Redundancy in synaptic connections enables neurons to learn optimally

Naoki Hiratani[1,2]* and Tomoki Fukai[1]

[1]Laboratory for Neural Circuit Theory, RIKEN Brain Science Institute, Wako, Saitama, Japan, 351-0198

[2]Gatsby Computational Neuroscience Unit, University College London, London, United Kingdom, W1T 4JG

*corresponding: N.Hiratani@gmail.com

Keywords: synaptic plasticity, connectomics, dendritic computation

Abstract

Recent experimental studies suggest that, in cortical microcircuits of the mammalian brain, the majority of neuron-to-neuron connections are realized by multiple synapses. However, it is not known whether such redundant synaptic connections provide any functional benefit. Here, we show that redundant synaptic connections enable near-optimal learning in cooperation with synaptic rewiring. By constructing a simple dendritic neuron model, we demonstrate that with multisynaptic connections, synaptic plasticity approximates a sample-based Bayesian filtering algorithm known as particle filtering, and wiring plasticity implements its resampling process. The derived synaptic plasticity rule accounts for many experimental observations, including the dendritic position dependence of spike-timing-dependent plasticity. The proposed framework is applicable to detailed single neuron models, and also to recurrent circuit models. Our study provides a novel conceptual framework for synaptic plasticity and rewiring.

Introduction

Synaptic connection between neurons is the fundamental substrate for learning and computation in neural circuits. Previous morphological studies suggest that in cortical microcircuits, often several synaptic connections are found between the presynaptic axons and the postsynaptic dendrites of two connected neurons (Deuchars et al., 1994; Markram et al., 1997; Feldmeyer et al., 1999). Recent connectomics studies confirmed these observations in somatosensory(Kasthuri et al., 2015) and visual(Lee et al., 2016) cortex, and also in hippocampus (Bartol et al., 2015). In particular, in barrel cortex, the average number of synapses per connection is estimated to be around 10(Markram et al., 2015). However, the functional importance of multisynaptic connections remains unknown. Especially, from a computational perspective, such redundancy in connection structure is potentially harmful for learning due to degeneracy (Watanabe, 2001; Amari et al., 2006). In this work, we study how neurons perform learning with multisynaptic connections and whether redundancy provides any benefit, from a Bayesian perspective.

Bayesian framework has been established as a candidate principle of information processing in the brain (Knill and Pouget, 2004; Körding and Wolpert, 2006). Many results further suggest that not only computation, but learning process is also near optimal in terms of Bayesian for given stream of information (Behrens et al., 2007; Lake et al., 2015; Madarasz et al., 2016), yet its underlying plasticity mechanism remains largely elusive. Previous theoretical studies revealed that Hebbian-type plasticity rules eventually enable neural circuits to perform optimal computation under appropriate normalization (Soltani and Wang, 2010; Nessler et al., 2013). However, these rules are not optimal in terms of learning, so that the learning rates are typically too slow to perform learning from a limited number of observations. Recently, some learning rules are proposed for rapid learning (Aitchison and Latham, 2014; Gütig, 2016), yet their biological plausibility are still disputable. Here, we propose a novel framework of non-parametric near-optimal learning using multisynaptic connections. We show that neurons can exploit the variability among synapses in a multisynaptic connection to accurately estimate the causal relationship between pre- and postsynaptic activity. The learning rule is first derived for a simple neuron model, and then implemented in a detailed single neuron model. The derived rule is consistent with many known properties of dendritic plasticity and synaptic organization. Furthermore, the model reveals potential functional roles of anti-Hebbian synaptic plasticity observed in distal dendrites (Letzkus et al., 2006; Sjöström and Häusser, 2006), and benefits of task-dependent dendritic synaptogenesis (Yang et al., 2009; Xu et al., 2009).

Results

A conceptual model of learning with multisynaptic connections

Let us first consider a model of two neurons connected with K numbers of synapses (Fig. 1A) to illustrate the concept of the proposed framework. In the model, synaptic connections from the presynaptic neuron are distributed on the dendritic tree of the postsynaptic neuron as observed in experiments (Markram et al., 1997; Feldmeyer et al., 1999). Although a cortical neuron receives synaptic inputs from several thousands of presynaptic neurons in reality, here we consider the simplified model to illustrate the conceptual novelty of the proposed framework. More realistic models will be studied in following sections.

The synapses generate different amplitudes of excitatory postsynaptic potentials at the soma mainly through two mechanisms. First, the amplitude of dendritic attenuation varies from synapse to synapse, because the distances from the soma are different (Stuart and Spruston, 1998; Segev and London, 2000). Let us denote this dendritic position dependence of synapse k as v_k, and call it as the unit EPSP, because v_k corresponds to the somatic potential caused by a unit conductance change at the synapse (i.e. somatic EPSP per AMPA receptor). As depicted in Figure 1A, unit EPSP v_k takes a small (large) value on a synapse at a distal (proximal) position on the dendrite. The second factor is the amount of AMPA receptors in the corresponding spine, which is approximately proportional to its spine size (Matsuzaki et al., 2004). If we denote this spine size factor as g_k, the somatic EPSP caused by a synaptic input through synapse k is written as $w_k = g_k v_k$. This means that even if the synaptic contact is made at a distal dendrite (i.e. even if v_k is small), if the spine size g_k is large, a synaptic input through synapse k has a strong impact at the soma (e.g. red synapse in Fig. 1A) or vice versa (e.g. cyan synapse in Fig. 1A).

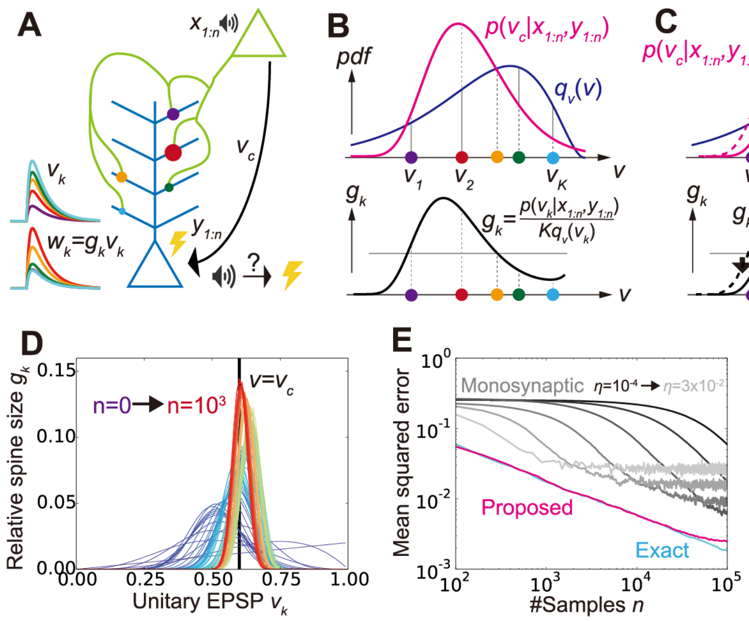

Figure 1. A conceptual model of multisynaptic learning

A) Schematic figure of the model consist of two neurons connected with K synapses. Curves on the left represent unit EPSP v_k (top) and the weighted EPSP $w_k=g_k v_k$(bottom) of each synaptic connection. Note that synapses are consistently colored throughout Figure 1 and 2. **B)** Schematics of non-parametric representation of the probability distribution by multisynaptic connections. In both graphs, x-axes are unit EPSP, and the left (right) side corresponds to distal (proximal) dendrite. The mean over the true distribution $p(v_c|x_{1:n},y_{1:n})$ can be approximately calculated by taking samples (i.e. synapses) from the unit EPSP distribution $q_v(v)$ (top), and then taking a weighted sum over the spine size factor g_k representing the ratio $p(v_c|x_{1:n},y_{1:n})/q_v(v)$ (bottom). **C)** Illustration of synaptic weight updating. When the distribution $p(v_c|x_{1:n+1},y_{1:n+1})$ comes to the right side of the original distribution $p(v_c|x_{1:n},y_{1:n})$, a synaptic weight g_k^{n+1} become larger (smaller) than g_k^n at proximal (distal) synapses. **D)** An example of learning dynamics at K=100 and q_v(v)=const. Each curve represents the distribution of relative spine size $\{g_k\}$, and the colors represent the growth of trial number. **E)** Comparison of performance among the proposed method, the monosynaptic rule, and the exact solution (see *A conceptual model of multisynaptic learning* in Methods for details). The monosynaptic learning rule was implemented with $\eta=10^{-4}$, $3\text{x}10^{-4}$, 10^{-3}, $3\text{x}10^{-3}$, 10^{-2}, $3\text{x}10^{-2}$, and the initial value was taken as $v_{rm}^0=1/2$. Lines were calculated by taking average over 100 independent simulations.

