

Neuron

Optimal Degrees of Synaptic Connectivity

Highlights

- Sparse synaptic wiring can optimize a neural representation for associative learning
- Maximizing dimension predicts the degree of connectivity for cerebellum-like circuits
- Supervised plasticity of input connections is needed to exploit dense wiring
- Performance of a Hebbian readout neuron is formally related to dimension

Authors

Ashok Litwin-Kumar,
Kameron Decker Harris, Richard Axel,
Haim Sompolinsky, L.F. Abbott

Correspondence

ak3625@columbia.edu

In Brief

The number of inputs a neuron receives varies widely across different neural circuits. Using analytic calculations, Litwin-Kumar et al. relate this degree of connectivity to the ability of neuronal populations to produce high-dimensional stimulus representations and support associative learning.



Optimal Degrees of Synaptic Connectivity

Ashok Litwin-Kumar,^{1,9,*} Kameron Decker Harris,² Richard Axel,^{1,3,4} Haim Sompolinsky,^{5,6,7} and L.F. Abbott^{1,8}

¹Mortimer B. Zuckerman Mind Brain Behavior Institute, Department of Neuroscience, Columbia University, New York, NY 10027, USA

²Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

³Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA

⁴Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA

⁵Edmond and Lily Safra Center for Brain Sciences, Hebrew University, Jerusalem 91904, Israel

⁶Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel

⁷Center for Brain Science, Harvard University, Cambridge, MA 02138, USA

⁸Department of Physiology and Cellular Biophysics, Columbia University, New York, NY 10032, USA

⁹Lead Contact

*Correspondence: ak3625@columbia.edu

<http://dx.doi.org/10.1016/j.neuron.2017.01.030>

SUMMARY

Synaptic connectivity varies widely across neuronal types. Cerebellar granule cells receive five orders of magnitude fewer inputs than the Purkinje cells they innervate, and cerebellum-like circuits, including the insect mushroom body, also exhibit large divergences in connectivity. In contrast, the number of inputs per neuron in cerebral cortex is more uniform and large. We investigate how the dimension of a representation formed by a population of neurons depends on how many inputs each neuron receives and what this implies for learning associations. Our theory predicts that the dimensions of the cerebellar granule-cell and *Drosophila* Kenyon-cell representations are maximized at degrees of synaptic connectivity that match those observed anatomically, showing that sparse connectivity is sometimes superior to dense connectivity. When input synapses are subject to supervised plasticity, however, dense wiring becomes advantageous, suggesting that the type of plasticity exhibited by a set of synapses is a major determinant of connection density.

INTRODUCTION

Extensive synaptic connectivity is often cited as a key element of neural computation, a prime example being the cerebral cortex, where neurons each receive thousands of inputs. However, the majority of neurons in the human brain, the 50 billion neurons that form the cerebellar granule-cell layer, each receive input from only about four of the mossy fibers innervating the cerebellum (Eccles et al., 1966; Llinás et al., 2003). Sparse input is a feature shared by neurons that play roles analogous to granule cells in other neural circuits with cerebellum-like structures, such as the dorsal cochlear nucleus, the electrosensory lobe of electric fish, and the insect mushroom body (Mugnaini et al., 1980; Bell et al., 2008; Keene and Waddell, 2007). What is the functional role of diversity in synaptic connectivity, and what deter-

mines the appropriate number of input connections to a given neuronal type? In this study, we address these questions by investigating the ability of populations of neurons with different degrees of connectivity to support associative learning.

Both cerebellar and cerebrocortical regions are involved in a variety of experience-dependent adaptive behaviors (Raymond et al., 1996; Buonomano and Merzenich, 1998). In cerebellar cortex and other cerebellum-like circuits, synaptic modifications associated with learning occur among the elaborate dendrites of densely connected output neurons—for example, cerebellar Purkinje cells (Ito et al., 1982) and the output neurons of the *Drosophila* mushroom body (Hige et al., 2015). Classic Marr-Albus theories of associative learning propose that the abundance of granule cells supports a high-dimensional representation of the information conveyed to the cerebellum by mossy fibers and that the large number of synapses received by Purkinje cells allows them access to this representation to form associations (Marr, 1969; Albus, 1971). These theories assume that the inputs to granule cells are random and are not modified during learning. Anatomical and physiological studies suggest that the handful of inputs received by granule cells in the electrosensory lobe of electric fish (Kennedy et al., 2014) and by Kenyon cells, the granule-cell analogs of the mushroom body, are a random subset of the afferents to these structures (Murthy et al., 2008; Caron et al., 2013; Gruntman and Turner, 2013). In many regions of cerebellar cortex, granule cells receive diverse (Huang et al., 2013; Chabrol et al., 2015; Ishikawa et al., 2015; but see Jörntell and Ekerot, 2006; Bengtsson and Jörntell, 2009; and Discussion), though not completely random (Billings et al., 2014), mossy-fiber input. In Marr-Albus theories, learning in cerebellar cortex relies exclusively on climbing-fiber-dependent modifications of the connections between parallel fibers and Purkinje cells, but unsupervised forms of plasticity have been reported for synapses from mossy fibers onto granule cells (Hansel et al., 2001; Schweighofer et al., 2001; Gao et al., 2012, 2016; D'Angelo, 2014; but see Rylkova et al., 2015).

The logic of experience-dependent circuit modifications is less clear in cerebral cortex, where densely connected neurons exhibit diverse forms of synaptic plasticity (Abbott and Nelson, 2000). Recent theoretical studies have proposed that populations of randomly connected cerebrocortical neurons support high-dimensional representations that enhance the ability of

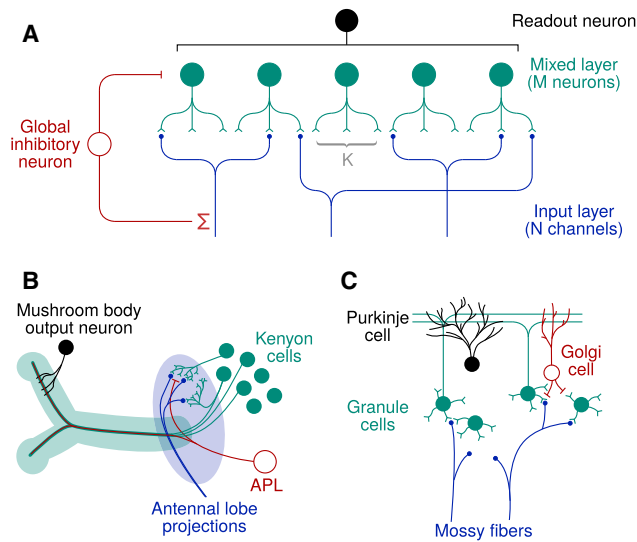


Figure 1. Schematic of Random Expansion

(A) N input channels project to a mixed layer of M neurons. Each mixed-layer neuron receives K connections from a random subset of input channels and inhibition via a globally connected inhibitory neuron.

(B) Schematic of *Drosophila* mushroom body. Kenyon cells (the mixed-layer neurons) receive inputs from antennal lobe projection neurons in the calyx (shaded blue region) and send bifurcated parallel-fiber axons that form the mushroom body peduncle and lobes (shaded green region). Mushroom body output neurons read out the activity of parallel fibers. Kenyon cells excite and are inhibited by the anterior paired lateral (APL) neuron.

(C) Schematic of cerebellar cortex. Mossy fibers project to the granule-cell (mixed) layer. Granule cells send parallel-fiber axons that are read out by Purkinje cells. Granule cells are inhibited by Golgi cells, which are excited by mossy and parallel fibers.

readout neurons to learn associations (Hansel and van Vreeswijk, 2012; Rigotti et al., 2013; Barak et al., 2013; Babadi and Sompolinsky, 2014), much as in theories of cerebellar cortex. Although these studies support the idea of random, high-dimensional representations as substrates for associative learning, they do not explain why the degree of synaptic connectivity in granule-cell and cerebrocortical layers is so different.

To address this issue, we explore the effects of degree of connectivity, balance of excitation and inhibition, and synaptic weight distribution on the ability of a large neural representation to support associative learning. We also investigate whether synaptic plasticity of input connections (e.g., mossy fiber to granule cell), augmenting plasticity of output connections (e.g., granule cell to Purkinje cell, as in Marr-Albus theories), improves performance. In these analyses, we distinguish between unsupervised synaptic plasticity that normalizes or otherwise modulates the gain of synaptic input without the aid of a feedback (“error”) signal and supervised synaptic plasticity that exploits feedback signals to reshape the neural representation based on prior experience.

Using a combination of analytic calculation and computer simulation, we find that the number of connections per neuron required to produce a high-dimensional representation increases slowly with the number of neurons. For a wide range of conditions (but not all), dimension and learning performance are maximized

if the number of inputs is small. These results apply when the input synapses onto the granule or granule-like cells are selected randomly or approximately randomly and are not modified by a supervised learning procedure. In contrast, if we permit supervised modification of these synapses during the learning of a task, dense connectivity becomes advantageous. Our theory predicts a degree of connectivity that maximizes dimension for cerebellum-like structures and quantitatively matches what is observed for cerebellar granule cells and Kenyon cells of the *Drosophila* mushroom body. It also predicts that supervised synaptic plasticity at multiple stages of processing is necessary to exploit the dense connectivity of cerebral cortex.

RESULTS

The “granule-cell” layer we consider consists of M neurons that each receive K excitatory connections from a random subset of N input channels and feedforward inhibition via a globally connected inhibitory neuron (Figure 1A). We assume that, as in the cerebellum and its analogs, $M > N$. We refer to K as the synaptic degree. Because the neurons in the second layer are selective to combinations of their inputs, we refer to the two layers as the input (e.g., mossy-fiber) and mixed (e.g., granule-cell) layers, respectively.

The ratio of the number of mixed-layer neurons to inputs, M/N , which we call the expansion ratio, is a critical parameter for our study. Our analysis assumes that the input channels are independent, so multiple channels with redundant activity are classified as a single input. This means that, for the *Drosophila* mushroom body, the number of distinct inputs is $N = 50$, an estimate of the number of antennal lobe glomeruli, while the estimated number of Kenyon cells is $M = 2,000$ (Keene and Waddell, 2007), yielding an expansion ratio of 40 (Figure 1B). For the cerebellum, the number of mossy fibers and granule cells presynaptic to a single Purkinje cell in the rat are estimated to be $N = 7,000$ and $M = 209,000$ (Marr, 1969; Harvey and Napper, 1991; Tyrrell and Willshaw, 1992; see STAR Methods), meaning the expansion ratio in this case is 30 if the mossy fibers are assumed to be independent (Figure 1C).

Heterogeneous Responses via Random Connectivity

A necessary requirement to produce a high-dimensional representation in the mixed layer is that its neurons should respond to different ensembles of inputs. To understand what is needed to produce this response heterogeneity, we first ask the simple question: what values of K ensure that, with high probability, each mixed-layer neuron receives a distinct subset of inputs? This question is a variant of the “birthday problem,” which concerns the likelihood of M people being born on unique days of the year. More generally, one can ask: what is the probability p of M draws (with replacement) from R equally likely possibilities all being different? In our case, the number of possibilities is $R = N$ choose K , the number of ways a mixed-layer neuron can choose K presynaptic partners from N inputs. The maximum value of this probability is always at $K = N/2$, but this maximum typically lies in the middle of a large range of K for which p is extremely close to 1. We therefore denote by K^* the smallest value of K for which p attains at least 95% of its maximum.

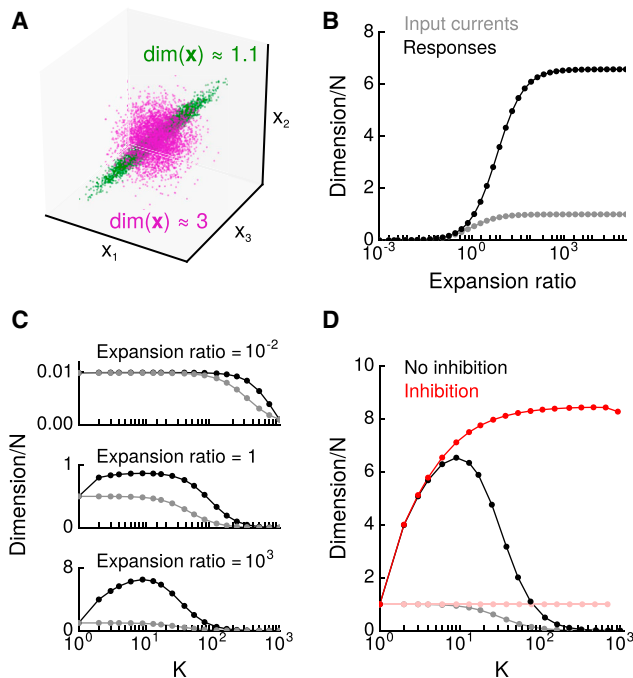


Figure 2. Dimension of Mixed-Layer Representation

(A) Illustration of dimension for a system with three degrees of freedom: x_1 , x_2 , and x_3 . Each point represents the response of the system to an input. The dataset consisting of the magenta points has a higher dimension than that of the green points.

