggestion that averaging can be used to minimize noise contributed by the cerebellum is consistent with a recent study that examined the trial by trial correlation between the firing of individual Purkinje cells and eye movements in awake rhesus monkeys (12). It was concluded that the majority of the variability in eye position was accounted for by the variability in the visual motion signals to the cerebellum and that the cerebellum and other downstream structures contribute little additional noise. It is likely that averaging is implemented to minimize the noise added by the cerebellum during pattern recognition, although in this case Purkinje cells converge onto neurons of the vestibular nuclei rather than the DCN.

Medina and Lisberger (12) also reported that for different Purkinje cells involved in the performance of the same motor task, as much as 50% of the variability in their firing rate was correlated. It is important to note that this correlation simply reflects the fact that these Purkinje cells receive common inputs and are similarly driven. Thus, the commonality of inputs underlying this measured correlation is not analogous to the noise discussed above for the decoding of a specific pattern, but from trial to trial is instead reflective of various patterns (signals) being presented to the different cells.

1. Steuber V, De Schutter E (2001) Steuber V, De Schutter E (2001) Long-term depression and recognition of parallel fibre patterns in a multi-compartmental model of a cerebellar Purkinje cell. *Neurocomputing* 38:383–388.
2. Steuber V, *et al.* (2007) Cerebellar LTD and pattern recognition by Purkinje cells. *Neuron* 54:121–136.
3. Walter JT, Khodakhah K (2006) The linear computational algorithm of cerebellar Purkinje cells. *J Neurosci* 26:12861–12872.
4. Ito M, Kano M (1982) Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neurosci Lett* 33:253–258.
5. Ito M, Sakurai M, Tongroach P (1982) Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *J Physiol* 324:113–134.
6. Ito M (2001) Cerebellar long-term depression: Characterization, signal transduction, and functional roles. *Physiol Rev* 81:1143–1195.
7. Dayan P, Willshaw DJ (1991) Optimising synaptic learning rules in linear associative memories. *Biol Cybern* 65:253–265.
8. Graham B (2001) Pattern recognition in a compartmental model of a CA1 pyramidal neuron. *Network* 12:473–492.
9. Chadderton, P., Margrie, T. W. & Hausser, M. (2004) Integration of quanta in cerebellar granule cells during sensory processing *Nature* 428:856–860.
10. Foster KA, Crowley JJ, Regehr WG (2005) The influence of multivesicular release and postsynaptic receptor saturation on transmission at granule cell to Purkinje cell synapses *J Neurosci* 25:11655–11665.
11. Eccles JC (1973) The cerebellum as a computer: Patterns in space and time. *J Physiol* 229:1–32.
12. Medina JF, Lisberger SG (2007) Variation, signal, and noise in cerebellar sensory-motor processing for smooth-pursuit eye movements. *J Neurosci* 27:6832–6842.