On this model, we consider a simplified fear-conditioning task as an example. Here, the presynaptic neuron activity represents a tone stimulus ($x_n \in \{0,1\}$), and the postsynaptic neuron activity represents an electric shock ($y_n \in \{0,1\}$), where $x_n = 1$ ($y_n = 1$) denotes the presence of the tone (shock), and subscript n stands for the trial number (Fig. 1A). In order to invoke appropriate fear responses, synaptic connections need to acquire the probability of the shock given the tone $v_c \equiv p(y_n = 1 | x_n = 1)$ (Madarasz et al., 2016). Below, we consider supervised learning of this parameter v_c by multisynaptic connections, from the tone and the shock stimuli represented by pre and postsynaptic activities respectively. From finite trials up to n, this conditional probability is estimated as $\bar{v}_c^n = \int v'_c p(v'_c | x_{1:n}, y_{1:n}) dv'_c$, where $x_{1:n}=\{x_1,x_2,\ldots,x_n\}$ and $y_{1:n}=\{y_1,y_2,\ldots,y_n\}$ are the histories of input and output activities respectively, and $p(v_c | x_{1:n}, y_{1:n})$ is the probability distribution of the hidden parameter v_c after n trials. Importantly, in general, it is impossible to get the optimal estimation of $\bar{v}_c^{n+1}$ directly from $\bar{v}_c^n$, because in order to calculate $\bar{v}_c^{n+1} = \int v'_c p(v'_c | x_{1:n+1}, y_{1:n+1}) dv'_c$, one first needs to calculate the distribution $p(v_c | x_{1:n+1}, y_{1:n+1})$ by integrating the previous distribution $p(v_c | x_{1:n}, y_{1:n})$ and the new observation at trial $n+1$: $\{x_{n+1}, y_{n+1}\}$. This means that for near-optimal learning, synaptic connections need to learn and represent the distribution $p(v_c | x_{1:n}, y_{1:n})$ instead of the point estimation $\bar{v}_c^n$. But, how can synapses achieve that? The key hypothesis of this paper is that redundancy in synaptic connections is the substrate for the non-parametric representation of this probabilistic distribution. Below, we show that dendritic summation over multisynaptic connections yields the optimal estimation from the given distribution $p(v_c | x_{1:n}, y_{1:n})$, and dendritic-position-dependent Hebbian synaptic plasticity updates this distribution.

Dendritic summation as importance sampling

We first consider how dendritic summation achieves the calculation of the mean conditional probability $\bar{v}_c^n = \int v'_c p(v'_c | x_{1:n}, y_{1:n}) dv'_c$. It is generally difficult to evaluate this integral by directly taking samples from the distribution $p(v_c | x_{1:n}, y_{1:n})$ in a biologically plausible way, because the cumulative distribution changes its shape at every trial. Nevertheless, we can still estimate the mean value by using an alternative distribution as the proposal distribution, and taking weighted samples from it. This method is called importance sampling(Robert and

Casella, 2013). In particular, here we can use the unit EPSP distribution $q_v(v)$ as the proposal distribution, because unit EPSPs $\{v_k\}$ of synaptic connections can be interpreted as samples depicted from the unit EPSP distribution (Fig. 1B top). Thus, the mean $\bar{v}_c^n$ is approximately calculated as

$$\bar{v}_c^n = \int v_c' p\left(v_c' \mid x_{1:n}, y_{1:n}\right) dv_c' \approx \frac{1}{K}\sum_{k=1}^{K} \frac{p\left(v_c = v_k \mid x_{1:n}, y_{1:n}\right)}{q_v\left(v_k\right)} v_k = \sum\nolimits_k g_k^n v_k = \sum\nolimits_k w_k^n, \qquad (1)$$

where $g_k^n = \frac{p\left(v_c = v_k \mid x_{1:n}, y_{1:n}\right)}{K q_v\left(v_k\right)}$. Therefore, if spine size g_k^n represents the relative weight of sample v_k, then dendritic summation over postsynaptic potentials $w_k^n \equiv g_k^n v_k$ naturally represents the desired value ($\bar{v}_c^n \approx \sum\nolimits_k w_k^n$). For instance, if the distribution of synapses is biased toward proximal side (i.e. if the mean $\bar{v}_c^n$ is overestimated by the distribution of unit EPSPs as in Fig. 1B top), then synapses at distal dendrites should possess large spine sizes, while the spine sizes of proximal synapses should be smaller (Fig. 1B bottom).

Synaptic plasticity as particle filtering

In the previous section, we showed that redundant synaptic connections can represent probabilistic distribution $p(v_c=v_k|x_{1:n},y_{1:n})$, if spine sizes $\{g_k\}$ coincide with their importance $g_k^n = \frac{p\left(v_c = v_k \mid x_{1:n}, y_{1:n}\right)}{K q_v\left(v_k\right)}$. But, how can synapses update their representation of the probabilistic distribution $p(v_c=v_k|x_{1:n},y_{1:n})$ based on a new observation $\{x_{n+1}, y_{n+1}\}$? Because $p(v_c=v_k|x_{1:n},y_{1:n})$ is mapped onto a set of spine sizes $\{g_k^n\}$ as in Equation 1, the update of the estimated distribution $p\left(v_k \mid x_{1:n}, y_{1:n}\right) \to p\left(v_k \mid x_{1:n+1}, y_{1:n+1}\right)$ can be performed by the update of spine sizes $\left\{g_k^n\right\} \to \left\{g_k^{n+1}\right\}$. By considering particle filtering(Doucet et al., 2000) on the parameter space (see *The learning rule for multisynaptic connections* in Methods for details), we can derive the learning rule for spine size as

$$g_k^{n+1} = \frac{1 + f\left(x_{n+1}, y_{n+1}; v_k\right)}{1 + f\left(x_{n+1}, y_{n+1}; w^n\right)} g_k^n, \quad f\left(x, y; v\right) \equiv \left(2v - 1\right) x \left(2y - 1\right). \qquad (2)$$

This rule is primary Hebbian, because the weight change depends on the product of pre and postsynaptic activity x_{n+1} and y_{n+1}. In addition to that, the change also depends on unit EPSP v_k. This dependence on unit EPSP reflects the dendritic position dependence of synaptic plasticity. In particular, for a distal synapse (i.e. for small v_k), the position-dependent term

($2v_k$–1) takes a negative value (note that $0 \le v_k < 1$), thus yielding an anti-Hebbian rule as observed in neocortical synapses (Letzkus et al., 2006; Sjöström and Häusser, 2006).

For instance, if the new data $\{x_{n+1}, y_{n+1}\}$ indicates that the value of v_c is in fact larger then previously estimated, then the distribution $p(v_c|x_{1:n+1},y_{1:n+1})$ shifts to the right side (upper panel of Fig. 1C). This means that the spine size g_k^{n+1} becomes larger then g_k^n at synapses on the right side (i.e. proximal side), whereas synapses get smaller on the left side (i.e. distal side; bottom panel of Fig. 1C). Therefore, pre- and postsynaptic activity causes LTP at proximal synapses induces LTD at distal synapses as observed in experiments (Letzkus et al., 2006; Sjöström and Häusser, 2006). The derived learning rule (Eq. 2) also depends on the total EPSP amplitude $w^n \equiv \sum_k w_k^n \equiv \sum_k g_k^n v_k$. This term reflects a normalization factor possibly modulated through redistribution of synaptic vesicles over the presynaptic axon(Staras et al., 2010).

We performed simulations by assuming a uniform spatial distribution for synapses; $q_v(v)$ = const. At an initial phase of learning, the distribution of spine size $\{g_k^n\}$ has a broad shape (purple lines in Fig. 1D), whereas the distribution gets skewed as evidence is accumulated through stochastic pre- and postsynaptic activities (red lines in Fig. 1D). Indeed, the estimation performance of the proposed method is nearly the same as that of the exact optimal estimation, and much better than that of the standard monosynaptic learning rule (Fig. 1E; see *Monosynaptic learning rule* in Methods for details).

Synaptogenesis as resampling

As shown above, weight modification in multisynaptic connections enables a near optimal learning. However, to represent the distribution accurately, many synaptic connections are required (gray line in Fig. 2B), while the number of synapses between a excitatory neuron pair is typically less than 10 in the cortical microcircuits. Moreover, even if many synapses are allocated between presynaptic and postsynaptic neurons, if the unit EPSP distribution is highly biased, the estimation is poorly performed (gray line in Fig. 2C). We next show that this problem can be avoided by introducing synaptogenesis (Holtmaat and Svoboda, 2009) into the learning rule.

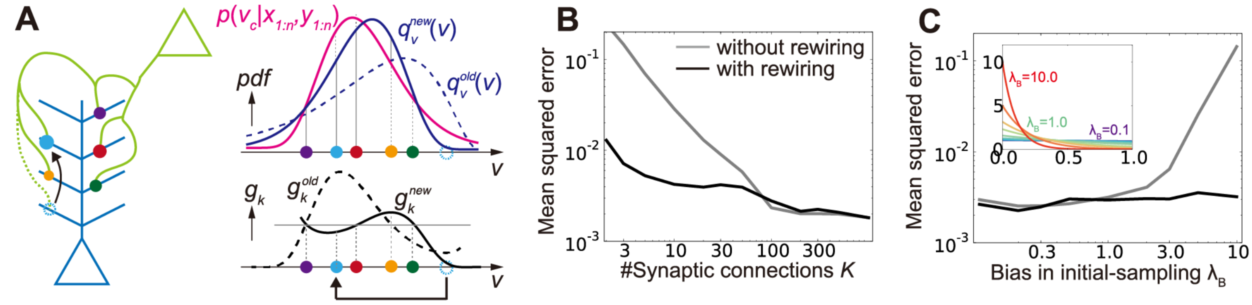

Figure 2. Synaptic rewiring for efficient learning

A) Schematic illustration of resampling. Dotted cyan circles represent an eliminated synapse, and the filled cyan circles represent a newly created synapse. **B, C)** Comparison of performance with/without synaptic rewiring at various synaptic multiplicity K **(B)**, and bias in initial-sampling λ_B **(C)**. For each bias parameter λ_B, the unit EPSP distribution $\{v_k\}$ was set as $v_{k'} = -\log\left(1-\left[1-e^{-\lambda_B}\right]\frac{k'}{K}\right)$, as depicted in the inset. Lines are the means over 100 simulations.