(B) The input-current dimension $\dim(\mathbf{h})$ and mixed-layer dimension $\dim(\mathbf{m})$, normalized by N , are plotted as a function of the expansion ratio M/N (light curve and dark curve, respectively). For each M , dimension is evaluated at the value of K that maximizes it. Results are shown for $N = 1,000$, $f = 0.1$, and homogeneous excitatory synaptic weights.

(C) Dependence of dimension on K .

(D) Dependence of dimension on K for networks with and without inhibition in the limit of a large expansion ratio.

The level of 95% is, of course, arbitrary, but the value of K^* typically changes very little, if at all, if p is varied over a reasonable range near 100% (see below).

Using numbers appropriate for the mushroom body, $N = 50$ and $M = 2,000$, leads to $K^* = 7$, equal to the average observed number of projection-neuron inputs received by Kenyon cells (Caron et al., 2013; p takes the values 0.88, 0.98, and 0.996 for $K = 6, 7$, and 8). Using the same criterion for the cerebellum ($N = 7,000$ and $M = 209,000$) leads to $K^* = 4$, equal to the typical number observed anatomically (Eccles et al., 1966; p takes values of 0.69, 0.9998, and greater than 0.9999 for $K = 3, 4$, and 5). That a small number of connections is sufficient to ensure each mixed-layer neuron receives a unique set of inputs owes to the rapid growth of N choose K with K . In other words, the combinatorial explosion in the number of possible wirings with synaptic degree permits even very sparsely connected systems to eliminate duplication in the mixed layer with high probability.

Dimension of Mixed-Layer Inputs and Responses

The combinatorial calculation in the previous section treated two mixed-layer neurons as distinct even if only one of their inputs

differed, but partially overlapping inputs may introduce correlations between neural responses that reduce the quality of the mixed-layer representation. To provide a more nuanced analysis, we define and calculate a quantitative measure that characterizes the mixed-layer representation: its dimension. We define the dimension of a system with M degrees of freedom, $\mathbf{x} = (x_1, x_2, \dots, x_M)$, as

$$\dim(\mathbf{x}) = \frac{\left(\sum_{i=1}^M \lambda_i\right)^2}{\sum_{i=1}^M \lambda_i^2}, \quad (\text{Equation 1})$$

where λ_i are the eigenvalues of the covariance matrix of \mathbf{x} computed by averaging over the distribution of inputs to the system (Abbott et al., 2011). If the components of \mathbf{x} are independent and have the same variance, all the eigenvalues are equal and $\dim(\mathbf{x}) = M$. Conversely, if the components are correlated so that the data points are distributed equally in each dimension of an m -dimensional subspace of the full M -dimensional space, only m eigenvalues will be nonzero and $\dim(\mathbf{x}) = m$ (Figure 2A). Later, we will show that this measure is closely related to the classification performance of a readout of mixed-layer responses.

Increased dimension of the mixed-layer representation requires, in addition to mixing of the inputs received by each neuron, a nonlinearity provided by their input-output response function. We begin by analyzing the effect of input mixing alone and then address the effect of nonlinearity for the case of purely excitatory input (later we will add inhibition). In studying the effect of mixing alone, the components of the vector we consider, which we call \mathbf{h} , are the total synaptic currents received by each of the mixed-layer neurons. These currents are given by $\mathbf{h} = \mathbf{J}\mathbf{s}$, where \mathbf{J} is the $M \times N$ matrix of synaptic weights describing the strengths of the connections from the input layer to the mixed layer and \mathbf{s} is the vector of activities for the input layer. We consider the case of uncorrelated inputs with a uniform variance across them. If the excitatory connections onto the mixed-layer neurons have homogeneous weights,

$$\dim(\mathbf{h}) \approx \frac{N}{1 + N/M + (K-1)^2/N}, \quad (\text{Equation 2})$$

when M and N are large. Two features of this result are noteworthy. First, $\dim(\mathbf{h}) \leq N$ even when $M \gg N$ (Figure 2B, light curve). This is because \mathbf{h} is a linear function of the activity of the input layer \mathbf{s} , so it cannot have dimension higher than N . Second, increasing the synaptic degree K reduces the dimension (Figure 2C, light curve). This is due to the final term in the denominator, which arises from correlations between the entries of \mathbf{h} . On average, two mixed-layer neurons share K^2/N of their K inputs, leading to an average correlation of K^2/N . Indeed, as K approaches N , all the currents received by mixed-layer neurons become identical, and $\dim(\mathbf{h})$ approaches 1. Thus, unlike the combinatorial calculation of the previous section, this analysis indicates that increasing K beyond 1 is detrimental because it reduces the dimension due to increased correlations even when the identities of the inputs to each neuron are different.

However, our real interest is the dimension of the nonlinear output of the mixed layer. We consider mixed-layer neurons with binary outputs given by $\mathbf{m} = \Theta(\mathbf{h} - \boldsymbol{\theta})$, where Θ is a step function (applied element-wise) and $\boldsymbol{\theta}$ is a vector of activity thresholds, one for each neuron. The thresholds are chosen so that each neuron is active with probability f , averaged across all the input patterns which, for now, we take to be standard Gaussians (we refer to the input-layer response to a particular stimulus as an input pattern). We call f the coding level of the mixed layer. Under these assumptions, the mixed-layer dimension is given by

$$\dim(\mathbf{m}) \approx \frac{1}{\frac{1}{M} + \langle \rho_{ij} \rangle^2 + \text{Var}(\rho_{ij})}, \quad (\text{Equation 3})$$

where ρ_{ij} is the correlation coefficient of the activities m_i and m_j of neurons i and j averaged over the distribution of input patterns (a more general expression holds when the coding levels of the neurons are not identical; see [STAR Methods](#)). Thus, the maximum mixed-layer dimension is limited by correlations among its neurons and saturates to a limiting value as the expansion ratio grows ([Equation 3](#); [Figure 2B](#), dark curve; also see [Babadi and Sompolinsky, 2014](#)). The limiting value scales linearly with N (since $\text{Var}(\rho_{ij}) \sim 1/N$; [Figure S1](#)). For a coding level of $f = 0.1$, the saturation suggests that expansion ratios beyond 10–50 do not increase the mixed-layer dimension. This maximum expansion ratio and the maximum dimension both increase as the coding level is reduced ([Figure S1](#)), but as we will see, representations with extremely low coding levels do not necessarily lead to improved discrimination ([Barak et al., 2013](#); [Babadi and Sompolinsky, 2014](#)).

We next investigate the dependence of mixed-layer dimension on K . When M is small, the mixed-layer dimension is similar to the input-current dimension and nearly constant over a wide range of K values ([Figure 2C](#), top). However, for larger M , dimension initially grows with K , as increased K results in each mixed-layer neuron being selective to different combinations of inputs ([Figure 2C](#), middle and bottom). The dimension is maximized for an intermediate value of K that depends on N , f , and the distribution of synaptic weights. For the case of homogeneous excitatory synaptic weights, this value is $K = 9$ for $N = 1,000$ and $f = 0.1$. Above this value, dimension decreases because of positive average correlations among the mixed-layer neurons ([Figure 2C](#); [Figure S1](#)). Thus, the detrimental effect of even small average correlations ($\langle \rho_{ij} \rangle$ in the denominator of [Equation 3](#)) on dimension leads to dimension being maximized at small K .

Many studies have shown that inhibition can decorrelate neural activity ([Ecker et al., 2010](#); [Renart et al., 2010](#); [Wiechert et al., 2010](#)), so we next investigate whether inhibition can increase dimension by reducing correlations among mixed-layer neurons. In the *Drosophila* mushroom body, a single GABAergic interneuron on each side of the brain inhibits the Kenyon-cell population globally ([Liu and Davis, 2009](#); [Figure 1B](#)). In the cerebellum, Golgi cells are vastly outnumbered by the granule cells they inhibit ([Eccles et al., 1966](#); [Figure 1C](#)). We therefore begin by introducing a single neuron that inhibits all mixed-layer neurons in proportion to its input, which is equal to the summed input-layer activity ([Figure 1A](#)).

When inhibition is tuned to balance excitation on average (see [STAR Methods](#)), the distribution of the correlation between inputs received by pairs of mixed-layer neurons has zero mean, although the variance remains finite ([Figure S1](#)). Consequently, $\dim(\mathbf{h})$ does not decrease with K , and $\dim(\mathbf{m})$ does not exhibit a peak at small K ([Figure 2D](#)). However, even in this case, dimension quickly approaches its maximum (which occurs at $K = N/2$), and when $N = 1,000$ and the expansion ratio is large, it attains 95% of its maximum value at $K = 29$, or approximately 3% connectivity. Furthermore, the increase in dimension due to inhibition is only substantial for sufficiently large K .

Thus, nonlinear mixed-layer neurons with small synaptic degree are sufficient to produce high-dimensional representations. This observation is consistent with the combinatorial argument of the first section, which showed that the explosion in possible wirings with synaptic degree leads to few redundant mixed-layer neurons, even when the degree is small. The present analysis also shows that positive average correlations limit dimension when mixed-layer neurons receive purely excitatory input, and that when K is large global inhibition can increase dimension through decorrelation. Although our analytic calculations are most easily performed for systems with feedforward inhibition, we verified with simulations that our qualitative results also hold for sufficiently strong recurrent inhibition ([Figure S2](#)).

Optimal Connectivity for Random Representations under Resource Constraints

The observations of the previous section suggest that a representation formed by many neurons with small synaptic degree may be higher dimensional than one formed by fewer neurons with large synaptic degree. Therefore, when constructing a randomly wired neural system with limited resources, the former strategy may be preferable. To formalize this intuition, we ask: what combination of mixed-layer neuron number M and synaptic degree K maximizes the dimension of the mixed-layer representation when the total number of connections $S = MK$ is limited to some maximum value? This is equivalent to fixing the number of presynaptic sites to which mixed-layer neurons can connect.

For $f = 0.1$ and $S = 14,000$, consistent with the *Drosophila* mushroom body ([Keene and Waddell, 2007](#); [Caron et al., 2013](#)), the optimum occurs at $K = 4$ when inhibition is absent or $K = 8$ when it is present ([Figure 3A](#)), close to the observed value of 7. For $f = 0.01$ and $S = 8.4 \times 10^5$, parameters consistent with the cerebellar granule-cell representation ([Eccles et al., 1966](#); [Harvey and Napper, 1991](#)), the optimum occurs at $K = 4$ ([Figure 3B](#)), equal to the observed value. In this case, the presence of inhibition has a vanishing effect when K is small. This is because, for $N = 7,000$ and small K , the average overlap of inputs received by granule-cell pairs is negligible, and decorrelation via inhibition is not necessary. Although we chose the mixed-layer coding level based on available experimental estimates ([Chadderton et al., 2004](#); [Honegger et al., 2011](#)), our conclusions do not depend strongly on this parameter ([Figures 3C](#) and [3D](#)). For both the *Drosophila* mushroom body and the cerebellum, therefore, allocating synaptic resources among randomly connected neurons to maximize dimension quantitatively predicts the observed synaptic degree.

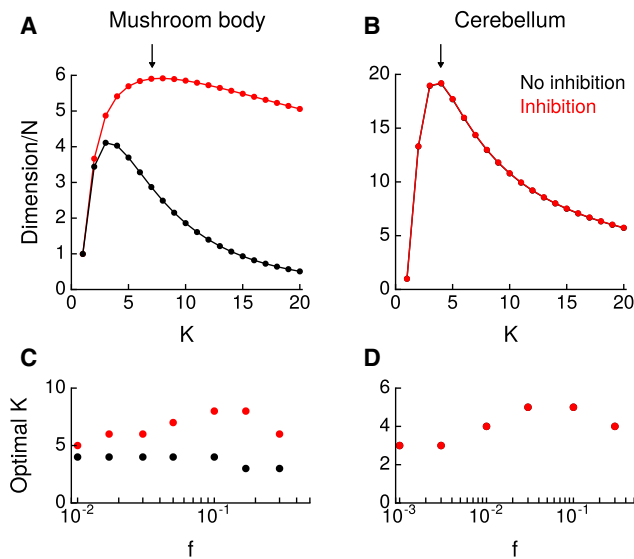


Figure 3. Optimal Connectivity with Constrained Connection Number

(A) For each K , $\dim(\mathbf{m})/N$ is plotted with the number of mixed-layer neurons set to $M = S/K$. S is equal to 14,000, an estimate of the total number of connections from projection neurons onto Kenyon cells in each hemisphere of the *Drosophila* mushroom body (Keene and Waddell, 2007; Caron et al., 2013). Other parameters are $N = 50$, an estimate of the number of antennal lobe glomeruli, and $f = 0.1$. Arrow indicates the average value of K observed anatomically.

(B) Same as (A), but for parameters consistent with the cerebellar granule-cell representation (Eccles et al., 1966; Harvey and Napper, 1991), $S = 8.4 \times 10^5$, $N = 7,000$, and $f = 0.01$.

(C) Optimal value of K for the *Drosophila* mushroom body representation, defined as the maximum of curves computed as in (A) for different mixed-layer coding levels.