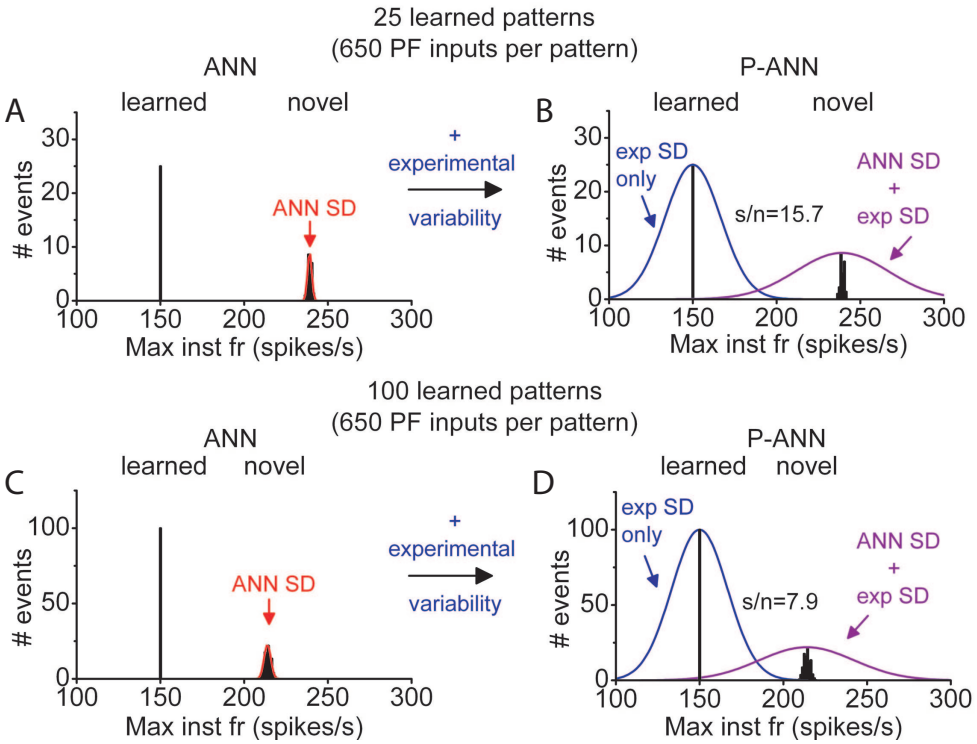

Fig. S1. Estimation of pattern recognition capacity of Purkinje cells using the implemented P-ANN. (*A*) The ANN was programmed to simulate a case where each pattern contained 650 PF inputs. Before learning, each pattern increased the firing rate to 250 spikes per second. The ANN was then simulated to recognize 25 patterns. As a consequence of LTD, each learned pattern was weakened such that it now increased the firing rate to only (but exactly) 150 spikes per second. Because some of the PFs that comprised the learned patterns were, at random, also assigned to some of the novel patterns, activation of novel patterns no longer increased the firing rate to 250 spikes per second but had a distribution with a slightly lower mean. (*B*) The experimentally determined response variability of Purkinje cells in cerebellar slices to repeated delivery of the same stimuli was incorporated with that of the ANN to determine the s/n of pattern recognition of the resulting P-ANN. (*C* and *D*) Same as above but with the system learning 100 patterns instead of 25.

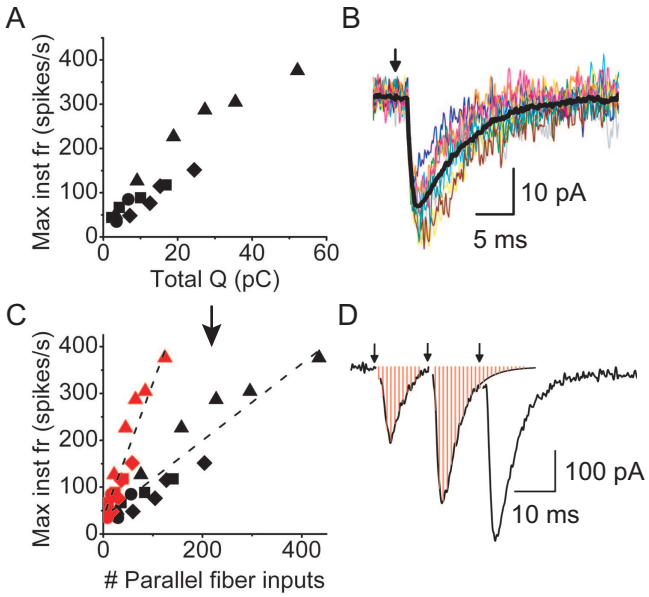

Fig. S2. Estimating the number of parallel fiber inputs that form a pattern. (*A*) The relationship between maximum firing rate of a Purkinje cell and the total charge injected by asynchronous activation of parallel fibers. (*B*) EPSCs recorded in a voltage-clamped Purkinje cell in response to repeated electrical activation of several parallel fibers. (*C*) Transformation of the data presented in *A* to correspond with the number of parallel fiber inputs needed to increase the firing rate of a Purkinje cell when each input releases neurotransmitter only once (black symbols) and when each input releases neurotransmitter twice (red symbols). (*D*) The impact of paired-pulse facilitation on the charge injected into Purkinje cells by parallel fibers.