In the proposed framework, when synaptic connections are fixed (i.e. when $\{v_k\}$ are fixed), some synapses quickly become useless for representing the distribution. For instance, in Figure 2A, (dotted) cyan synapse is too proximal to contribute for the representation of $p(v_c|x,y)$. Therefore, by removing the cyan synapse and creating a new synapse at a random site, on average, the representation becomes more effective (Fig. 2A). Importantly, in our framework, synaptic weight is proportional to its informatic importance by definition, thus optimal rewiring is achievable simply by removing the synapse with the smallest spine size. Ideally, the new synapse should be sampled from $p(v_c|x,y)$ for an efficient rewiring, but random creation is more biologically plausible (Holtmaat and Svoboda, 2009), and indeed sufficient as long as elimination is selectively performed as mentioned above (see also Hiratani and Fukai, 2016).

By introducing this resampling process, the model is able to achieve high performance even if the total number of synaptic connection is just around three (black line in Fig. 2B), or if the initial distribution of $\{v_k\}$ is poorly taken (black line in Fig. 2C).

Detailed single neuron model of learning from many presynaptic neurons

In the previous sections, we found that synaptic plasticity in multisynaptic connections can achieve non-parametric near-optimal learning in a simple model with one presynaptic neuron. To investigate its biological plausibility, we next extend the proposed framework to a detailed single neuron model receiving inputs from many presynaptic neurons. To this end, we constructed an active dendritic model using NEURON simulator(Hines and Carnevale, 1997) based on a previous model of a pyramidal neuron in the layer 2/3 (Smith et al., 2013). We randomly distributed 500 excitatory synaptic inputs from 50 presynaptic neurons on the dendritic tree of the postsynaptic neuron, while fixing synaptic connections per presynaptic neuron at K=10 (Fig. 3A; see *Morphology* in Methods for the details of the model). First, we added a small constant conductance for each synapse, and then measured the somatic potential change, which corresponds to unit EPSP in the model. As observed in cortical neurons(Stuart and Spruston, 1998), inputs at distal dendrite tended to show large attenuation at the soma, though variability was quite high (Fig. 3B). The calculated unit EPSP distribution was rather skewed, because more branches were located on distal dendrites (Fig. 3C top). From this distribution, we set the initial weight distribution accordingly (Fig. 3C bottom), as described in Figure 1B.

On this model, we considered a supervised classification task: the neuron should fire if the stochastic presynaptic spikes are generated from the target pattern not from a distractor (Fig. 3D). The supervised signal was stochastically given if the generate presynaptic spike pattern resembled the target pattern (see *Classification task* in Methods). We first applied the proposed synaptic plasticity rule without rewiring(see *The learning rule for the detailed model* in Methods). After a sufficient number of trials, indeed, the neuron learned to show large depolarization only for the target pattern not for distractors (Fig. 3E). Classification performance was better when EPSP was decoded from its height, compared to the decoding from the total EPSP area (Fig. 3F), suggesting an advantage of spike-based information processing with a threshold mechanism.

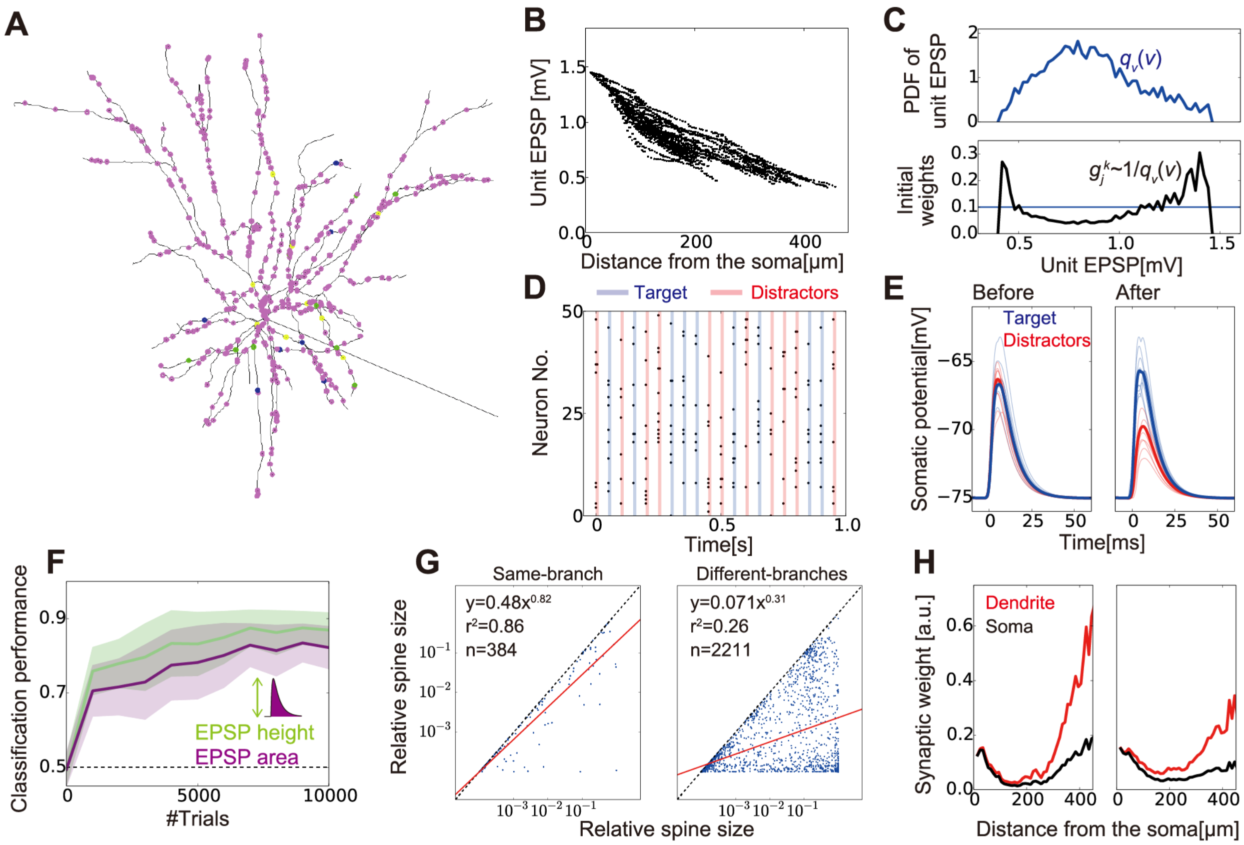

Figure 3. A detailed model of multisynaptic learning with multiple presynaptic neurons **A)** Schematic figure of the detailed neuron model. Colored points on the dendritic trees represent 500 synaptic inputs from 50 presynaptic neurons. Green, yellow, and blue points show examples of initial distributions of inputs from three presynaptic neurons. **B)** Dendritic position dependence of unit EPSP. Each black dot represents a synaptic contact on the dendritic tree. **C)** Unit EPSP distribution (top) and the corresponding initial weights distribution (bottom) for the detailed neuron model. This figure corresponds to Figure 1B. **D)** Examples of input spike trains generated from the target and distractor stimuli. Vertical bars indicate each stimulation trial. Note that in the actual simulations, variables were initialized after each stimulation trial. See *Classification task* in Methods for details of the task configuration. **E)** Somatic membrane dynamics before and after learning. Thick lines represent the average response curves over 100 trials, and thin lines are trial-by-trial responses. **F)** The learning curves calculated from peak EPSP height (light-green line) and EPSP area (purple line). Error bars represent the standard derivations calculated over 50 simulations. **G)** Correlations of relative spine sizes between two synapses projected from the same presynaptic neuron onto the same dendritic branch (left), and onto different dendritic branches (right). Branches longer than 40μm were excluded from the analysis. **H)** Mean

distribution of the relative strength of synapse at its dendritic position g_k (red) and at the soma $g_k v_k / v_{max}$ (black). In the left panel, we used α_B=0.2 as in the rest of the figures, while in the right panel, we used α_B=0.6 to reproduce a less sparse input configuration (see *Input configuration* in Methods for details).