(D) Same as (C), but for the cerebellar granule-cell representation. In (B) and (D), points for inhibition and no inhibition lie on top of each other when $K \ll N$.

Heterogeneous Synaptic Weights

EPSP amplitudes recorded *in vitro* are not homogeneous but broadly distributed (Sargent et al., 2005; Song et al., 2005). To examine if our results are valid in the presence of this heterogeneity, we ask: what changes when synaptic weights are drawn from a distribution with mean $\langle w \rangle$ and variance $\text{Var}(w)$? We first examine the dimension of the input currents. When $M \gg N \gg 1$, and in the absence of inhibition,

$$\dim(\mathbf{h}) \approx \frac{N}{1 + \frac{(K-1)^2}{N} \left(\frac{\langle w \rangle^2}{\langle w \rangle^2 + \text{Var}(w)} \right)^2} \quad (\text{Equation 4})$$

First, we observe that the input-current dimension still decreases as K increases (compare with Equation 2). However, this decrease occurs more slowly when $\text{Var}(w)$ is large. This makes sense because, with heterogeneous weights, even neurons with identical input ensembles can receive different input currents. Accordingly, when considering the nonlinear mixed-layer responses, the maximum dimension occurs for a larger value of $K = 19$ when synaptic weights are sampled from a log-normal distribution matching recorded EPSP amplitudes in

neocortical slices (Song et al., 2005) and inhibition is absent (Figure 4A). This is in contrast to $K = 9$ for the case of homogeneous weights (Figure 2D).

Next, we add global inhibition. We assume that the synaptic weights of the connections from the input layer onto both mixed-layer neurons and the global inhibitory neuron are heterogeneous, while inhibitory weights are homogeneous. In this case, the addition of one inhibitory neuron only modestly increases the mixed-layer dimension (Figure 4A). To understand why this improvement is weaker than for homogeneous weights (Figure 2D), we again examine the correlation of the input currents received by pairs of mixed-layer neurons. When the weights of the connections received by the inhibitory neuron from the input layer are chosen from a distribution with mean $\langle w_i \rangle$ and variance $\text{Var}(w_i)$, the average input-current correlation is proportional to

$$\frac{K}{N} \left(1 - \frac{\langle w_i \rangle^2}{\langle w_i \rangle^2 + \text{Var}(w_i)} \right), \quad (\text{Equation 5})$$

when inhibition is tuned to minimize correlation (see STAR Methods). Only when all the connections received by the inhibitory neuron have homogeneous weights (so that $\text{Var}(w_i) = 0$) can perfect decorrelation occur. This is because, when these weights are heterogeneous, inhibition serves as an extra source of correlated input, precluding full decorrelation.

However, this heterogeneity can be averaged out by the presence of multiple inhibitory neurons, leading to more effective decorrelation. When the number of inhibitory neurons is increased, the maximum dimension increases, and the maximum is located at a larger value of K (Figure 4B). Notably, the advantage of inhibition is restricted to large K , as in the case of homogeneous weights (Figure 2D).

There seems to be little benefit to the presence of inhibition when K is small and N is large, at least as far as dimensionality is concerned. However, inhibition may be useful for different purposes. For example, if inputs convey signals with different coding levels, it is difficult to set an activity threshold that assures that neurons respond to signals with a low coding level without saturating for signals with a high coding level (in other words, the dynamic range of the mixed layer is low). Global inhibition can mitigate this problem by increasing a neuron's effective threshold for dense signals (Marr, 1969; Albus, 1971; Billings et al., 2014). In our model, one global inhibitory neuron is sufficient to increase the dynamic range of the mixed layer, even if it sums its inputs through connections with heterogeneous weights (Figure 4C).

Inhibition may therefore serve two purposes in neural circuits implementing random expansions: normalization and decorrelation. The former can be accomplished with a single inhibitory neuron that provides global graded inhibition (Figure 4C), while the latter requires either precisely tuned inhibitory synaptic weights (Figure 2D) or sufficiently many inhibitory neurons to average heterogeneities in synaptic weights (Figure 4B). The benefits of decorrelation are more pronounced as synaptic degree grows, but even in this case, dimension increases at best logarithmically (Figure 4B). Thus, neural circuits containing many mixed-layer neurons with relatively small synaptic degrees and few global inhibitory neurons are capable of providing many of the benefits of a random expansion. Accordingly, even for

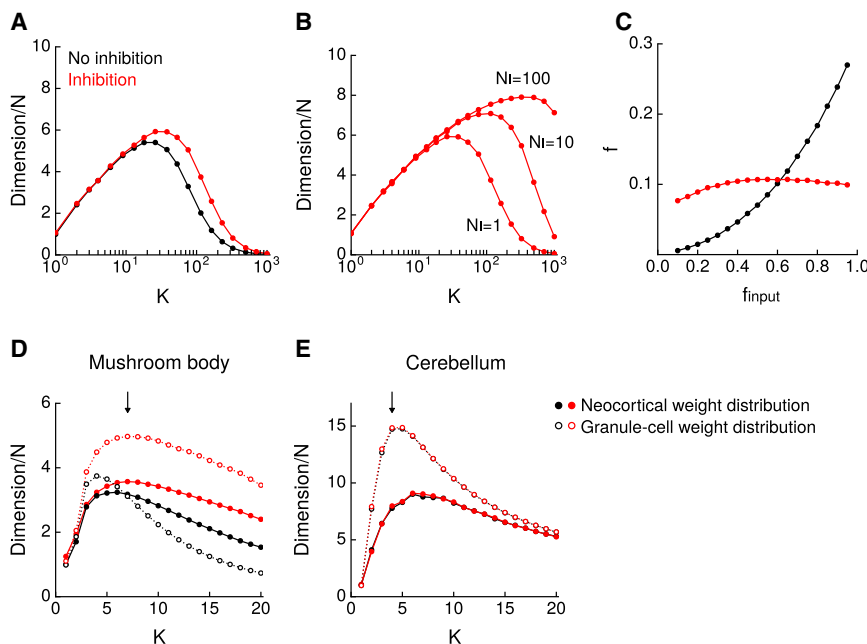


Figure 4. The Case of Heterogeneous Synaptic Weights

(A) Dependence of dimension on K for log-normal synaptic weights chosen from a distribution consistent with recordings from neocortical neurons (Song et al., 2005), with $N = 1,000$, $f = 0.1$, and in the limit of a large expansion ratio.

(B) Similar to (A) for different numbers of global inhibitory neurons N_i .

(C) Dependence of mixed-layer coding level f on input-layer coding level for $K = 10$. Input-layer activities were sampled from a standard Gaussian distribution truncated at zero with a coding level of f_{input} . Mixed-layer activity thresholds were chosen to obtain an average coding level of $f = 0.1$ across input patterns when f_{input} is drawn uniformly from the range (0,1) for each pattern.

(D) Dimension of the *Drosophila* mushroom body representation as in Figure 3A but for log-normal synaptic weights ($N_i = 1$ for the case of inhibition). Parameters of the log-normal distribution were chosen to match data from neocortical neurons (Song et al., 2005; filled circles) or cerebellar granule cells (Sargent et al., 2005; open circles).

(E) Same as (D), but for the cerebellar granule-cell representation. In this case, the presence of inhibition or the number of inhibitory neurons N_i does not affect the shape of the curves.

broadly distributed synaptic weights, the optimal value of K for the mushroom body and cerebellum calculated in the manner of Figure 3 is no more than 7 (Figures 4D and 4E). Notably, when the distribution of weights is chosen to match recordings from cerebellar granule cells (Figures 4D and 4E, open circles; Sargent et al., 2005), rather than recordings from neocortical neurons that predict a broader distribution (Figures 4D and 4E, filled circles; Song et al., 2005), dimension is increased, and the optimal synaptic degrees more closely match the values observed anatomically.

Local Connectivity in Cerebellar Cortex

The inputs to *Drosophila* mushroom-body Kenyon cells appear to be random (Caron et al., 2013), but the extent to which this is true for cerebellar cortex is unclear. We extended a recent model (Billings et al., 2014) to determine the dimension of the representation formed by cerebellar granule cells when their inputs are chosen in accordance with the constraints imposed by the spatial arrangement of granule cells and mossy fibers. In the model, synaptic connections are formed depending on the distance between granule cells and mossy-fiber rosettes (Figure 5A; see STAR Methods).

This connectivity rule leads to an increased proportion of nearby granule-cell pairs that have multiple shared mossy-fiber inputs (Figure 5B). Consistent with this increased redundancy, the dimension of the resulting representation decreases by approximately 50% compared to the case of random connectivity (Figure 5C), suggesting that spatially structured wiring limits dimension. Nonetheless, the optimal synaptic degree still occurs at $K = 4$. Thus, sparse connectivity can be sufficient to create high-dimensional representations even when dimension is limited by nonrandom wiring.

Classification of Mixed-Layer Responses

Dimensionality provides a convenient measure of the quality of a mixed-layer representation because it makes few assumptions about how the representation is used. A less abstract measure is the ability of a decoder to correctly classify input patterns on the basis of mixed-layer responses. Consider a set of P patterns of activity \mathbf{s}^μ for $\mu = 1 \dots P$ in the input layer, each of which is associated with a positive or negative valence, denoted by $v^\mu = \pm 1$, with equal probability. Patterns associated with positive or negative valences might correspond to neural activity evoked by appetitive or aversive stimuli, for example. We examine the performance of a fully connected Hebbian classifier of mixed-layer responses trained to discriminate between input patterns of opposing valences (Figure 6A; see STAR Methods). During the training phase, the classifier updates its synaptic weights according to the mixed-layer activity and the valence so that, after training, the weights are given by $\sum_{\mu=1}^P (\mathbf{m}^\mu - f) v^\mu$, where \mathbf{m}^μ is the mixed-layer activity pattern evoked by input pattern \mathbf{s}^μ (the subtraction is performed element-wise). To assess the performance of this classifier during the test phase, we compute its error rate when classifying the activity evoked by instances of previously learned input patterns that have been corrupted by noise. Input noise produces noise in the mixed layer, which we quantify by a measure $\Delta = d/2f(1-f)$, where d is the mean-squared distance between the mixed-layer activity during training and testing (see STAR Methods). If noise is not present, $\Delta = 0$, while if noise is so large that training and test patterns are uncorrelated, $\Delta = 1$.

For the case of uncorrelated input patterns and valences, mixed-layer neurons with a fixed coding level, and input-layer to mixed-layer connections that are not modified by supervised learning, there is a particularly simple relationship between the

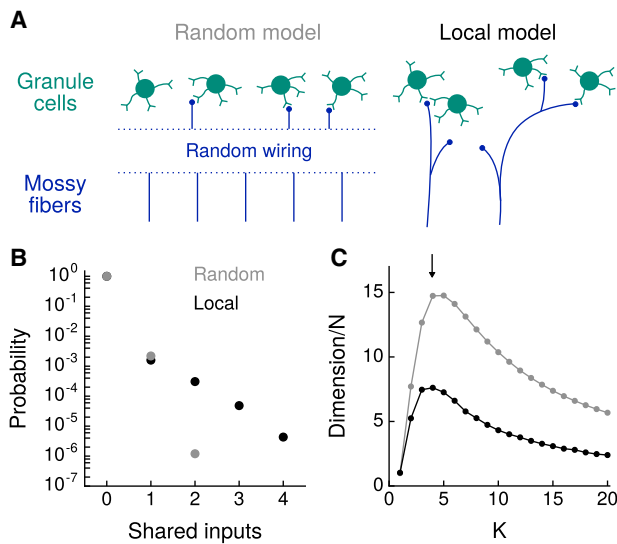


Figure 5. Spatial Model of Cerebellar Cortex

(A) Schematic of random model (left) and local model (right) of mossy-fiber to granule-cell connectivity based on Billings et al. (2014).

(B) Distribution of the number of shared mossy-fiber inputs received by pairs of granule cells when connections are chosen randomly (random model, gray) or based on the distance between mossy-fiber rosettes and granule cells in the local model (black).

(C) Dimension as a function of K as in Figure 3B for the two models. For the local model, the granule-cell density is scaled inversely with K to fix the total number of connections. Results are shown for excitatory synaptic weights with a distribution of strengths chosen to match recordings from cerebellar granule cells (Sargent et al., 2005), but the results do not change if global inhibition is added to the model (see Figure 4E).

performance of a Hebbian classifier and mixed-layer dimension (Equation 3). The error rate of the classifier is expressed in terms of a signal-to-noise ratio (SNR) for the mixed layer through the relation $P(\text{error}) = 0.5\text{erfc}(\sqrt{\text{SNR}/2})$ (Babadi and Sompolinsky, 2014). The SNR is proportional to the mixed-layer dimension (see STAR Methods) and given by

$$\text{SNR} = \frac{\text{dim}(\mathbf{m}) \cdot (1 - \Delta)^2}{P}, \quad (\text{Equation 6})$$

when P is large. Thus, SNR and error rate, which is a monotonically decreasing function of SNR, depend only on the mixed-layer dimension, the number of input patterns, and Δ . We have already explored the dependence of dimension on synaptic degree (Figures 2, 3, 4, and 5). All that remains is to understand the dependence of Δ , which is determined by the difference between the mixed-layer representation of an input and that same input corrupted by noise. For a fixed coding level and Gaussian noise in the input layer, the dependence is trivial; Δ is independent of synaptic degree (Figure 6B, left). Thus, the synaptic degree that maximizes dimension for a given coding level also minimizes error rate, regardless of the level of noise in the input layer (Figure 6C, left). We note that Δ is high when the coding level is low (Barak et al., 2013; Babadi and Sompolinsky, 2014; Figure S3), meaning that representations with a very low coding

level have a reduced SNR even though they have a higher dimension (Figure S1).