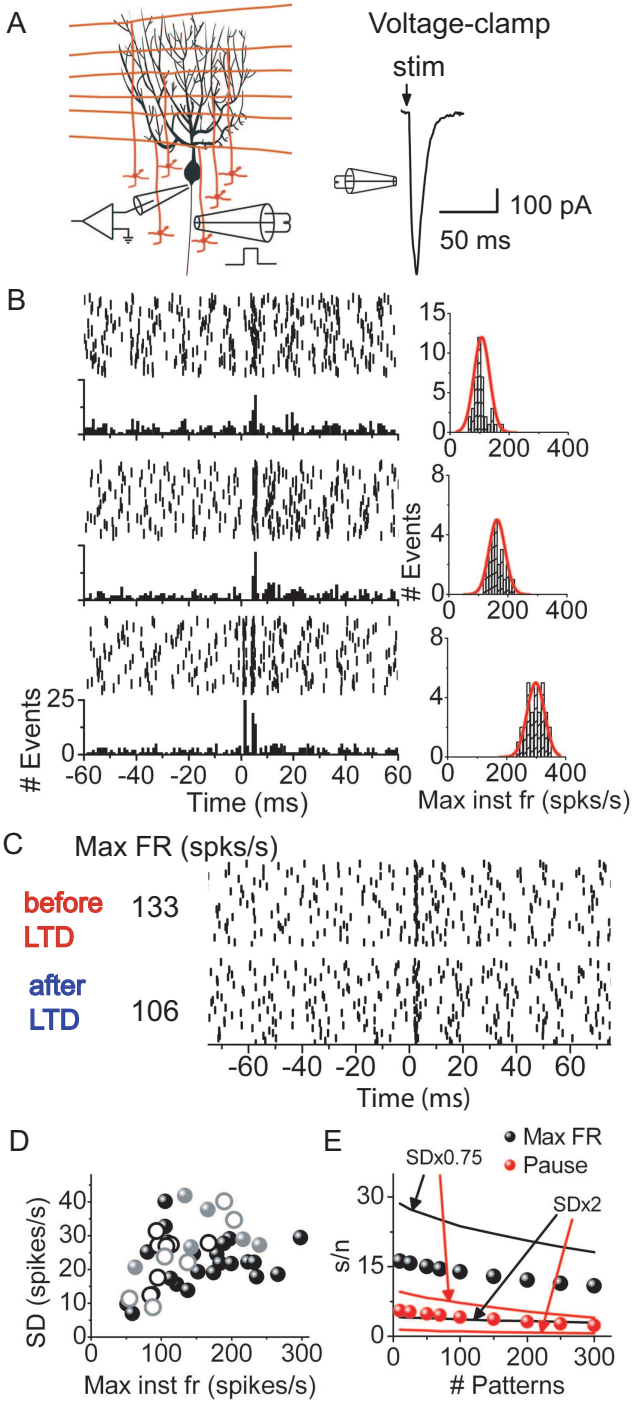

Fig. S3. Pattern recognition capacity of Purkinje cells with synchronous inputs. (*A*) In voltage-clamped Purkinje cells, synchronous activation of granule cells with a 200-μs electrical pulse resulted in EPSCs that had fast kinetics. (*B*) Raster plots of the response of a Purkinje cell to 50 repeated presentations of the same electrical stimulus. Vertical bars indicate the time of occurrence of each action potential. Below each raster plot, the associated population histogram is shown. The histograms to the right of each raster plot show the resulting maximum instantaneous firing rate distribution after stimulus. Each distribution was fit well by a Gaussian function (red line). (*C*) Raster plot of an experiment demonstrating a reduction in the maximum firing rate response following induction of LTD. LTD was produced using a standard protocol (5-min train of 1-Hz stimulation with parallel fiber stimulation preceding that of climbing fiber by 50 ms). (*D*) The scatter plot of the standard deviations of the maximum firing rates after stimulus determined from a number of experiments similar to that described above. Symbols refer to experiments where the standard deviation was estimated before (solid symbols) and after (open symbols) induction of LTD. (*E*) The experimentally determined standard deviations were implemented in the P-ANN to estimate the s/n in distinguishing learned patterns from novel ones as the number of patterns to be learned was increased (black symbols). The impact of altering the Purkinje cell response variability on their pattern recognition capacity was estimated by increasing and decreasing the experimentally determined standard deviations.

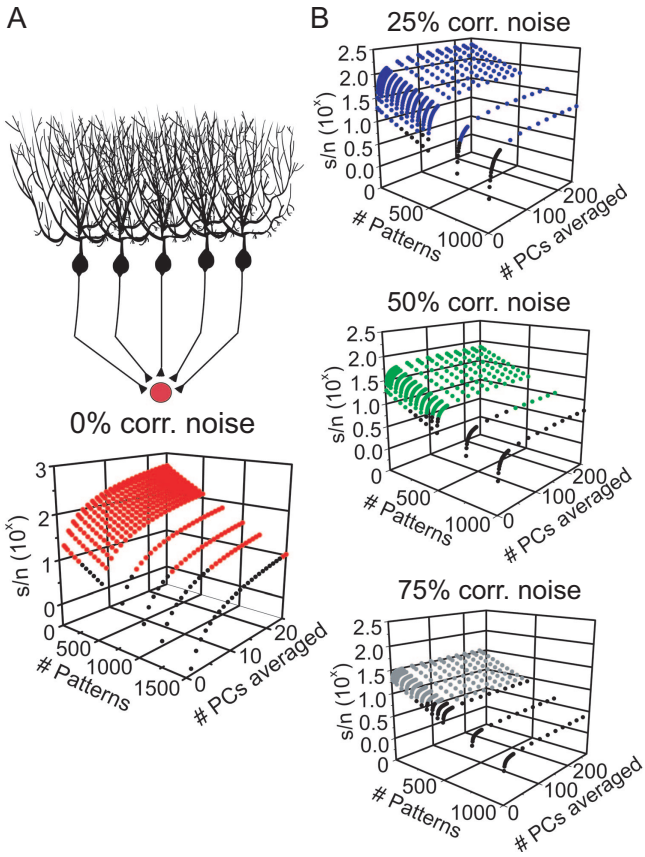

Fig. S4. Impact of correlated noise on averaging. (*A*) Purkinje cells carrying the same information were assumed to converge onto a single output neuron. The improvement in s/n for pattern recognition as a consequence of averaging is shown as the number of Purkinje cells that converged onto the output neuron was increased. In this instance, it was assumed that the noise in each cell was not correlated to any other (0% Corre. Noise). (*B*) The same as that in *A* expect with different extents of correlated noise (from 25% to 75%).