We further investigated consistency with experimental results. As observed in hippocampal neurons (Bartol et al., 2015), when two synaptic connections were made from the same presynaptic neuron to the same dendritic branch, their spine sizes were highly correlated (Fig. 3G left). In contrast, if synapses were connected to different branches of the postsynaptic neuron, spine size correlation vanished almost completely (Fig. 3G right). Furthermore, it is known that distal synaptic connections tend to have larger spine sizes than proximal connections, so that their somatic impacts are nearly the same with proximal synapses in hippocampus, and somewhat smaller in neocortex (Williams and Stuart, 2003). Our model replicated this correlation between the average spine size and the dendritic position (Fig. 3H) due to sparse stimulus representation used in the model. In addition, we found that under a denser stimulus representation, the average spine size of distal synapses became smaller, resulting in a spine size distribution closer to that of cortical neurons than of hippocampal neurons (Fig. 3H right compared to the left). This result is consistent with experimental observations of non-sparse selectivity at cortical pyramidal neurons (Rigotti et al., 2013).

We next changed the shape of distribution of synapses on dendritic tree. As expected, when synapses are biased toward the distal side, the performance improved faster than the opposite case, because the unit EPSP distribution provides a better prior distribution of optimal EPSPs (Fig. 4A). However, the performance after a long training was better when the distribution was skewed toward the proximal side (Fig. 4A), because strong signals are better represented when most synapses are proximal.

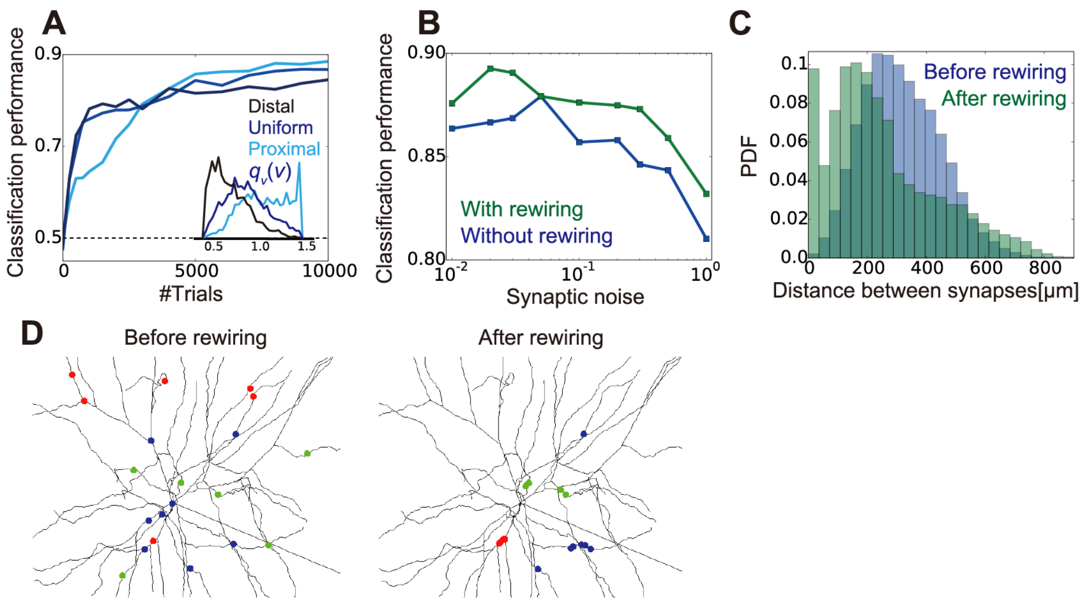

Figure 4. Effects of dendritic synaptic distribution and rewiring in the detailed model
A) Learning curves for different synaptic distributions generated from three values of biased parameter: λ_B=0.1, 1.0, 1.9 (from light blue to black). Note that blue line in the center corresponds to the light-green line in Figure 3F. The inset represents the unit EPSP distributions in the three settings. **B)** Classification performance after 10000 trials with or without rewiring at various synaptic noise levels. **C)** Distributions of distance along the dendrite between two synapses projected from the same presynaptic neuron before (blue) and after (green) 10000 trials of rewiring. **D)** Effect of synaptic rewiring on the dendritic distributions of synapses from the same presynaptic neuron. 30 synaptic contacts from 3 representative presynaptic neurons (color-coded) are depicted. In **B-D**, resampling threshold was set at g_{th}=0.0001, and the potential location of newly created spines were limited to the dendritic branches to which the corresponding presynaptic neuron initially projected, because typically presynaptic axons and postsynaptic dendrites have a limited number of close contacts (see *Details of the NEURON simulations* in Methods for details of rewiring). All data points in **A-C** were calculated by taking average over 50 independent simulations.

Finally, we investigated effects of synaptic rewiring. Here we introduced synaptic noise that reflects stochastic signal transmission at individual synapses (see *Details of the NEURON simulations* in Methods). Although rewiring did not improve performance much except in the regime of high synaptic noise (Fig. 4B), we found that the rewiring causes clustering of synapses from the same presynaptic neuron (Fig. 4D), and reduces the mean dendritic distance between two synapses projected from the same presynaptic neuron (Fig. 4C). This result suggests that the clustering of synaptic contacts made by the same presynaptic neuron observed in adult neocortex (Kasthuri et al., 2015; Lee et al., 2016) could be the result of developmental synaptogenesis.

Recurrent circuit model of unsupervised learning

Results so far demonstrated that the proposed learning rule works efficiently in supervised learning tasks. However, supervised signals are often not available in the actual brain. Hence, we next show that the proposed framework is also applicable for unsupervised learning in recurrent networks in which lateral interactions enable self-organized learning. Let us consider acquisition of stochastic sequences by mutually connected firing-rate units (Fig. 5A). In the model, external states are updated according to a hidden Markov model (Fig. 5A top). The chosen state is observed and represented stochastically by a layer of binary units (Fig. 5A middle), and then their outputs project to the recurrent network (Fig. 5A bottom). Here, we used an all-to-all recurrent network in which each neuron-to-neuron connection is realized by five synapses. The task is to infer both parameters and states of the hidden Markov process from a given sequence of observations. On this task, we implemented our multisynaptic learning rule, as well as previously proposed rules (Rabiner, 1989; Mongillo and Deneve, 2008) for comparison (see *Recurrent circuit model* in Methods for details).

Although the learning performance of the proposed rule was not as good as that of the batch EM algorithm (Rabiner, 1989) due to lack of memory, still the proposed rule performed better than a stochastic gradient descent (SGD) rule during the most epochs of learning (Fig. 5B). A virtue of the proposed rule is that it does not require fine-tuning of learning rate unlike SGD learning, or online EM algorithm (Mongillo and Deneve, 2008)(Fig. 5C).

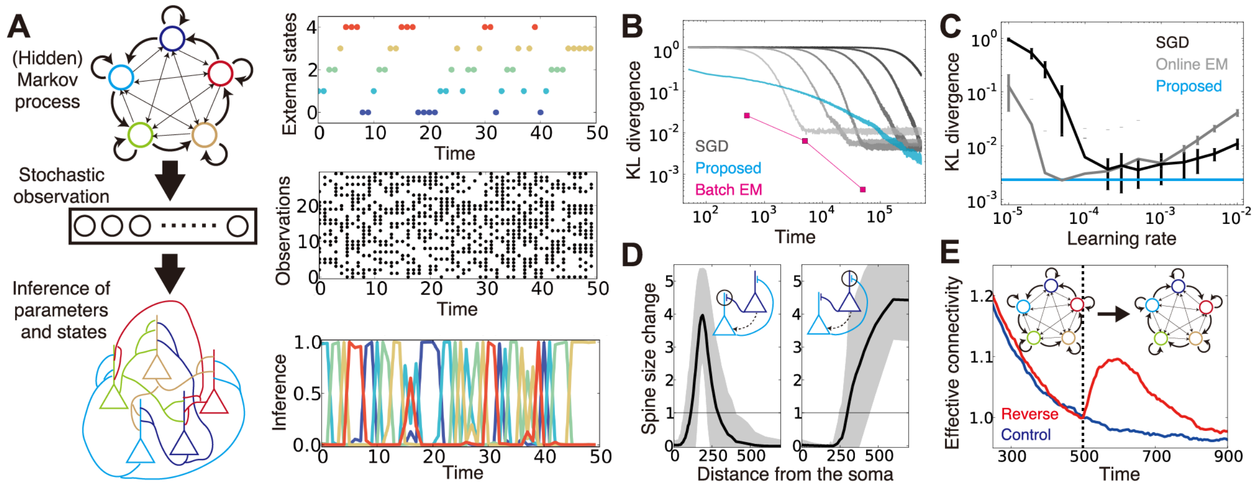

Figure 5. Unsupervised learning of probabilistic sequences by a recurrent network with multisynaptic connections

A) Schematic illustration of the model organization (left) and examples of model behaviors (right). External states generated from a hidden Markov model (top) are observed by stochastic binary units (middle), which in turn projected to a recurrent circuit in which every unit pair is connected with five synapses (bottom). **B, C)** Performance comparisons of the model and other methods at different times (trial numbers)**(B)** and learning rates**(C)**. In **C**, KL divergence were calculated at t=5x10^5, and error bars are standard deviations calculated over trials. **D)** Relationship between spine size change and the distance from the soma, at a synapse where presynaptic input and postsynaptic firing having strong causal relationship (left), or anti-causal relationship (right). Here, rewiring was turned off to observe the pure effect of synaptic plasticity alone. Spine sizes were calculated at t=2500, and the distance from the soma was artificially determined by setting the characteristic length as 200μm. **E)** The influence of retraining on effective connectivity. We changed the hidden Markov process from a counter-clockwise dynamics to clockwise dynamics after 500 trials (inset). The effective connectivity was calculated by discarding the silent synapses with spine size g_k < 0.01. Averages were taken over 50 **(B, C)**, and 5000 **(D,E)** simulations. In **D** and **E**, the feedforward projections were fixed, and only the recurrent connections were learned. See *Recurrent circuit model* in Methods for further details.