Gaussian input patterns and noise may apply when inputs received by mixed-layer neurons are communicated via smoothly varying firing rates. However, for neurons with small synaptic degree, a few EPSCs are often sufficient to generate a spiking response (Chadderton et al., 2004). In this case, it is more appropriate to describe the input patterns as discrete spike counts. While the qualitative dependence of mixed-layer dimension on synaptic degree is similar for discrete and Gaussian inputs (Figure S4), the dependence of Δ is not (Figure 6B, right). Surprisingly, sparse connectivity leads to lower Δ than dense connectivity in this case (except when K is very small; see Figure 6 caption). This further biases the optimal connectivity toward small synaptic degree (Figure 6C, right). Noise reduction at small K arises from the fact that, when coding level is low, many mixed-layer neurons are just slightly above or below threshold and thus sensitive to noise. When K is large, all neurons experience noise in a fraction of their inputs, while when K is small, there is a subpopulation of neurons whose input is not strongly affected by noise and that can respond reliably. The effect is maximal for low noise and high mixed-layer coding level (Figure S3). Thus, small synaptic degree can curtail the amplification of input noise if this noise arises from low-rate spiking activity.

For the above analyses, we assumed that only fluctuations in the input layer contribute to noise in the mixed layer. However, if synaptic release is unreliable and uncorrelated across release sites (Markram et al., 1997; Sargent et al., 2005), Δ can be reduced by increasing synaptic degree to average over these fluctuations (Figure S5). This averaging biases the optimal synaptic degree toward larger values. Cerebellar granule cells appear to be capable of responding reliably to mossy fiber activation (Chadderton et al., 2004), but quantifying the relative contributions of input-layer and synaptic fluctuations to noise in the mixed layer requires further analysis.

Learned Mixed-Layer Representations

We have shown that small synaptic degree is sufficient to obtain optimal or near-optimal classification performance when the connections and synaptic weights of the mixed-layer neurons are set randomly and the activities of input-layer and mixed-layer neurons are normalized (meaning that their variances and coding levels are homogeneous). What happens if synaptic weights can be modified by plasticity directed by a learning process? We considered the performance of a classifier after both unsupervised and supervised modifications of the connections received by mixed-layer neurons.

Unsupervised learning has been proposed to improve classification by regulating the activity of the cerebellar granule-cell layer (Schweighofer et al., 2001; D'Angelo, 2014). We hypothesized that such learning could be particularly important when the activities of different inputs are not normalized. In this case, synaptic plasticity could increase the influence of presynaptic inputs whose activity only weakly activates mixed-layer neurons, ensuring each mixed-layer neuron conveys information relevant to a task (Schweighofer et al., 2001). We implemented a plasticity rule that bidirectionally adjusts synaptic weights based on recent activity to compensate for changes in activity level as

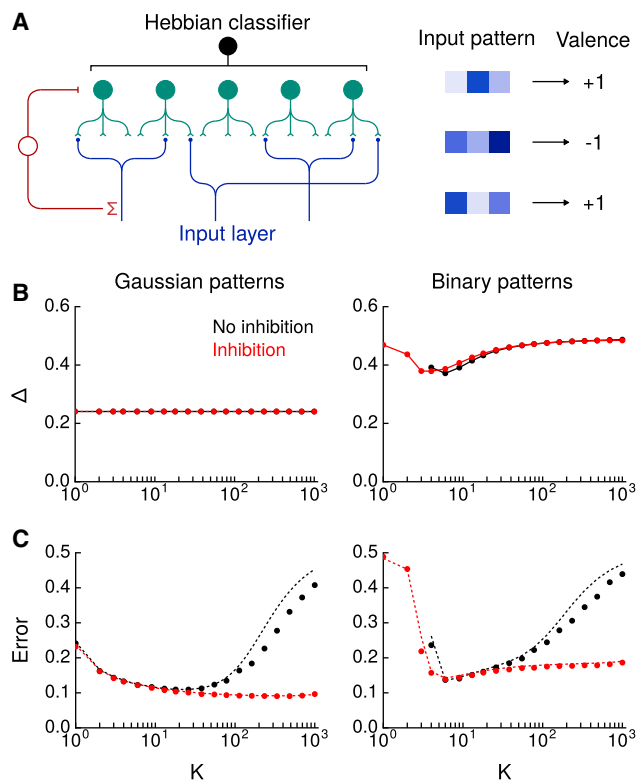


Figure 6. Classification of Mixed-Layer Responses

(A) Schematic of network with a classifier that computes a weighted sum of mixed-layer activity to determine the valence of an input pattern.

(B) Δ for $N = 1,000$ and an expansion ratio of $M/N = 5$ for log-normal synaptic weights (Song et al., 2005) without inhibition (black) and with $N_i = 100$ global inhibitory neurons (red). At left are Gaussian input patterns (curves with and without inhibition lie on top of each other); at right are binary input patterns with coding level of $f_{\text{input}} = 0.5$. For Gaussian patterns, the standard deviation of the noise at the input layer is 0.3 times that of the signal; for binary patterns, noise is introduced by randomly flipping the activity of 10% of the inputs. For binary patterns and no inhibition, the mixed-layer coding level cannot be adjusted to exactly match the desired level of $f = 0.1$ because the input to each mixed-layer neuron takes only discrete values. We therefore choose the threshold that leads to the highest $f \leq 0.1$. For $K = 1, 2, 3$, it is not possible to find such a threshold with $f > 0$, so data is not shown.

(C) Error rate of a Hebbian classifier when classifying $P = 1,000$ random input patterns. Circles denote simulations, and dashed lines denote prediction from Equation 6.

well as one that adjusts the excitability of mixed-layer neurons to maintain a desired coding level (see STAR Methods). After learning, the error rate of the classifier is reduced (Figure 7A). Notably, small synaptic degree is still sufficient to attain high performance.

We next asked if supervised learning could further reduce the error rate for a system in which inputs and mixed-layer responses have already been normalized. To implement learning, the strengths of connections received by mixed-layer neurons were modified with a Hebbian rule to associate each input pattern with a target pattern of activity in the mixed layer (see STAR Methods). In this scenario, the problem of learning is to determine the appropriate target patterns. Unlike the unsuper-

vised case, this form of learning requires a feedback signal to determine the target patterns.

A previous study proposed associating each input pattern with a random and uncorrelated target pattern (Babadi and Sompolinsky, 2014). This approach attempts to reduce the overlap of any two target patterns, which may otherwise be large if mixed-layer neurons are highly correlated. Such learning reduces the error rate, but only for large synaptic degree (Figure 7B, blue curve). The magnitude of the reduction is greater for very low coding levels (Figure S6), although it is still restricted to large synaptic degree.

Improved performance can be achieved if mixed-layer target patterns incorporate information about valence. When target patterns are selective for conjunctions of valence and input, classification performance improves dramatically for large synaptic degree (Figure 7B, cyan curve). This improvement demonstrates the utility of mixed representations that incorporate information about both input and desired outcome. Such representations can be generated in a biologically plausible manner by computing linear combinations of input and valence and applying a threshold (see STAR Methods).

The mixed-layer representation is not uncorrelated with valence or the input patterns presented during training in the case of supervised learning; therefore, the error rate is no longer simply related to dimension (Equation 6). In fact, the dimensions of the representations described above, computed over the distribution of all random input patterns, are lower than the dimension of a random representation (Figure 7C). This reflects a trade-off between minimizing classification error for known patterns and creating a high-dimensional representation of arbitrary novel patterns. Surprisingly, the benefits of these learned representations, when compared to random connectivity with Gaussian statistics, are only evident when $K > 10$, suggesting that when inputs to mixed-layer neurons are appropriately normalized and synaptic degree is small, implementing a supervised learning procedure may not improve performance on associative learning tasks.

DISCUSSION

Our analysis demonstrates that the sparse connectivity observed in the cerebellar granule-cell layer and analogous structures is well suited for producing high-dimensional representations that can be read out by densely connected output neurons. This conclusion is supported by simple combinatorial arguments, analysis of dimension (Figures 2, 3, 4, and 5), and the relationship between dimension and classification error (Figure 6C), all of which demonstrate that the performance of randomly wired neural systems quickly saturates as synaptic degree grows. While random wiring is unable to exploit synaptic degrees beyond this saturation point, supervised Hebbian modifications of input synapses can dramatically improve classification performance for large synaptic degrees (Figure 7B). Such modifications may therefore be present in cerebrocortical circuits, where input and output neurons are both densely connected.

Expansion via Sparse Connectivity in Neural Systems

Classic theories of cerebellar cortex were primarily interested in the mechanisms that maintain a low coding level (that is, sparse

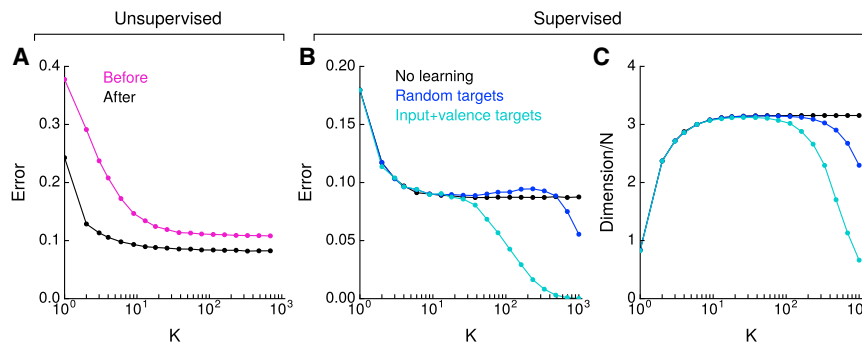


Figure 7. Learning Mixed-Layer Representations

(A) Error rate of a Hebbian classifier before (magenta) and after (black) unsupervised learning. The input-layer activity and mixed-layer coding level are initially heterogeneous (see STAR Methods). Parameters are $N = 1,000$, $M/N = 5$, $P = 1,000$, and $f = 0.1$. Global inhibition is present, and excitatory synaptic weights are initialized to be homogeneous before learning.

(B) Effect of supervised learning on error rate. Results are shown for learning that associates inputs with either random mixed-layer target patterns (blue) or mixed-layer target patterns in which

neurons are sensitive to combinations of input and valence (cyan). Results are compared to the case of synaptic weights drawn randomly from a standard Gaussian distribution (black). Other parameters are as in (A).

(C) Same as (B) but for mixed-layer dimension averaged over the distribution of random input patterns.

activity) in the granule-cell layer and did not consider its dimension (Marr, 1969; Albus, 1971). While a low coding level can increase dimension (Figure S1), it is not sufficient to do so, because correlations in activity can limit dimension regardless of the coding level (Figure 2A). The representations we considered exhibit low coding levels, but we note that a low coding level does not require or imply sparse connectivity. Marr (1969) implied that sparse connectivity can ensure that the granule-cell coding level does not saturate as the level of mossy fiber activity is varied, and a similar argument has been made for the mutual information between mossy-fiber input and granule-cell responses (Billings et al., 2014). However, the importance of sparse connectivity for this purpose is reduced if inhibition is included (see Figure 4C) or if preprocessing normalizes the activity of the inputs (see below). Our result that small synaptic degree maximizes dimension does not depend strongly on coding level (Figures 3C and 3D).

Our analysis shows that when synaptic wiring is random, the dimension of a representation formed by a large number of sparsely connected neurons is typically higher than that of a smaller number of densely connected neurons (Figure 3). This intuition holds as long as the ratio of mixed-layer neurons to inputs is below 10–50. Above this expansion ratio, increases in dimension are modest and require both well-tuned inhibition and large increases in synaptic degree (Figure S1). The observed expansion ratio in the *Drosophila* mushroom body and an estimate of the expansion ratio of the granule cells presynaptic to a single Purkinje cell in cerebellar cortex both lie within this range of 10–50 (Keene and Waddell, 2007; Marr, 1969; Tyrrell and Willshaw, 1992). Thus, the extensive dendritic arbors of the output neurons in both circuits may provide near-optimal classification ability, allowing cerebellum-like circuits to operate close to the limit of performance for sparsely connected neural systems. While we established a formal link between dimension and one specific computation, binary classification (Figure 6), high-dimensional representations may be useful for a variety of tasks (Rigotti et al., 2013) and thus desirable even in circuits whose output neurons do not perform binary classification.