We further looked into correspondence with the experimental results. In burst-dependent STDP, distal synapses show an anti-Hebbain STDP time window, while proximal synapses show the ordinary Hebbian STDP time window (Letzkus et al., 2006). Correspondingly, in the model, when the pre- and postsynaptic neurons had causal relationship, synapses on proximal dendrites were more likely potentiated compared to distal synapses (Fig. 5D left). This was because large unit EPSPs (i.e. proximal synapses) are preferred when the total synaptic weight should be large. In contrast, when two neurons had an anti-causal relationship, distal synapses were more likely potentiated (Fig. 5D right). Secondly, it is known that novel training increases the observed number of spines in task-related neurons (Yang et al., 2009; Xu et al., 2009). We forced novel learning on the network model by reversing the order of transitions among hidden states (Fig. 5E inset). Our model showed a similar increase in the number of effective connectivity right after the changes in the external environment (Fig. 5E). This is because, when the input structure changes due to novel training, the distributions of possible parameter values become broader, as a result, previously silent synapses are employed for representing these wider probabilistic distributions.

Discussion

In this work, first we have used a simple conceptual model to show: (i) Multisynaptic connections provide a non-parametric representation of probabilistic distribution of the hidden parameter using redundancy in synaptic connections (Fig. 1AB); (ii) Updating of probabilistic distribution given new inputs can be performed by a Hebbian-type synaptic plasticity when the output activity is supervised (Fig. 1C-E); (iii) Elimination and creation of spines is crucial for efficient representation and fast learning (Fig. 2A-C). In short, synaptic plasticity and rewiring at multisynaptic connections naturally implements an efficient sample-based Bayesian filtering algorithm. Secondly, we have demonstrated that the proposed multisynaptic learning rule works well in a detailed single neuron model receiving stochastic spikes from many neurons (Fig. 3). Moreover, the model suggests that the dendritic distribution of multisynaptic inputs provides a prior distribution of the expected synaptic weight (Fig. 4A), and rewiring of synaptic connection supports robust information processing under synaptic noise (Fig. 4B). We have further extended the framework for unsupervised learning in recurrent circuits (Fig. 5A-C), and then shown that the model reproduces the experimentally known dendritic position dependences of plasticity, including anti-Hebbian plasticity at distal dendrites (Fig. 5D).

Distribution of multisynaptic projections

Our study provides several experimentally testable predictions on dendritic synaptic plasticity, and the resultant synaptic distribution. First, the model suggests developmental convergence of synaptic connections from each presynaptic neuron (Fig. 4CD). It is indeed known that in adult cortex, synaptic connections from the same presynaptic neuron are often clustered (Kasthuri et al., 2015; Lee et al., 2016). Our model interprets synaptic clustering as a result of an experience-dependent resampling process by synaptic rewiring, and predicts that synaptic connections are less clustered in immature animal. In addition, although the model does not provide direct insights on dendritic clustering of inputs from different presynaptic neurons (Takahashi et al., 2012), our results indicate that if two presynaptic inputs are tightly correlated with each other, these presynaptic neurons are more likely to make synaptic contacts on similar positions on the dendritic tree.

Our result also suggests that position on the dendritic tree acts as a prior of the expected total connection strength, and supports rapid acquisition of desired synaptic weights. (Fig. 4A). For instance, primary inputs to the postsynaptic neuron should be proximal, since these inputs are typically expected to have stronger impacts on the soma than modulatory inputs. This is consistent with the synaptic organization on the dendrite of pyramidal cell where primary inputs are often projected to proximal dendrite, while modulatory inputs are typically more distal (Bittner et al., 2015; Manita et al., 2015).

Anti-Hebbian plasticity at distal synapse (Letzkus et al., 2006; Sjöström and Häusser, 2006) can be interpreted in a similar way. Modulatory inputs are typically not tightly correlated with the output spike trains, because these inputs usually carry contextual information(Bittner et al., 2015), or delayed feedback signals(Manita et al., 2015). Hence, anti-Hebbian plasticity at distal synapse potentially helps neurons to select appropriate modulatory inputs.

Related works

Previous theoretical studies often explain synaptic plasticity as stochastic gradient descent on some objective functions (Pfister et al., 2006; Nessler et al., 2013; Urbanczik and Senn, 2014; Hiratani and Fukai, 2016), but these models require fine-tuning of the learning rate for explaining near-optimal learning performance observed in humans (Behrens et al., 2007; Lake et al., 2015) and rats (Madarasz et al., 2016), unlike our model. Moreover, in this study, we proposed synaptic dynamics during learning as a sample-based inference process, in contrast to previous studies in which sample-based interpretations were applied for neural dynamics(Orbán et al., 2016).

On the anti-Hebbian plasticity at distal synapse, previous modeling studies have revealed its potential phenomenological origins (Graupner and Brunel, 2012), but its functional benefits, especially optimality, have not been well investigated before. Particle filtering is an established method in machine learning (Doucet et al., 2000), and has been applied to artificial neural networks (Freitas et al., 2000), yet its biological correspondence had been elusive.

Previous computational studies on dendritic computation have been emphasizing the importance of active dendritic process (Segev and London, 2000), especially for performing inference from correlated inputs (Ujfalussy et al., 2015), or for computation at terminal tufts of cortical layer 5 or CA1 neurons (Urbanczik and Senn, 2014). Nevertheless, experimental studies suggest the summation of excitatory input through dendritic tree is approximately linear (Cash and Yuste, 1999; Hao et al., 2009). Indeed, we have shown that a linear summation of synaptic inputs is suitable for implementing importance sampling. Moreover, we have demonstrated that even in a detailed neuron model with active dendrites, a learning rule assuming a linear synaptic summation works well.

Author Contributions

NH and TF conceived the study, NH designed and performed the modeling, NH and TF wrote the manuscript.

Acknowledgements

The authors thank to Peter Latham for discussions and comments on the manuscript. This work was partly supported by CREST, JST (JPMJCR13W1 to TF) and Grants-in-Aid for Scientific Research (KAKENHI) from MEXT (no 15H04265 and 16H01289 to TF).

References

Aitchison, L., and Latham, P.E. (2014). Bayesian synaptic plasticity makes predictions about plasticity experiments in vivo. ArXiv14101029 Q-Bio.

Amari, S., Park, H., and Ozeki, T. (2006). Singularities Affect Dynamics of Learning in Neuromanifolds. Neural Comput. *18*, 1007–1065.

Bartol, T.M., Bromer, C., Kinney, J.P., Chirillo, M.A., Bourne, J.N., Harris, K.M., and Sejnowski, T.J. (2015). Nanoconnectomic upper bound on the variability of synaptic plasticity. eLife e10778.

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. *10*, 1214–1221.

Bittner, K.C., Grienberger, C., Vaidya, S.P., Milstein, A.D., Macklin, J.J., Suh, J., Tonegawa, S., and Magee, J.C. (2015). Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. Nat. Neurosci. *18*, 1133–1142.

Cash, S., and Yuste, R. (1999). Linear Summation of Excitatory Inputs by CA1 Pyramidal Neurons. Neuron *22*, 383–394.

Churchland, A.K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., and Shadlen, M.N. (2011). Variance as a Signature of Neural Computations during Decision Making. Neuron *69*, 818–831.

Deuchars, J., West, D.C., and Thomson, A.M. (1994). Relationships between morphology and physiology of pyramid-pyramid single axon connections in rat neocortex in vitro. J. Physiol. *478*, 423–435.

Doucet, A., Godsill, S., and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. Stat. Comput. *10*, 197–208.

Everett, B. (2013). An Introduction to Latent Variable Models (Springer Science & Business Media).

Feldmeyer, D., Egger, V., Lübke, J., and Sakmann, B. (1999). Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single "barrel" of developing rat somatosensory cortex. J. Physiol. *521*, 169–190.

Freitas, J.F.G. de, Niranjan, M., Gee, A.H., and Doucet, A. (2000). Sequential Monte Carlo Methods to Train Neural Network Models. Neural Comput. *12*, 955–993.

Graupner, M., and Brunel, N. (2012). Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. Proc. Natl. Acad. Sci. *109*, 3991–3996.

Gütig, R. (2016). Spiking neurons can discover predictive features by aggregate-label learning. Science *351*, aab4113.

Hao, J., Wang, X., Dan, Y., Poo, M., and Zhang, X. (2009). An arithmetic rule for spatial summation of excitatory and inhibitory inputs in pyramidal neurons. Proc. Natl. Acad. Sci. *106*, 21906–21911.

Hines, M.L., and Carnevale, N.T. (1997). The NEURON Simulation Environment. Neural Comput. *9*, 1179–1209.

Hiratani, N., and Fukai, T. (2016). Hebbian Wiring Plasticity Generates Efficient Network Structures for Robust Inference with Synaptic Weight Plasticity. Front. Neural Circuits *10*, 41.