The Cerebellar Granule-Cell Representation

Classic theories of cerebellar cortex assumed that granule cells employ fixed, random connections from mossy fibers to produce

a high-dimensional representation (Marr, 1969; Albus, 1971). The spatial organization of cerebellar cortex (Billings et al., 2014; Figure 5) rules out the extreme case in which every mossy-fiber input can be assigned randomly to any granule cell. However, within defined spatial domains, inputs to individual granule cells appear heterogeneous (Huang et al., 2013; Chabrol et al., 2015; Ishikawa et al., 2015), supporting a spatially restricted form of randomness. This is sufficient for our theory because granule cells represent heterogeneous combinations of mossy-fiber input relevant to the classification performed by their postsynaptic Purkinje cells, which are unlikely to require all possible input combinations. Some studies have suggested that granule-cell input is more homogeneous (Jörntell and Ekerot, 2006; Bengtsson and Jörntell, 2009) and that granule cells function as noise filters (Ekerot and Jörntell, 2008) or improve generalization (Spanne and Jörntell, 2015). However, even when we impose a corresponding level of similarity of tuning in our cerebellar cortex model, the synaptic degree that maximizes dimension is still small ($K \leq 4$), although the dimension of the resulting representation is reduced (Figure S7). Another study related to ours, of which we recently became aware, shows that sparse connectivity also improves convergence speed in a model of associative learning in cerebellar cortex (N.A. Cayco-Gajic et al., 2016, BMC Neurosci. abstract).

Classic models of cerebellar cortex assume that the inputs to granule cells are fixed (Marr, 1969; Albus, 1971), but plasticity has been reported at mossy-fiber-to-granule-cell synapses (Hansel et al., 2001; Schweighofer et al., 2001; Gao et al., 2012, 2016; D'Angelo, 2014), although these synapses appear morphologically stable (Rylkova et al., 2015). Prior modeling studies have suggested that these modifications reflect an unsupervised learning process that allows granule cells to respond to inputs relevant to a task (Schweighofer et al., 2001). Consistent with this, unsupervised learning improves classification performance in our model (Figure 7A). However, the performance of these models still saturates when synaptic degree is small, in contrast to models subject to supervised learning (Figure 7B). The fact that climbing-fiber error signals target Purkinje cells but not granule cells is consistent with this observation. It is also possible that mossy-fiber plasticity is useful for functions other than improving associative learning performance, and further work is necessary to test this hypothesis (D'Angelo, 2014).

The distribution of the strengths of connections received by cerebellar granule cells from mossy fibers appears to be less variable than the very broad, log-normal distribution observed for neocortical neurons (Sargent et al., 2005; Song et al., 2005). In our model, this reduced heterogeneity increases the dimension of the granule-cell representation (Figure 4E). In contrast, the broader distribution for neocortical neurons may be the outcome of a supervised learning process rather than the basis of a high-dimensional representation, which would be consistent with our predictions (Figure 7) and other theoretical analyses showing that similar distributions can be generated through associative learning (Chapeton et al., 2012; Brunel, 2016).

Neural Representations for Associative Learning

We considered the simple case of learning associations between outputs and random, uncorrelated input patterns with a fixed coding level, a standard benchmark for learning systems (Gardner, 1988; Barak et al., 2013; Babadi and Sompolinsky, 2014). Under these assumptions, we derived a direct relationship between the performance of a Hebbian classifier and dimension (Equation 6), thus making explicit the previously assumed relationship between these two quantities that characterize random mixed-layer representations (Rigotti et al., 2013). Our qualitative results also hold for more sophisticated pseudoinverse and maximum-margin classifiers, which can achieve low error rates when mixed-layer neurons have heterogeneous coding levels (Figure S8).

The activities of the inputs to biological systems will inevitably exhibit some level of correlation, violating the simplifying assumptions of our model. However, the magnitude of this correlation may be reduced by decorrelating mechanisms within pre-processing circuitry such as the *Drosophila* antennal lobe (Bhandawat et al., 2007). Similarly, normalizing circuitry may ensure that input coding levels remain relatively constant (Carandini and Heeger, 2011). For more complicated distributions of input patterns and noise that strongly violate our assumptions, different analyses are required, and the results will depend on the exact forms of these distributions. Some simple forms of correlation among input channels do not substantially affect the location of the optimal synaptic degree (Figure S7). For a classification problem related to the challenging “parity” task, in which only knowledge of the pattern of activity in every one of a set of independent input streams provides any information about valence (Barak et al., 2013), larger synaptic degree is needed when there are many input streams (Figure S8). Dense connectivity or structural plasticity mechanisms capable of selecting relevant inputs may be necessary for such tasks.

Our analysis also did not consider the complex temporal dynamics that underlie many forms of associative learning. To extend our model to the temporal domain for systems that lack strong recurrence, each input pattern could be interpreted as the averaged activity during a period comparable to the integration time constant of a mixed-layer neuron. In this case, it is necessary to consider an ensemble of patterns with correlations determined by the temporal statistics of the input. Short-term synaptic plasticity (Xu-Friedman and Regehr, 2003; Chabrol et al., 2015) and unsupervised learning mechanisms (Schweighofer

et al., 2001) are likely to control these statistics to ensure well-timed input to mixed-layer neurons. In general, we expect the qualitative result that small synaptic degree is sufficient for high performance to hold, provided that the temporal overlap of inputs is sufficient to allow mixed-layer neurons to respond to heterogeneous input combinations.

In our model, supervised modifications of the synaptic weights received by mixed-layer neurons, in addition to those of the classifier, are required to exploit large synaptic degree, but they provide little improvement when synaptic degree is small (Figure 7B). We also found little improvement at small synaptic degree using a state-of-the-art implementation of a gradient descent algorithm for modifying synaptic weights (Figure S9; Kingma and Ba, 2014). We cannot rule out the existence of other learning rules that lead to improved performance, but, if they exist, these rules may not be easily implementable by biological systems.

Learning in Cerebellum-like and Cerebrocortical Systems

These observations suggest that supervised learning of the inputs to cerebellar granule cells or analogous neurons in cerebellum-like circuits may be unnecessary to achieve near-optimal performance for learning associations when input-layer activity is normalized (Figure 7). As described above, the modifications of mossy-fiber to cerebellar granule-cell synaptic transmission that have been identified in vitro and in vivo may underlie an unsupervised learning process (Schweighofer et al., 2001). Another study reported enhanced Kenyon-cell responses to conditioned odors in honeybees, but this change may reflect learning at other stages of the olfactory processing hierarchy (Szyszka et al., 2008). In cephalopods, plasticity in the vertical lobe, a cerebellum-like structure, appears to be confined to output neurons in cuttlefish but to be present only in the dendrites of intermediate-layer, not output-layer, neurons in *Octopus vulgaris* (Shomrat et al., 2011). This does not contradict our analysis, which predicts that supervised learning of mixed-layer responses is dispensable only in the presence of output-layer plasticity. Experiments suggest that Kenyon cells in the locust mushroom body are densely connected (Jortner et al., 2007), unlike in *Drosophila*, implying that their input synapses may potentially exhibit associative modifications.

For densely connected neural systems, we investigated various forms of supervised modifications of input-layer to mixed-layer connections and found that learning based on mixed-layer conjunctions of valence and input provides the greatest improvement. This is consistent with the observation that both task-relevant variables and reward are represented in higher neocortical regions (Saez et al., 2015). Although our analysis neglects features of cerebral cortex, including recurrence and complex temporal dynamics, it suggests that supervised learning of both inputs and outputs is critical to take advantage of the dense connectivity of cerebrocortical systems. Thus, these systems are likely to exploit supervised synaptic plasticity at multiple levels of processing, similar to artificial neural networks that attain high performance on complex categorization tasks (LeCun et al., 2015).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **METHOD DETAILS**
 - Combinatorics
 - Neuron Model
 - Dimension
 - Classifier Performance
 - Unsupervised Learning of Mixed-Layer Inputs
 - Supervised Learning of Mixed-Layer Inputs
 - Calculation of Input-Current Dimension
 - Calculation of Mixed-Layer Dimension
 - Spatial Model of Mossy Fiber-Granule Cell Connectivity
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND SOFTWARE AVAILABILITY**

SUPPLEMENTAL INFORMATION

Supplemental Information includes nine figures and can be found with this article online at <http://dx.doi.org/10.1016/j.neuron.2017.01.030>.

AUTHOR CONTRIBUTIONS

A.L.-K., K.D.H., R.A., H.S., and L.F.A. conceived the study and wrote the paper. A.L.-K. and K.D.H. performed simulations and analyses.

ACKNOWLEDGMENTS

We are grateful to Natasha Cayco-Gajic for discussions and advice concerning modeling of cerebellar cortex. We thank Stefano Fusi and Nate Sawtell for comments on the manuscript and Mark Goldman and Hannah Payne for discussions that helped shape an early version of this study. We also acknowledge the support of the Marine Biological Laboratories in Woods Hole, where the work was partially completed. Research was supported by NIH grant F32DC014387 (A.L.-K.), the University of Washington NIH Big Data for Genomics and Neuroscience training grant (T32LM012419; K.D.H.), a Boeing Fellowship (K.D.H.), the Howard Hughes Medical Institute (R.A.), the Mathers Foundation (R.A. and L.F.A.), NIH grant U01NS090449 (H.S.), the Gatsby Charitable Foundation and the Simons Foundation under the Simons Collaboration for the Global Brain (H.S. and L.F.A.), CRCNS grant NSF-1430065 (L.F.A.), the Swartz Foundation (L.F.A.), and the Kavli Foundation (L.F.A.).

Received: July 25, 2016

Revised: November 3, 2016

Accepted: January 30, 2017

Published: February 16, 2017

REFERENCES

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems. *arXiv*, arXiv:1603.04467. <https://arxiv.org/abs/1603.04467>.
- Abbott, L.F., and Nelson, S.B. (2000). Synaptic plasticity: taming the beast. *Nat. Neurosci.* 3, 1178–1183.
- Abbott, L.F., Rajan, K., and Sompolinsky, H. (2011). Interactions between intrinsic and stimulus-evoked activity in recurrent neural networks. In *The Dynamic Brain: An Exploration of Neuronal Variability and Its Functional Significance*, M. Ding and D.L. Gliman, eds. (Oxford University Press), pp. 65–82.
- Albus, J.S. (1971). A theory of cerebellar function. *Math. Biosci.* 10, 25–61.
- Babadi, B., and Sompolinsky, H. (2014). Sparseness and expansion in sensory representations. *Neuron* 83, 1213–1226.
- Barak, O., Rigotti, M., and Fusi, S. (2013). The sparseness of mixed selectivity neurons controls the generalization-discrimination trade-off. *J. Neurosci.* 33, 3844–3856.
- Bell, C.C., Han, V., and Sawtell, N.B. (2008). Cerebellum-like structures and their implications for cerebellar function. *Annu. Rev. Neurosci.* 31, 1–24.
- Bengtsson, F., and Jörntell, H. (2009). Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. *Proc. Natl. Acad. Sci. USA* 106, 2389–2394.
- Bhandawat, V., Olsen, S.R., Gouwens, N.W., Schlieff, M.L., and Wilson, R.I. (2007). Sensory processing in the *Drosophila* antennal lobe increases reliability and separability of ensemble odor representations. *Nat. Neurosci.* 10, 1474–1482.
- Billings, G., Piasini, E., Lőrincz, A., Nusser, Z., and Silver, R.A. (2014). Network structure within the cerebellar input layer enables lossless sparse encoding. *Neuron* 83, 960–974.
- Brunel, N. (2016). Is cortical connectivity optimized for storing information? *Nat. Neurosci.* 19, 749–755.
- Buonomano, D.V., and Merzenich, M.M. (1998). Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186.
- Carandini, M., and Heeger, D.J. (2011). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62.
- Caron, S.J.C., Ruta, V., Abbott, L.F., and Axel, R. (2013). Random convergence of olfactory inputs in the *Drosophila* mushroom body. *Nature* 497, 113–117.
- Chabrol, F.P., Arenz, A., Wiechert, M.T., Margrie, T.W., and DiGregorio, D.A. (2015). Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nat. Neurosci.* 18, 718–727.
- Chadderton, P., Margrie, T.W., and Häusser, M. (2004). Integration of quanta in cerebellar granule cells during sensory processing. *Nature* 428, 856–860.
- Chapeton, J., Fares, T., LaSota, D., and Stepanyants, A. (2012). Efficient associative memory storage in cortical circuits of inhibitory and excitatory neurons. *Proc. Natl. Acad. Sci. USA* 109, E3614–E3622.
- D'Angelo, E. (2014). The organization of plasticity in the cerebellar cortex: from synapses to control. *Prog. Brain Res.* 270, 31–58.
- Eccles, J.C., Llinás, R., and Sasaki, K. (1966). The mossy fibre-granule cell relay of the cerebellum and its inhibitory control by Golgi cells. *Exp. Brain Res.* 1, 82–101.
- Eccles, J.C., Ito, M., and Szentagotai, J., eds. (1967). *The Cerebellum as a Neuronal Machine* (Springer-Verlag).
- Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587.
- Ekerot, C.F., and Jörntell, H. (2008). Synaptic integration in cerebellar granule cells. *Cerebellum* 7, 539–541.
- Fox, C.A., Hillman, D.E., Siegesmund, K.A., and Dutta, C.R. (1967). The primate cerebellar cortex: a Golgi and electron microscopic study. *Prog. Brain Res.* 25, 174–225.
- Gao, Z., van Beugen, B.J., and De Zeeuw, C.I. (2012). Distributed synergistic plasticity and cerebellar learning. *Nat. Rev. Neurosci.* 13, 619–635.
- Gao, Z., Proietti-Onori, M., Lin, Z., Ten Brinke, M.M., Boele, H.J., Potters, J.W., Ruigrok, T.J., Hoebeek, F.E., and De Zeeuw, C.I. (2016). Excitatory cerebellar nucleocortical circuit provides internal amplification during associative conditioning. *Neuron* 89, 645–657.
- Gardner, E. (1988). The space of interactions in neural network models. *J. Phys. A: Math. Gen.* 21, 257.

- Gruntman, E., and Turner, G.C. (2013). Integration of the olfactory code across dendritic claws of single mushroom body neurons. *Nat. Neurosci.* 16, 1821–1829.
- Hansel, D., and van Vreeswijk, C. (2012). The mechanism of orientation selectivity in primary visual cortex without a functional map. *J. Neurosci.* 32, 4049–4064.
- Hansel, C., Linden, D.J., and D'Angelo, E. (2001). Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum. *Nat. Neurosci.* 4, 467–475.
- Harvey, R.J., and Napper, R.M. (1991). Quantitative studies on the mammalian cerebellum. *Prog. Neurobiol.* 36, 437–463.
- Hige, T., Aso, Y., Modi, M.N., Rubin, G.M., and Turner, G.C. (2015). Heterosynaptic plasticity underlies aversive olfactory learning in *Drosophila*. *Neuron* 88, 985–998.
- Honegger, K.S., Campbell, R.A., and Turner, G.C. (2011). Cellular-resolution population imaging reveals robust sparse coding in the *Drosophila* mushroom body. *J. Neurosci.* 31, 11772–11785.
- Huang, C.C., Sugino, K., Shima, Y., Guo, C., Bai, S., Mensh, B.D., Nelson, S.B., and Hantman, A.W. (2013). Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. *eLife* 2, e00400.
- Ishikawa, T., Shimuta, M., and Häusser, M. (2015). Multimodal sensory integration in single cerebellar granule cells in vivo. *eLife* 4, e12916.
- Ito, M., Sakurai, M., and Tongroach, P. (1982). Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *J. Physiol.* 324, 113–134.
- Jörlt, H., and Ekerot, C.F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *J. Neurosci.* 26, 11786–11797.
- Jortner, R.A., Farivar, S.S., and Laurent, G. (2007). A simple connectivity scheme for sparse coding in an olfactory system. *J. Neurosci.* 27, 1659–1669.
- Keene, A.C., and Waddell, S. (2007). *Drosophila* olfactory memory: single genes to complex neural circuits. *Nat. Rev. Neurosci.* 8, 341–354.
- Kennedy, A., Wayne, G., Kaifosh, P., Alviña, K., Abbott, L.F., and Sawtell, N.B. (2014). A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat. Neurosci.* 17, 416–422.
- Kingma, D., and Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv*, arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444.
- Liu, X., and Davis, R.L. (2009). The GABAergic anterior paired lateral neuron suppresses and is suppressed by olfactory learning. *Nat. Neurosci.* 12, 53–59.
- Llinás, R.R., Walton, K.D., and Lang, E.J. (2003). Cerebellum. In *The Synaptic Organization of the Brain*, 5th edition, G.M. Shepherd, ed. (Oxford University Press), pp. 271–309.
- Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.* 500, 409–440.
- Marr, D. (1969). A theory of cerebellar cortex. *J. Physiol.* 202, 437–470.
- Mugnaini, E., Osen, K.K., Dahl, A.L., Friedrich, V.L., Jr., and Korte, G. (1980). Fine structure of granule cells and related interneurons (termed Golgi cells) in the cochlear nuclear complex of cat, rat and mouse. *J. Neurocytol.* 9, 537–570.
- Murthy, M., Fiete, I., and Laurent, G. (2008). Testing odor response stereotypy in the *Drosophila* mushroom body. *Neuron* 59, 1009–1023.
- Raymond, J.L., Lisberger, S.G., and Mauk, M.D. (1996). The cerebellum: a neuronal learning machine? *Science* 272, 1126–1131.
- Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590.
- Rigotti, M., Barak, O., Warden, M.R., Wang, X.J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590.
- Rylkova, D., Crank, A.R., and Linden, D.J. (2015). Chronic in vivo imaging of ponto-cerebellar mossy fibers reveals morphological stability during whisker sensory manipulation in the adult rat. *eNeuro* 2, ENEURO.0075–15.2015.
- Saez, A., Rigotti, M., Ostojic, S., Fusi, S., and Salzman, C.D. (2015). Abstract context representations in primate amygdala and prefrontal cortex. *Neuron* 87, 869–881.
- Sargent, P.B., Saviane, C., Nielsen, T.A., DiGregorio, D.A., and Silver, R.A. (2005). Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse. *J. Neurosci.* 25, 8173–8187.
- Schweighofer, N., Doya, K., and Lay, F. (2001). Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control. *Neuroscience* 103, 35–50.
- Shomrat, T., Graindorge, N., Bellanger, C., Fiorito, G., Loewenstein, Y., and Hochner, B. (2011). Alternative sites of synaptic plasticity in two homologous “fan-out fan-in” learning and memory networks. *Curr. Biol.* 21, 1773–1782.
- Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3, e68.
- Spanne, A., and Jörlt, H. (2015). Questioning the role of sparse coding in the brain. *Trends Neurosci.* 38, 417–427.
- 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, 2123–2130.
- Szyska, P., Galkin, A., and Menzel, R. (2008). Associative and non-associative plasticity in Kenyon cells of the honeybee mushroom body. *Front. Syst. Neurosci.* 2, 3.
- Tyrrell, T., and Willshaw, D. (1992). Cerebellar cortex: its simulation and the relevance of Marr's theory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 239–257.
- Wiechert, M.T., Judkewitz, B., Riecke, H., and Friedrich, R.W. (2010). Mechanisms of pattern decorrelation by recurrent neuronal circuits. *Nat. Neurosci.* 13, 1003–1010.
- Xu-Friedman, M.A., and Regehr, W.G. (2003). Ultrastructural contributions to desensitization at cerebellar mossy fiber to granule cell synapses. *J. Neurosci.* 23, 2182–2192.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--|---------------------|---|
| Software and Algorithms | | |
| TensorFlow machine learning library | Abadi et al., 2016 | https://www.tensorflow.org |
| Adam gradient descent optimizer | Kingma and Ba, 2014 | http://arxiv.org/abs/1412.6980 |
| Algorithms for calculation of dimension and error rate | This paper | http://www.columbia.edu/~ak3625/code/litwin-kumar_et_al_dimension_2017.zip |

CONTACT FOR REAGENT AND RESOURCE SHARING

Further requests for resources should be directed to and will be fulfilled by the Lead Contact, Ashok Litwin-Kumar (ak3625@columbia.edu).

METHOD DETAILS

Combinatorics

Suppose we draw M items with replacement from R possibilities. The probability p that all items are distinct is given by

$$p = \frac{R(R-1)(R-2)\dots(R-M+1)}{R^M} = \frac{R!}{(R-M)!R^M}. \quad (\text{Equation 7})$$

For the case of a synaptic degree of K and N input channels, setting $R = \binom{N}{K}$ yields the probability of all mixed-layer neurons being distinct.

Neuron Model

The mixed-layer responses are given by

$$\mathbf{m} = \Theta(\mathbf{J}\mathbf{s} - \boldsymbol{\theta}), \quad (\text{Equation 8})$$

where \mathbf{J} are the synaptic weights, \mathbf{s} is the input-layer activity, $\boldsymbol{\theta}$ are the activity thresholds, and Θ is the Heaviside step function (applied element-wise). When inhibition is present, $\mathbf{J} = \mathbf{J}^+ + \mathbf{J}^-$, where \mathbf{J}^+ represents excitatory synaptic weights and $\mathbf{J}^- = \mathbf{J}_m \mathbf{J}_s^T$ represents global inhibition. \mathbf{J}_m is a vector of length M with each element equal to $-\alpha$, and \mathbf{J}_s is a vector of length N with each element either equal to 1 (homogeneous inhibition) or equal to $(1/N_I) \sum_{j=1}^{N_I} w_{I,j}$ (heterogeneous inhibition), where $w_{I,j}$ is drawn from the same distribution as the excitatory synaptic weights and N_I is the number of inhibitory neurons. When $\alpha = K/N$, inhibition and excitation are balanced. For all analyses of dimension, we choose the value of α that minimizes input-current correlations; this value depends on the distribution of inhibitory synaptic weights and is less than K/N except for homogeneous weights (see [Calculation of Input-Current Dimension](#)). We note that our model is insensitive to the mean of the distribution of input activities, since the activity thresholds can be adjusted to compensate for a change in mean (specifically, if $s_j \leftarrow s_j + \mu$, the thresholds may be redefined as $\theta_i \leftarrow \theta_i + \mu \sum_{j=1}^N J_{ij}$).

For heterogeneous weights, synaptic weights were drawn from a log-normal distribution with $\mu = -.702$, $\sigma = .936$ (consistent with neocortical data; [Song et al., 2005](#)) or $\mu = 0$, $\sigma = 0.438$ (leading to a CV of 0.46 consistent with data from cerebellar granule cells; [Sargent et al., 2005](#)).

Dimension

To compute the dimension of a system whose state is defined by \mathbf{x} , we construct the covariance matrix $\mathbf{C} = \langle (\mathbf{x} - \langle \mathbf{x} \rangle)(\mathbf{x} - \langle \mathbf{x} \rangle)^T \rangle$. Using $\text{Tr } \mathbf{C} = \sum_i \lambda_i$, where $\{\lambda_i\}$ are the eigenvalues of \mathbf{C} , and noting that \mathbf{C} is symmetric, we find

$$\dim(\mathbf{x}) = \frac{(\text{Tr } \mathbf{C})^2}{\text{Tr } \mathbf{C}^2} = \frac{M \langle C_{ii} \rangle^2}{\langle C_{ii}^2 \rangle + (M-1) \langle C_{ij}^2 \rangle}, \quad (\text{Equation 9})$$

where C_{ii} and C_{ij} are random variables representing the diagonal and off-diagonal elements of \mathbf{C} , respectively. For the case of the input currents, analytic expressions can be obtained in terms of the distribution of synaptic weights (see [Calculation of Input-Current Dimension](#)), leading to [Equations 2](#) and [4](#). For the case of binary mixed-layer responses with identical coding levels for each neuron, $\langle C_{ii} \rangle = f(1-f)$, $\text{Var}(C_{ii}) = 0$, and the expression reduces to [Equation 3](#) when M is large. The distribution of C_{ij} can be obtained exactly

for the case of homogeneous weights, while for heterogeneous weights we estimated the distribution by sampling (see [Heterogeneous Weights](#)).

Classifier Performance

The synaptic weights of the Hebbian classifier are given by $\mathbf{w} = \sum_{\mu=1}^P (\mathbf{m}^{\mu} - f) v^{\mu}$, where $\mathbf{m}^{\mu} - f$ is short for element-wise subtraction of f from \mathbf{m}^{μ} . We assume that the training patterns and noisy test patterns have a mean-squared distance (equivalent to the Hamming distance divided by M) of d and equal coding levels. Denoting the classifier input by $\mathbf{g}^{\mu} = \mathbf{w} \cdot (\widehat{\mathbf{m}}^{\mu} - f)$, where $\widehat{\mathbf{m}}^{\mu}$ represents a noisy test pattern,

$$\langle v^{\mu} \mathbf{g}^{\mu} \rangle = \langle (\mathbf{m}^{\mu} - f) \cdot (\widehat{\mathbf{m}}^{\mu} - f) \rangle = M(f(1-f) - d/2). \quad (\text{Equation 10})$$

The average is taken over noise realizations. The variance $\text{Var}(v^{\mu} \mathbf{g}^{\mu})$ is equal to

$$\text{Var} \left(\sum_{\nu=1}^P v^{\mu} v^{\nu} (\mathbf{m}^{\nu} - f) \cdot (\widehat{\mathbf{m}}^{\mu} - f) \right). \quad (\text{Equation 11})$$

When P is large, this quantity can be approximated by the sum over $\nu \neq \mu$. In the special case of random and uncorrelated patterns and noise, then if $\nu \neq \mu$, the statistics of $v^{\mu} v^{\nu} (\mathbf{m}^{\nu} - f) \cdot (\widehat{\mathbf{m}}^{\mu} - f)$ are identical to those of $O_{\mu\nu} \equiv (\mathbf{m}^{\nu} - f) \cdot (\mathbf{m}^{\mu} - f)$, which are Gaussian distributed with mean zero. Therefore $\text{Var}(v^{\mu} \mathbf{g}^{\mu}) \approx P \text{Var}(O_{\mu\nu})$. $O_{\mu\nu}$ is a random variable representing the overlap of the mixed-layer responses to two random input patterns. We now relate these overlaps to the covariance matrix of mixed-layer responses, and thus to dimension.