Holtmaat, A., and Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. Nat. Rev. Neurosci. *10*, 647–658.

Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, et al., 2015. Saturated Reconstruction of a Volume of Neocortex. Cell 162, 648–661. doi:10.1016/j.cell.2015.06.054

Knill, D.C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. *27*, 712–719.

Körding, K.P., and Wolpert, D.M. (2006). Bayesian decision theory in sensorimotor control. Trends Cogn. Sci. *10*, 319–326.

Lake, B.M., Salakhutdinov, R., and Tenenbaum, J.B. (2015). Human-level concept learning through probabilistic program induction. Science *350*, 1332–1338.

Lee, W.-C.A., Bonin, V., Reed, M., Graham, B.J., Hood, G., Glattfelder, K., and Reid, R.C.

(2016). Anatomy and function of an excitatory network in the visual cortex. Nature *532*, 370–374.

Letzkus, J.J., Kampa, B.M., and Stuart, G.J. (2006). Learning Rules for Spike Timing-Dependent Plasticity Depend on Dendritic Synapse Location. J. Neurosci. *26*, 10420–10429.

Madarasz, T.J., Diaz-Mataix, L., Akhand, O., Ycu, E.A., LeDoux, J.E., and Johansen, J.P. (2016). Evaluation of ambiguous associations in the amygdala by learning the structure of the environment. Nat. Neurosci. 19, 965–972.

Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., Ota, K., Matsubara, C., Inutsuka, A., Sato, M., et al. (2015). A Top-Down Cortical Circuit for Accurate Sensory Perception. Neuron *86*, 1304–1316.

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.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M.W., Abdellah, M., et al., 2015. Reconstruction and Simulation of Neocortical Microcircuitry. Cell 163, 456–492. doi:10.1016/j.cell.2015.09.029

Matsuzaki, M., Honkura, N., Ellis-Davies, G.C.R., and Kasai, H. (2004). Structural basis of long-term potentiation in single dendritic spines. Nature *429*, 761–766.

Mongillo, G., and Deneve, S. (2008). Online Learning with Hidden Markov Models. Neural Comput. *20*, 1706–1716.

Nessler, B., Pfeiffer, M., Buesing, L., and Maass, W. (2013). Bayesian Computation Emerges in Generic Cortical Microcircuits through Spike-Timing-Dependent Plasticity. PLoS Comput Biol *9*, e1003037.

Orbán, G., Berkes, P., Fiser, J., and Lengyel, M. (2016). Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex. Neuron *92*, 530–543.

Pfister, J.-P., Toyoizumi, T., Barber, D., and Gerstner, W. (2006). Optimal Spike-Timing-Dependent Plasticity for Precise Action Potential Firing in Supervised Learning. Neural Comput. *18*, 1318–1348.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE *77*, 257–286.

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.

Robert, C., and Casella, G. (2013). Monte Carlo Statistical Methods (Springer Science & Business Media).

Segev, I., and London, M. (2000). Untangling Dendrites with Quantitative Models. Science *290*, 744–750.

Sjöström, P.J., and Häusser, M. (2006). A Cooperative Switch Determines the Sign of Synaptic Plasticity in Distal Dendrites of Neocortical Pyramidal Neurons. Neuron *51*, 227–238.

Smith, S.L., Smith, I.T., Branco, T., and Häusser, M. (2013). Dendritic spikes enhance stimulus selectivity in cortical neurons in vivo. Nature *503*, 115–120.

Soltani, A., and Wang, X.-J. (2010). Synaptic computation underlying probabilistic inference. Nat. Neurosci. *13*, 112–119.

Staras, K., Branco, T., Burden, J.J., Pozo, K., Darcy, K., Marra, V., Ratnayaka, A., and Goda, Y. (2010). A Vesicle Superpool Spans Multiple Presynaptic Terminals in Hippocampal Neurons. Neuron *66*, 37–44.

Stuart, G., and Spruston, N. (1998). Determinants of Voltage Attenuation in Neocortical Pyramidal Neuron Dendrites. J. Neurosci. *18*, 3501–3510.

Takahashi, N., Kitamura, K., Matsuo, N., Mayford, M., Kano, M., Matsuki, N., and Ikegaya, Y. (2012). Locally Synchronized Synaptic Inputs. Science *335*, 353–356.

Tsubo, Y., Isomura, Y., and Fukai, T. (2012). Power-Law Inter-Spike Interval Distributions Infer a Conditional Maximization of Entropy in Cortical Neurons. PLOS Comput Biol *8*, e1002461.

Ujfalussy, B.B., Makara, J.K., Branco, T., and Lengyel, M. (2015). Dendritic nonlinearities are tuned for efficient spike-based computations in cortical circuits. eLife e10056.

Urbanczik, R., and Senn, W. (2014). Learning by the Dendritic Prediction of Somatic Spiking. Neuron *81*, 521–528.

Watanabe, S. (2001). Algebraic Analysis for Nonidentifiable Learning Machines. Neural Comput. *13*, 899–933.

Williams, S.R., and Stuart, G.J. (2003). Role of dendritic synapse location in the control of action potential output. Trends Neurosci. *26*, 147-154.

Xu, T., Yu, X., Perlik, A.J., Tobin, W.F., Zweig, J.A., Tennant, K., Jones, T., and Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. Nature *462*, 915-919.

Yang, G., Pan, F., and Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. Nature *462*, 920-924.

Methods

A conceptual model of multisynaptic learning

The learning rule for multisynaptic connections

In the model, tone stimulus and electric shock were represented by binary variables $x_n \in \{0,1\}$ and $y_n \in \{0,1\}$. At each trial *n*, tone was delivered with $\Pr[x_n = 1] = \pi_x$, and electric shock was given only when x_n=1, with probability $\Pr[y_n = 1 | x_n = 1] = v_c$. For this task, the update rule for the spine size factor $g_k^{n+1} = \frac{1}{Kq_v(v_k)} p(v_c = v_k | x_{1:n+1}, y_{1:n+1})$ is given as,

$$\begin{aligned} g_k^{n+1} &= \tfrac{1}{Kq_v(v_k)} p(v_c = v_k | x_{1:n+1}, y_{1:n+1}) \\ &\propto \tfrac{1}{Kq_v(v_k)} p(x_{n+1}, y_{n+1} | v_c = v_k) p(v_c = v_k | x_{1:n}, y_{1:n}) \\ &\propto p(y_{n+1} | x_{n+1}, v_c = v_k) \left(\tfrac{1}{Kq_v(v_k)} p(v_c = v_k | x_{1:n}, y_{1:n}) \right) \\ &= p(y_{n+1} | x_{n+1}, v_c = v_k) g_k^n. \end{aligned}$$

In particular, in our problem setting, v_c does not provide any information about y_n when x_n=0, thus approximately (see Supplementary Information for the proof of convergence),

$$\begin{aligned} p(y_{n+1} | x_{n+1}, v_c = v_k) &\approx x_{n+1}\left[v_k y_{n+1} + (1-v_k)(1-y_{n+1})\right] + \tfrac{1}{2}(1-x_{n+1}) \\ &\propto 1 + (2v_k - 1) x_{n+1} (2y_{n+1} - 1). \end{aligned}$$

Because the normalization factor is determined by

$$1 = \int p(v_c' | x_{1:n}, y_{1:n}) dv_c' \approx \frac{1}{K} \sum_k \frac{p(v_c' = v_k | x_{1:n}, y_{1:n})}{q_v(v_k)} = \sum_k g_k^n,$$

the sum of $\{g_k^{n+1}\}$ should also be normalized to 1. Thus the update rule is given as

$$g_k^{n+1} = \frac{\left[1 + f(x_{n+1}, y_{n+1}; v_k)\right] g_k^n}{\sum_{k'} \left[1 + f(x_{n+1}, y_{n+1}; v_{k'})\right] g_{k'}^n} = \frac{1 + f(x_{n+1}, y_{n+1}; v_k)}{1 + f(x_{n+1}, y_{n+1}; w^n)} g_k^n,$$

where $f(x, y; v) \equiv (2v-1)x(2y-1)$ and $w^n \equiv \sum_k w_k^n = \sum_k g_k^n v_k$. As for the resampling process, at every trial *n*, if spine *k* satisfied $g_k < g_{th}$, unit EPSP was resampled uniformly from [0,1), and the spine size was set at $g_k = g_{th}$.

Monosynaptic learning rule

For comparison, we implemented a monosynaptic learning rule, by expanding the exact solution $\bar{v}_c^n = \sum_{n'} x_{n'} y_{n'} / \sum_{n'} x_{n'}$ as

$$\bar{v}_c^n = \left(x_n y_n + \sum_{n'=1}^{n-1} x_{n'} y_{n'}\right) \Big/ \left(x_n + \sum_{n'=1}^{n-1} x_{n'}\right) \approx \bar{v}_c^{n-1} \left(1 + x_n \left(y_n - \bar{v}_c^{n-1}\right) \Big/ \sum_{n'=1}^{n-1} x_{n'} y_{n'}\right).$$

Hence, by using a single variable v_{rm}^n, the learning rule is given as $v_{rm}^n = v_{rm}^{n-1}\left(1 + \eta x_n \left(y_n - \bar{v}_{rm}^{n-1}\right)\right)$, where η represents the learning rate (Nessler et al., 2013).