Consider a set of Q mixed-layer responses to random input patterns $\{\mathbf{m}^{\mu}\}$, $\mu = 1 \dots Q$, and construct the matrix \mathbf{M} , whose μ th column is equal to $\mathbf{m}^{\mu} - f$. The covariance matrix of mixed-layer responses can be written as $\mathbf{C} = \lim_{Q \rightarrow \infty} (1/Q) \mathbf{M} \mathbf{M}^T$. Furthermore, $\text{Var}(O_{\mu\nu})$ is equal to the variance of the off-diagonal elements of $\mathbf{O} = \mathbf{M}^T \mathbf{M}$. Therefore,

$$\begin{aligned} \text{Tr } \mathbf{C}^2 &= \lim_{Q \rightarrow \infty} \frac{1}{Q^2} \text{Tr}(\mathbf{M} \mathbf{M}^T)^2 = \lim_{Q \rightarrow \infty} \frac{1}{Q^2} \text{Tr}(\mathbf{M}^T \mathbf{M})^2 = \lim_{Q \rightarrow \infty} \frac{1}{Q^2} \text{Tr } \mathbf{O}^2 \\ &= \lim_{Q \rightarrow \infty} \frac{1}{Q^2} \left(Q \langle O_{\mu\mu} \rangle + Q(Q-1) \langle O_{\mu\nu}^2 \rangle \right) = \langle O_{\mu\nu}^2 \rangle = \text{Var}(O_{\mu\nu}). \end{aligned} \quad (\text{Equation 12})$$

Thus,

$$\text{Var}(v^{\mu} \mathbf{g}^{\mu}) \approx P \text{Var}(O_{\mu\nu}) = P(\text{Tr } \mathbf{C}^2).$$

We define the signal-to-noise ratio

$$\text{SNR} = \frac{\langle v^{\mu} \mathbf{g}^{\mu} \rangle^2}{\text{Var}(v^{\mu} \mathbf{g}^{\mu})} \approx \frac{M^2(f(1-f) - d/2)^2}{P(\text{Tr } \mathbf{C}^2)}. \quad (\text{Equation 13})$$

For the case of a fixed coding level, $\text{Tr } \mathbf{C} = M f(1-f)$. We can therefore write

$$\text{SNR} \approx \frac{(\text{Tr } \mathbf{C})^2 \left(1 - \frac{d}{2f(1-f)}\right)^2}{(\text{Tr } \mathbf{C}^2) P} = \frac{\dim(\mathbf{m}) \cdot (1 - \Delta)^2}{P}. \quad (\text{Equation 14})$$

Unsupervised Learning of Mixed-Layer Inputs

To study unsupervised learning in the presence of heterogeneous input-layer activity, we calculated the error rate of the classifier before and after an unsupervised learning period consisting of 1,000 presentations of random input patterns. During this learning period, but not during the calculation of error rate, synaptic weights and the thresholds of mixed-layer neurons are modified.

The activity of the j th input channel is drawn from a Gaussian distribution with mean 0 and variance σ_j^2 , with σ_j uniformly distributed between 0.25 and 1.75. We define a variable $u_{ij}(t)$ for each connection, with $u_{ij}(0) = 1$. When an input pattern $\mathbf{s}(t)$ is presented, u_{ij} is updated according to

$$u_{ij}(t) = (1 - \alpha) u_{ij}(t-1) + \alpha (s_j(t))^2, \quad (\text{Equation 15})$$

with $\alpha = 0.01$ and $t = 1, 2, \dots, 1,000$. The variable $u_{ij}(t)$ thus estimates the variance of input channel j . Synaptic weights are set equal to $J_{ij}(t) = (u_{ij}(t))^{-1/2}$.

To update neuronal thresholds, we initially set $\theta_i(0)$ to be identical for all i , leading to heterogeneous coding levels across neurons (the initial value is chosen so that the average coding level is 0.1). When an input pattern is presented, thresholds are updated according to

$$\theta_i(t) = \begin{cases} \theta_i(t-1) + \eta & \text{if neuron } i \text{ is active, and} \\ \theta_i(t-1) - \eta \frac{f_t}{1-f_t} & \text{otherwise,} \end{cases} \quad (\text{Equation 16})$$

with $\eta = 0.01$. This learning rule ensures that each neuron approaches a target coding level of $f_t = 0.1$.

Supervised Learning of Mixed-Layer Inputs

We compared classification performance when input-layer to mixed-layer connections were chosen from a standard Gaussian distribution to performance when they were modified by Hebbian learning. Since the learning rule does not constrain the sign of the synaptic weights, we did not require the elements of \mathbf{J} to be positive. To implement Hebbian learning, each input pattern \mathbf{s}^μ is associated with a target mixed-layer pattern \mathbf{t}^μ . The synaptic weights are then set according to

$$\mathbf{J} = \sum_{\mu=1}^P (\mathbf{t}^\mu - f) (\mathbf{s}^\mu)^T. \quad (\text{Equation 17})$$

Target patterns are either random binary patterns with coding level f or patterns that mix input and valence. For the latter case, target patterns are generated by setting

$$\mathbf{t}^\mu = \Theta(\mathbf{J}_0 \mathbf{s}^\mu + \mathbf{J}_v \mathbf{v}^\mu - \theta_t). \quad (\text{Equation 18})$$

θ_t is chosen so that each mixed-layer neuron is active for a fraction f of the target patterns. \mathbf{J}_0 is a $M \times N$ matrix with entries drawn independently from a zero-mean Gaussian distribution with variance $1/N$, while \mathbf{J}_v is a M -dimensional vector with entries drawn independently from a zero-mean Gaussian distribution with variance 100.

Calculation of Input-Current Dimension

If the input patterns have zero mean, then $\mathbf{h}^\mu = \mathbf{J} \mathbf{s}^\mu$ has zero mean. The (i, j) th element of the covariance matrix of the input currents is therefore

$$C_{ij} = \langle h_i^\mu h_j^\mu \rangle = \langle (\mathbf{J}_i \cdot \mathbf{s}^\mu) (\mathbf{J}_j \cdot \mathbf{s}^\mu) \rangle = \mathbf{J}_i \langle \mathbf{s}^\mu (\mathbf{s}^\mu)^T \rangle (\mathbf{J}_j)^T, \quad (\text{Equation 19})$$

where \mathbf{J}_i is the i th row of \mathbf{J} and the average is taken over input patterns for a fixed weight matrix. If the input patterns are also uncorrelated, $\langle \mathbf{s}^\mu (\mathbf{s}^\mu)^T \rangle = \mathbf{I}$, and the above expression leads to $\mathbf{C} = \mathbf{J} \mathbf{J}^T$. This is the case we analyze.

To compute the input-current dimension, we calculate

$$\dim(\mathbf{h}) = \frac{\left(\sum_{i=1}^N \lambda_i \right)^2}{\sum_{i=1}^N \lambda_i^2}, \quad (\text{Equation 20})$$

where $\{\lambda_i\}$ are the eigenvalues of \mathbf{C} , and therefore

$$\dim(\mathbf{h}) = \frac{(\text{Tr } \mathbf{C})^2}{\text{Tr } \mathbf{C}^2}. \quad (\text{Equation 21})$$

If C_{ii} is a random variable representing the diagonal elements of \mathbf{C} and C_{ij} the off-diagonal elements, then

$$\dim(\mathbf{h}) = \frac{M \langle C_{ii} \rangle^2}{\langle C_{ii}^2 \rangle + (M-1) \langle C_{ij}^2 \rangle} = \frac{M \langle C_{ii} \rangle^2}{\langle C_{ii} \rangle^2 + \text{Var}(C_{ii}) + (M-1) (\langle C_{ij} \rangle^2 + \text{Var}(C_{ij}))}. \quad (\text{Equation 22})$$

The averages are taken over the population of mixed-layer neurons. Thus, the dimension is determined by the mean and variance of the distributions of $C_{ii} = \mathbf{J}_i \cdot \mathbf{J}_i$ and $C_{ij} = \mathbf{J}_i \cdot \mathbf{J}_j$ across the population.

Homogeneous Weights

We first consider the case where the rows of \mathbf{J} each have K nonzero elements, at random locations. At these locations, $J_{ij} = 1$. Thus, $C_{ii} = K$ and $\text{Tr } \mathbf{C} = MK$. For $i \neq j$, $C_{ij} \sim \text{Hypergeom}(N, K, K)$, and $C_{ij} = C_{ji}$. This yields

$$\langle C_{ii} \rangle = K$$

$$\text{Var}(C_{ii}) = 0$$

$$\langle C_{ij} \rangle = \frac{K^2}{N}$$

$$\text{Var}(C_{ij}) = \frac{K^2}{N} \left(1 - \frac{K}{N}\right) \frac{N-K}{N-1}. \quad (\text{Equation 23})$$

Thus,

$$\dim(\mathbf{h}) = \frac{M K^2}{K^2 + (M-1) \left(\frac{K^4}{N^2} + \frac{K^2}{N} \left(1 - \frac{K}{N}\right) \frac{N-K}{N-1} \right)}. \quad (\text{Equation 24})$$

When $M, N \gg K$ this simplifies to

$$\dim(\mathbf{h}) = \frac{1}{\frac{1}{M} + \frac{1}{N} \left(1 + \frac{(K-1)^2}{N}\right)}. \quad (\text{Equation 25})$$

In the limit of a large expansion, when $M \gg N \gg K$,

$$\dim(\mathbf{h}) = \frac{N}{1 + \frac{(K-1)^2}{N}}. \quad (\text{Equation 26})$$

Homogeneous Weights with Homogeneous Inhibition

We now suppose that all neurons receive inhibition proportional to the summed activity of the input layer. This is equivalent to adding a negative offset to all weights. Let $\mathbf{J} = \mathbf{J}^+ + \mathbf{J}^-$, where \mathbf{J}^+ is determined as above for homogeneous weights and $J_{ij}^- = -\alpha$. Using the results of the previous section leads to

$$\langle C_{ii} \rangle = K(1 - \alpha)^2 + (N - K)\alpha^2$$

$$\text{Var}(C_{ii}) = 0$$

$$\langle C_{ij} \rangle = \frac{K^2}{N} - 2K\alpha + N\alpha^2$$

$$\text{Var}(C_{ij}) = \frac{K^2}{N} \left(1 - \frac{K}{N}\right) \frac{N-K}{N-1}. \quad (\text{Equation 27})$$

When $\alpha = K/N$, the row sums of \mathbf{J} are equal to zero and excitation and inhibition are “balanced.” In this case, $\langle C_{ii} \rangle = K(1 - (K/N))$ and $\langle C_{ij} \rangle = 0$. Hence

$$\dim(\mathbf{h}) = \frac{M K^2 \left(1 - \frac{K}{N}\right)^2}{K^2 \left(1 - \frac{K}{N}\right)^2 + (M-1) \left(\frac{K^2}{N} \left(1 - \frac{K}{N}\right) \frac{N-K}{N-1} \right)}. \quad (\text{Equation 28})$$

When $M, N \gg K$ this simplifies to

$$\dim(\mathbf{h}) = \frac{1}{\frac{1}{M} + \frac{1}{N}}. \quad (\text{Equation 29})$$

Heterogeneous Weights

Now we suppose each nonzero element of \mathbf{J} is drawn from a distribution so that $J_{ij} \sim w$, where w is a random variable. In this case,

$$\langle C_{ii} \rangle = K \langle w^2 \rangle$$

$$\text{Var}(C_{ii}) = K \left(\langle w^4 \rangle - \langle w^2 \rangle^2 \right). \quad (\text{Equation 30})$$

For a given $i \neq j$, the number of nonzero elements in the sum $\sum_{k=1}^M J_{ij} J_{jk}$ is again distributed as Hypergeom(N, K, K). Also, the individual nonzero elements have mean $\langle w \rangle^2$ and second moment $\langle w^2 \rangle^2$. This leads to

$$\langle C_{ij} \rangle = \langle w \rangle^2 \frac{K^2}{N},$$

$$\text{Var}(C_{ij}) = \frac{K^2}{N} \left(\langle w^2 \rangle^2 + \langle w \rangle^4 \left(-1 + \left(1 - \frac{K}{N} \right) \frac{N-K}{N-1} \right) \right). \quad (\text{Equation 31})$$

Thus, the dimension is

$$\dim(\mathbf{h}) = \frac{M K^2 \langle w^2 \rangle^2}{K \langle w^4 \rangle + K(K-1) \langle w^2 \rangle^2 + (M-1) \frac{K^2}{N} \left(\langle w^2 \rangle^2 + \langle w \rangle^4 \left(-1 + \frac{K^2}{N} + \left(1 - \frac{K}{N} \right) \frac{N-K}{N-1} \right) \right)}. \quad (\text{Equation 32})$$

When $M, N \gg 1$, this simplifies to

$$\dim(\mathbf{h}) = \frac{1}{\frac{1}{M} \left(\frac{1}{K} \frac{\langle w^4 \rangle}{\langle w^2 \rangle^2} + \frac{K-1}{K} \right) + \frac{1}{N} \left(1 + \frac{(K-1)^2}{N} \left(\frac{\langle w \rangle^2}{\langle w^2 \rangle} \right)^2 \right)}. \quad (\text{Equation 33})$$

In the limit of a large expansion, when $M \gg N \gg K$,

$$\dim(\mathbf{h}) = \frac{N}{1 + \frac{(K-1)^2}{N} \left(\frac{\langle w \rangle^2}{\langle w^2 \rangle} \right)^2}. \quad (\text{Equation 34})$$

Heterogeneous Weights with Homogeneous Inhibition

Again, we assume $\mathbf{J} = \mathbf{J}^+ + \mathbf{J}^-$, where \mathbf{J}^+ is determined as above for heterogeneous weights and $J_{ij}^- = -\alpha$. Then,

$$\langle C_{ij} \rangle = K \langle w^2 \rangle - 2K \langle w \rangle \alpha + N \alpha^2$$

$$\text{Var}(C_{ij}) = K \left(\langle w^4 \rangle - \langle w^2 \rangle^2 - 4\alpha \langle w^3 \rangle + 4\alpha \langle w \rangle \langle w^2 \rangle + 4\alpha^2 \langle w^2 \rangle - 4\alpha^2 \langle w \rangle^2 \right)$$

$$\langle C_{ij} \rangle = \frac{K^2}{N} \langle w \rangle^2 - 2K \alpha \langle w \rangle + N \alpha^2. \quad (\text{Equation 35})$$

To determine the variance of $C_{ij} = \sum_k (J_{ik}^+ - \alpha)(J_{jk}^+ - \alpha)$, observe that the terms that contribute to the variance are those for which either J_{ik}^+ or J_{jk}^+ are nonzero. We therefore define $C'_{ij} = \sum_{k, \text{nonzero}} (J_{ik}^+ - \alpha)(J_{jk}^+ - \alpha)$ as the sum over only these entries. This sum can be decomposed into those indices for which both J_{ik}^+ and J_{jk}^+ are nonzero and those for which only one is nonzero. We denote the number of these as n_{ij}^{++} and $2(K - n_{ij}^{++})$, respectively. Given n_{ij}^{++} ,

$$\langle C'_{ij} \mid n_{ij}^{++} = n \rangle = n (\langle w \rangle - \alpha)^2 - 2\alpha(K - n)(\langle w \rangle - \alpha)$$

$$\text{Var}(C'_{ij} \mid n_{ij}^{++} = n) = n \left(2(\langle w \rangle - \alpha)^2 (\langle w^2 \rangle - \langle w \rangle^2) + (\langle w^2 \rangle - \langle w \rangle^2)^2 \right) + 2\alpha^2(K - n)(\langle w^2 \rangle - \langle w \rangle^2). \quad (\text{Equation 36})$$

Now, observing that n_{ij}^{++} is distributed as Hypergeom(N, K, K), we use the law of total variance to find

$$\text{Var}(C_{ij}) = \frac{K^2}{N} \left(2(\langle w \rangle - \alpha)^2 (\langle w^2 \rangle - \langle w \rangle^2) + (\langle w^2 \rangle - \langle w \rangle^2)^2 \right) + 2\alpha^2 \left(K - \frac{K^2}{N} \right) (\langle w^2 \rangle - \langle w \rangle^2) + (\langle w \rangle^2 - \alpha^2)^2 \frac{K^2}{N} \left(1 - \frac{K}{N} \right) \frac{N-K}{N-1}. \quad (\text{Equation 37})$$

When $\alpha = (K \langle w \rangle / N)$, inhibition balances excitation and $\langle C_{ij} \rangle = 0$. In this case and in the limit of $N \gg K$,

$$\langle C_{ij} \rangle = K \langle w^2 \rangle$$

$$\text{Var}(C_{ij}) = K (\langle w^4 \rangle - \langle w^2 \rangle^2)$$

$$\text{Var}(C_{ij}) = \frac{K^2 \langle w^2 \rangle^2}{N}$$

$$\Rightarrow \dim(\mathbf{h}) = \frac{1}{\frac{1}{M} \left(1 + \frac{1}{K} \left(\frac{\langle w^4 \rangle}{\langle w^2 \rangle^2} - 1 \right) \right) + \frac{1}{N}}. \quad (\text{Equation 38})$$

Heterogeneous Weights with Heterogeneous Inhibition

Let us suppose that $\mathbf{J}^- = \mathbf{J}_m \mathbf{J}_s^T$, where the entries of \mathbf{J}_m are homogeneous and equal to $-\alpha$, while the entries of \mathbf{J}_s are drawn from a distribution with mean $\langle w_I \rangle$ and variance $\text{Var}(w_I)$. This represents the case of global inhibition, with heterogeneous weights of the connections from the input layer onto the inhibitory neuron but homogeneous weights of the connections from the inhibitory neuron onto the mixed-layer neurons.

The average correlation between the input currents is

$$\langle C_{ij} \rangle = \frac{K^2}{N} \langle w \rangle^2 - 2K\alpha \langle w \rangle \langle w_I \rangle + N\alpha^2 \langle w_I^2 \rangle. \quad (\text{Equation 39})$$

The value of α that minimizes this quantity is

$$\alpha = \frac{K \langle w \rangle \langle w_I \rangle}{N \langle w_I^2 \rangle}. \quad (\text{Equation 40})$$

Inserting this expression into Equation 39 leads to

$$\langle C_{ij} \rangle = \frac{K^2 \langle w \rangle^2}{N} \left(1 - \frac{\langle w_I \rangle^2}{\langle w_I^2 \rangle} \right). \quad (\text{Equation 41})$$

Also, in this case the average variance is

$$\langle C_{ii} \rangle = K \langle w^2 \rangle - \frac{K \langle w \rangle^2 \langle w_I \rangle^2}{N \langle w_I^2 \rangle}. \quad (\text{Equation 42})$$

Calculation of Mixed-Layer Dimension

For Gaussian patterns and a homogeneous coding level across the mixed-layer neurons, we can evaluate $C_{ij} = \langle (m_i - f)(m_j - f) \rangle = \langle m_i m_j \rangle - f^2$ by determining the variance and covariance of the input currents. The input to neuron i is Gaussian distributed with $\langle h_i \rangle = 0$ and $\text{Var}(h_i) = \sum_k \mathbf{J}_{ik}^2$. This determines the threshold of neuron i through $\theta_i = \sqrt{2\text{Var}(h_i)} \text{erfc}^{-1}(2f)$. The covariance of h_i and h_j is $\text{Cov}(h_i, h_j) = \mathbf{J}_i \cdot \mathbf{J}_j$. These quantities are sufficient to determine C_{ij} using the approach described below. We use this approach to calculate the dimension in the main text for the case of Gaussian patterns. For non-Gaussian patterns, we compute the dimension by direct simulation.

Suppose h_i, h_j are zero-mean Gaussian random variables with variances σ_i^2, σ_j^2 and covariance ρ . We can evaluate $\langle m_i m_j \rangle = \langle \Theta(h_i - \theta_i) \Theta(h_j - \theta_j) \rangle$ by defining three auxiliary unit Gaussian random variables x_i, x_j , and y and writing $h_{i,j} = \sqrt{\sigma_{i,j}^2 - \eta_{i,j}^2} x_{i,j} + \eta_{i,j} y$, where $\eta_i = \sqrt{|\rho| \sigma_i / \sigma_j}$ and $\eta_j = \text{sign}(\rho) \sqrt{|\rho| \sigma_j / \sigma_i}$. The x_i and y variables represent the independent and common components of the input-current fluctuations, respectively. Then

$$\begin{aligned} \langle \Theta(h_i - \theta_i) \Theta(h_j - \theta_j) \rangle &= \\ &= \int_y Dy \left(\int_{x_i} dx_i P \left(\sqrt{\sigma_i^2 - \eta_i^2} x_i + \eta_i y > \theta_i \right) \right) \left(\int_{x_j} dx_j P \left(\sqrt{\sigma_j^2 - \eta_j^2} x_j + \eta_j y > \theta_j \right) \right) \\ &= \int_y Dy \frac{1}{4} \text{erfc} \left(\frac{\theta_i - \eta_i y}{\sqrt{2(\sigma_i^2 - \eta_i^2)}} \right) \text{erfc} \left(\frac{\theta_j - \eta_j y}{\sqrt{2(\sigma_j^2 - \eta_j^2)}} \right), \end{aligned} \quad (\text{Equation 43})$$

where

$$Dy = dy \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

Homogeneous Weights

When weights are homogeneous, the correlation between the responses of a pair of mixed-layer neurons depends only on how many shared inputs they receive. The number of inputs shared by neurons i and j is distributed as $n_{ij}^{++} \sim \text{Hypergeom}(N, K, K)$. If inhibition is absent, $\sigma_i^2 = K$ and $\rho = n_{ij}^{++}$, while if it is present, $\sigma_i^2 = K(1 - K/N)$ and $\rho = n_{ij}^{++} - K^2/N$. These quantities and Equation 43 are sufficient to evaluate the distribution of correlations and thus the dimension.

Heterogeneous Weights

For the case of heterogeneous weights, we obtain random samples from the distribution of C_{ij} by generating random rows of \mathbf{J} , computing σ_i , σ_j , and ρ , and using Equation 43 to compute $\langle m_i m_j \rangle$. These random samples are used to estimate the mean and variance of the distribution.

Spatial Model of Mossy Fiber-Granule Cell Connectivity

We adapted the model described in Billings et al. (2014) to develop an anatomically realistic connectivity matrix between mossy fibers and granule cells. In the model, connections are formed based on the distances between mossy-fiber rosettes and granule cells. We extended the model beyond the 80 μm -diameter sphere in Billings et al. (2014) to cover a region containing all granule cells presynaptic to a single model Purkinje cell.

A single Purkinje cell forms a large dendritic tree approximately 250 μm in diameter (Eccles et al., 1967). We therefore consider a cylinder with diameter 250 μm , representing the unfolded granular-layer region containing granule cells that send parallel fibers to the Purkinje cell. The length of the cylinder is chosen so that the number of granule cells contained within it is equal to the value of 209,000 reported to synapse onto a single Purkinje cell in rats (Harvey and Napper, 1991). Using a granule-cell density of $1.9 \cdot 10^6$ per mm^3 (Harvey and Napper, 1991; Billings et al., 2014), the resulting length is 2,240 μm , which agrees with the typically reported parallel fiber extent of 2,000 – 3,000 μm (Eccles et al., 1967; Albus, 1971). Here we have assumed that all parallel fibers that pass through the Purkinje-cell dendritic tree form connections with it. Anatomical studies indicate that this is true for greater than half of parallel fibers passing through a typical Purkinje-cell dendritic tree (Harvey and Napper, 1991).

Glomeruli containing mossy-fiber rosettes are placed within the cylinder using the procedure in Billings et al. (2014) which we now describe. The density of mossy fibers is equal to the observed glomerular density of $6.6 \cdot 10^5$ per mm^3 divided by the number of rosettes per mossy fiber. In the 80 μm -diameter sphere of Billings et al. (2014), mossy fibers typically formed 2–3 rosettes, but anatomical studies indicate that mossy fibers form on average approximately 20 rosettes in total (Eccles et al., 1967; Fox et al., 1967). We assume that mossy fibers form 10 rosettes in the area we consider, but varying this number does not influence our qualitative results. The value of 10 yields 7,257 unique mossy fibers in the region, consistent with previous estimates by Marr (1969), who estimated 7,000 (Tyrrell and Willshaw, 1992 estimated 13,000). For each mossy fiber, the first rosette is placed uniformly within the cylinder, while subsequent rosettes are displaced relative to the previous rosette by a distance that is exponentially distributed in the x , y , and z directions (the axis of the cylinder is oriented along the z direction). The average displacements are 2 μm , 58 μm , and 21 μm for the x , y , and z directions (Sultan, 2001; Billings et al., 2014), reflecting the anisotropic orientations of the paths of the mossy fibers.

Granule cells and their connections are formed using a procedure identical to Billings et al. (2014). Granule cells are placed uniformly within the cylinder according to the granule-cell density. Each forms connections with $K = 4$ mossy-fiber rosettes. The K connected rosettes are those whose distances to the granule cell are closest to 15 μm , representing the typical granule-cell dendritic length. Granule cells are disallowed from forming multiple connections to a single mossy fiber via multiple rosettes.

QUANTIFICATION AND STATISTICAL ANALYSIS

Results for dimension, error, and other quantities are first averaged over random input patterns for a fixed network architecture, then over random network architectures. Unless otherwise noted, the standard error of the mean across network architectures for the plotted quantities is smaller than the width of the marks.

DATA AND SOFTWARE AVAILABILITY

Software was written in the Julia (<http://julialang.org>) and Python (<http://python.org>) programming languages. Implementations of algorithms used to compute quantities presented in this study are available at: http://www.columbia.edu/~ak3625/code/litwin-kumar_et_al_dimension_2017.zip.