Details of the conceptual model

In the simulations, we used π_x=0.3, and v_c was randomly chosen from [0,1) uniformly at each simulation (not at each trial). The number of connections was kept at K=100 except for Figure 2B in which K=2 to 1000 were used. Initial value of k-th connection v_k was set as $v_k = k/K$ except for Figure 2C in which the initial distribution was biased by choosing v_k as $v_{k'} = -\log\left(1-\left[1-e^{-\lambda_B}\right]\frac{k'}{K}\right)$ where λ_B is the bias parameter. Resampling was performed with the threshold g_{th}=0.0001, and a new unit EPSP v_k was uniformly sampled from [0,1).

Detailed single neuron model

Morphology

We constructed a detailed neuron model based on a model of layer 2/3 pyramidal neuron with active dendrites (Smith et al., 2013) using NEURON simulator (Hines and Carnevale, 1997). We distributed 500 excitatory synaptic inputs from 50 presynaptic neurons randomly on the dendrite. Synaptic input was modeled as a double exponential conductance change with the rise time τ_{rise}=0.5ms and the decay time τ_{decay}=2.5ms. For each synapse k from presynaptic neuron j, we first applied a synaptic input with a constant weight factor γ_g=1.5nS, and then determined the unit EPSP v_j^k of synapse k by measuring somatic membrane potential change. In the simulation of the task, using malleable spine size factor g_j^k, we set the weight factor of synapse k as $\gamma_g g_j^k$.

Classification task

Using this neuron model, we considered learning of pattern classification. In particular, here we defined the problem as an acquisition of a latent variable model (Everett, 2013). We first constructed a discrete latent variable model $p(y^t) = \sum_{j=1}^{M} p(y^t \mid x^t = x_j)\, p(x^t = x_j)$ where $x \in \{x_1, \ldots, x_M\}$ and $y \in \{0,1\}$ represent the input and the output variable, M is the total number of presynaptic neurons, and t represents trial number. Suppose the firing rate of presynaptic neurons r_j^t represent $p(x^t=x_j)$, and the total synaptic weight from presynaptic neuron j satisfies $w_j \propto p(y=1 \mid x=x_j)$, then dendritic summation over all presynaptic inputs naturally reflects the probability $p(y^t=1)$ (i.e. $\sum_j w_j r_j^t \propto p(y^t=1)$). Let us define the true latent model (or the target of learning) as $w_j^{trg} \equiv p(y=1 \mid x=x_j)$, and the distribution of the hidden variable x at trial t (i.e. the firing rate distribution of input neurons at trial t) as $\rho_j^t \equiv p(x^t=x_j)$.

In this configuration, the task can be defined as the acquisition of the target model $\{w_j^{trg}\}$ from presynaptic spikes generated from $\{\rho^t_j\}$ and the stochastic teaching signal y^t given by $\Pr[y^t=1]=\sum_{j=1}^{M} w_j^{trg}\rho_j^t$. When input signal $\{\rho^t_j\}$ is generated from the target model $\{w_j^{trg}\}$ as $\rho_j^t \propto w_j^{trg}$, the probability $\Pr[y^t=1]$ typically takes a large value, while the probability becomes small if $\{\rho^t_j\}$ is randomly generated. Thus, we can evaluate learning performance by considering the classification of inputs generated by the target distribution from those generated by other distributions (i.e. distractors), through observation of somatic membrane dynamics. In the simulation, we first performed training of synaptic weights by presenting stimuli generated from both the target and the distractors with corresponding stochastic supervised signals. Then, evaluated the performance by comparing the somatic responses for the target stimuli and the distractors.

The learning rule for the detailed model

We next derived the multisynaptic learning rule for this task. By Bayesian filtering,

$$p\left(w_j^{trg}=w_j^k \mid y^{1:t},\rho_j^{1:t}\right) \propto p\left(y^t,\rho_j^t \mid w_j^{trg}=w_j^k, y^{1:t-1},\rho_j^{1:t-1}\right) p\left(w_j^{trg}=w_j^k \mid y^{1:t-1},\rho_j^{1:t-1}\right).$$

Because ρ^t_j does not depend on w_j^{trg} nor previous activities $\{y^{1:t-1}, \rho_j^{1:t-1}\}$,

$$\begin{aligned} p\left(y^t,\rho_j^t \mid w_j^{trg}=w_j^k, y^{1:t-1},\rho_j^{1:t-1}\right) &\propto p\left(y^t \mid w_j^{trg}=w_j^k, \rho_j^t, y^{1:t-1},\rho_j^{1:t-1}\right) \\ &\cong y^t\left[\bar{w}+\left(w_j^k-\bar{w}\right)\rho_j^t\right]+\left(1-y^t\right)\left[\left(1-\bar{w}\right)-\left(w_j^k-\bar{w}\right)\rho_j^t\right], \end{aligned} \tag{3}$$

where $\bar{w}=\left\langle w_j^{trg}\right\rangle_j$. In the neuronal implementation, presynaptic activities are not directly given as the firing rates $\rho_j^t \equiv p\left(x^t=x_j\right)$, but given as spike trains, yet we can still apply this rule by approximating ρ^t_j by the spike count s^t_j as $\rho_j^t \approx s_j^t / A_p$, with amplification factor A_p. More specifically, we have chosen the factor A_p as $A_p = 0.2M$, so that each presynaptic neuron emits 0.2 spikes in each trial on average. Similarly, representation of w^k_j by unit EPSP of the *k*-th synapse v^k_j can be implemented as $w_j^k = \gamma_V\left(v_j^k - v_{\min}\right)$. In the model, v_{min} was set at 0.41mV (Fig. 3B), and γ_v was defined as $\gamma_v \equiv 1/(v_{\max} - v_{\min})$ where v_{max}=1.45mV. Thus, by representing the importance of each sample by the spine size factor as $g_j^{k,t} = p\left(w_j^{trg} = \gamma_v\left(v_j^k - v_{\min}\right) | \, y^{1:t}, \rho_j^{1:t}\right) / K q_v\left(v_j^k\right)$, from Equation (3), the learning rule of $\{g_j^{k,t}\}$ is approximately given

$$g_j^{k,t+1} = \begin{cases} \left(\bar{w}^t + \left[\gamma_v\left(v_j^k - v_{\min}\right) - \bar{w}^t\right] s_j^t / A_p\right) g_j^{k,t} / Z_j^t & \text{(if } y^t = 1) \\ \left(\left[1 - \bar{w}^t\right] - \left[\gamma_v\left(v_j^k - v_{\min}\right) - \bar{w}^t\right] s_j^t / A_p\right) g_j^{k,t} / Z_j^t & \text{(if } y^t = 0), \end{cases}$$

where

$$Z_j^t = \begin{cases} \left(\bar{w}^t + \left[\gamma_v\left(\bar{v}_j^t - v_{\min}\right) - \bar{w}^t\right] s_j^t / A_p\right) & \text{(if } y^t = 1) \\ \left(\left[1 - \bar{w}^t\right] - \left[\gamma_v\left(\bar{v}_j^t - v_{\min}\right) - \bar{w}^t\right] s_j^t / A_p\right) & \text{(if } y^t = 0). \end{cases}$$

Mean values $\{\bar{v}_j^t\}$ and $\bar{w}^t$ were estimated as $\bar{v}_j^t = \sum_{k=1}^{K} g_j^{k,t} v_j^k$ and $\bar{w}^t = \sum_{j=1}^{M} \gamma_v\left(\bar{v}_j^t - v_{\min}\right) / M$.

Input configuration

In the simulations, we first constructed the mean presynaptic spike probabilities $\{\rho_j\}$ for target and distractor stimuli, and then generated input spike trains $\{s^t_j\}$ according to $\{\rho_j\}$. Mean spike probabilities $\{\rho_j^{trg}\}$ for the target stimulus were randomly generated from a Beta distribution as $\rho_j^{trg} = \tilde{\rho}_j^{trg} / Z_{trg}$ where $\tilde{\rho}_j^{trg} \leftarrow Beta(\alpha_B, 1)$ and $Z_{trg} = (\alpha_B + 1)\sum_{j=1}^{M} \tilde{\rho}_j^{trg} / M\alpha_B$,with α_B being the sparseness parameter. Here, Beta distributions were used due to the constraint on $\{\rho_j^{trg}\}$ ($0 \le \rho_j^{trg} < 1$). Mean responses for the distractors $\{\rho_j^{dst,\mu}\}$ (μ=1,…,10) were defined in the same way. Based on these models, we generated presynaptic spikes $\{s^t_j\}$ with a doubly stochastic process to reproduce high variability typically observed in cortical activity (Churchland et al., 2011; Tsubo et al., 2012). In trial *t* with the target stimulus, we determined the number of spikes s_j^t emitted from presynaptic neuron *j* as $s_j^t = \lfloor A_p \rho_j^t + \zeta \rfloor$ where $\rho_j^t = \tilde{\rho}_j^t / \sum_j \tilde{\rho}_j^t$, $\tilde{\rho}_j^t \leftarrow \text{Gamma}\left(\rho_j^{trg}, 1\right)$, ζ is a random variable uniformly sampled from [0,1), and $\lfloor x \rfloor$ is the largest integer smaller than *x*. In a distractor trial, we instead sampled $\tilde{\rho}_j^t$ from one of the distractor distributions $\{\rho_j^{dst,\mu}\}$ as $\tilde{\rho}_j^t \leftarrow \text{Gamma}\left(\rho_j^{dst,\mu}, 1\right)$, and generated $\{s^t_j\}$ from the chosen distribution. In either trial, supervising signal y^t was stochastically given with probability $\Pr\left[y^t=1\right]=\sum_{j=1}^{M} w_j^{trg}\rho_j^t$. Finally, spike timings of the *m*-th spike from presynaptic neuron *j* at trial *t* was determined as $\left(0.1\zeta_G^{j,t}+m-1\right)\Delta t_{stimulus}/s_j^t$ where $\Delta t_{stimulus}$=10ms, and $\zeta_G^{j,t}$ is a Gaussian random variable.

During training phase, the target stimulus $\{\rho_j^{trg}\}$ were presented in 20% of trials, and each distractor $\{\rho_j^{dst,\mu}\}$ (μ=1,...,10) were presented in 8% of trials. In the test phase, we provided 200 stimuli, of which 100 stimuli were generated from $\{\rho_j^{trg}\}$, and the rest were from $\{\rho_j^{dst,\mu}\}$. The classification performance was measured by the ratio of target trials in which the maximum EPSP height Δv_n^{trg} exceeded the threshold $\theta=\left(m_{trg}/\sigma_{trg}^2+m_{dst}/\sigma_{dst}^2\right)/\left(1/\sigma_{trg}^2+1/\sigma_{dst}^2\right)$, to the total of 100 trials, where $m_{trg}=E\left[\Delta v_n^{trg}\right]$ and $\sigma_{trg}^2=\mathrm{Var}\left[\Delta v_n^{trg}\right]$ were calculated over 100 test stimuli. Although the evaluations were made solely on false negatives, we also observed significant decrease of false positives during learning (Fig. 3E). In the purple line of Figure 3F, we used the total EPSP area instead of the maximum EPSP height for the measurement.

Details of the NEURON simulations

Initial values of spine sizes $\{g_j^k\}$ were chosen such that $g_j^k \sim 1/q_v(v_j^k)$ is satisfied. To this end, we first estimated the unit EPSP distribution $q_v(v)$ through a sampled-based approximation $q_v(n)\equiv\sum_{i=1}^{N}\sum_{k=1}^{K}\left[n\cdot dv\le v_j^k<(n+1)\cdot dv\right]_+\propto q_v\left(n\cdot dv\le v_j^k<(n+1)\cdot dv\right)$, and then calculated g_j^k by $g_j^k=\dfrac{1/q_v\left(n_j^k\right)}{\sum_k 1/q_v\left(n_j^k\right)}$, where n_j^k is the integer that satisfies $n_j^k\cdot dv\le v_j^k<\left(n_j^k+1\right)\cdot dv$. In Figure 4A, to generate a biased synaptic distribution, we randomly sampled a position from the whole dendritic tree with probability $\left(\frac{L'}{L_{\max}}\right)^{\lambda_B-1}\cdot\left(\frac{L'}{L_{\max}}\right)^{1-\lambda_B}/10\cdot B\left(\lambda_B,2-\lambda_B\right)$, and added a synapse until 500 synapses are created on the dendritic tree. Here, *L'* is the distance from the soma, L_{max} is its maximum length, λ_B is the bias parameter, and B(x,y) is the Beta function.

In Figures 4B and C, we replicated synaptic noise independent of presynaptic activity by introducing a fluctuation term, $\xi_j^k\leftarrow Gamma\left(1/s_{noise},s_{noise}\right)$, into the spine-size factor as $\gamma_g g_j^k\to\gamma_g g_j^k\xi_j^k$. The threshold for synaptic elimination was set as g_{th} =0.0001, and the spine size of the new synapse was initialized at $g_k = g_{th}$. In the detailed model, we restricted the position of newly created synapse within the dendritic branches for which the presynaptic neuron was initially projected to. Thus, in the presented simulations, one presynaptic neuron can make synaptic contact with at most on 10 dendritic branches. We introduced this restriction in order to reproduce limited number of close contacts between the axons and the dendrite (Markram et al., 1997; Feldmeyer et al., 1999). Further details of the model are available at ModelDB (http://modeldb.yale.edu/225075 with access code "1234").

Recurrent circuit model

In the model, the hidden Markov model had five states, and the transition probabilities among them $a_{\mu\nu} \equiv p\left(x^t = \mu \mid x^t = \nu\right)$ were defined as $a_{\mu\nu}=0.5$ if $\mu=\nu$, $a_{\mu\nu}=0.4$ if $\mu=\nu+1$ (mod. p), and $a_{\mu\nu}=0.333$ otherwise. In the latter half of Figure 5E, we instead set as $a_{\mu\nu}=0.4$ if $\mu=\nu-1$ (mod. p). Stochastic observation $\{y^t\}$ was defined as $p\left(y^t \mid x^t = \mu\right) = \prod_{i=1}^{N} p\left(y_i^t \mid x^t = \mu\right) \equiv \prod_{i=1}^{N} h\left(y_i^t; b_{i\mu}\right)$ where $N=30$, and $h(y;b)$ is a Bernoulli process with probability b. In the simulation, observation matrix $\{b_{i\mu}\}$ was randomly generated from a uniform distribution [0.1, 0.9). In Figures 5B and C, both transition matrix $\{a_{\mu\nu}\}$ and observation matrix $\{b_{i\mu}\}$ were learned to compare performance with other learning method, while in Figures 5D and E, only the transition matrix was acquired by synaptic plasticity. KL divergence in Figures 5B and C was evaluated as $\underset{perm}{\mathrm{argmin}} \sum_{\mu=1}^{p} x_{\mu}^{opt} \left(\log x_{\mu}^{opt} - \log x_{perm(\mu)}^{est}\right)$, where x_μ^{opt} is the estimation from the true internal model, x_μ^{est} is the estimation by each learning method, and *perm* denotes permutation over hidden states. See the Supplementary Information for the further details of the model.

Supplementary Information

1.Proof of convergence of the learning rule for the conceptual model

The derived learning rule can be rewritten as

$$\log p\left(v_c = v_k \mid x_{1:n}, y_{1:n}\right) = \sum_{n'} \log\left[1+(2v_k-1)x_{n'}(2y_{n'}-1)\right] + \text{const},$$

so in order to prove convergence, we need to show that $\varphi(v) \equiv \left\langle \log\left[1+(2v_k-1)x_{n'}(2y_{n'}-1)\right]\right\rangle_{n'}$ is maximized at true v_c. By considering Taylor expansion, the above equation is expanded as $\left\langle \log(1+z)\right\rangle = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m}\left\langle z^m \right\rangle$. In this form, the average is calculated as

$$\begin{aligned}\left\langle \left((2v_k-1)x_{n'}(2y_{n'}-1)\right)^m \right\rangle &= (2v_k-1)^m \left\langle x_{n'}y_{n'} + (-1)^m x_{n'}(1-y_{n'})\right\rangle \\ &= (2v_k-1)^m v_c \pi_x + (1-2v_k)^m (1-v_c)\pi_x\end{aligned}$$

Note that $(x_n)^m = x_n$ if $m>0$, because $x_n=0$ or 1. Thus, by substituting the above equation into the Taylor expansion form,

$$\begin{aligned}\varphi(v) &= \pi_x v_c \log\left[1+(2v-1)\right] + \pi_x(1-v_c)\log\left[1+(1-2v)\right] \\ &= \pi_x\left[v_c \log v + (1-v_c)\log(1-v)\right] + \text{const.}\end{aligned}$$

Therefore, φ(v) is maximized at $v = v_c$.

2. Details of recurrent circuit model

Let us consider a hidden Markov model in which state x^t is updated with $x^t \sim p(x^t|x^{t-1})$, and the observation y^t is given as $y^t \sim p(y^t|x^t)$. Here, we denote the total number of hidden state as p, the number of independent observation as N, transition probabilities as $p\left(x^t = \mu \mid x^{t-1} = v\right) \equiv a_{\mu v}$, and the probabilistic distribution of the observation as $p\left(y^t \mid x^t = \mu\right) = \prod_{i=1}^{N} p\left(y_i^t \mid x^t = \mu\right) \equiv \prod_{i=1}^{N} h\left(y_i^t; b_{i\mu}\right)$. The objective of the task is to estimate $A=\{a_{\mu v}\}$, $B=\{b_{i\mu}\}$, and $x^{1:t}=\{x^1,x^2,\ldots,x^t\}$ from given observations $y^{1:t}=\{y^1,y^2,\ldots,y^t\}$. Note that due to symmetry, there are at least $p!$ numbers of $\{A, B\}$ which gives the same system with the true $\{A^*, B^*\}$, and in that sense, the problem is ill-posed. However, it is still possible to acquire one of such $\{A, B\}$ asymptotically (Rabiner, 1989).

2.1 Particle filtering in parameter space

From Bayes rule, inference of x^t is given as

$$\begin{aligned} p\left(x^t \mid y^{1:t}\right) &\propto p\left(y^t \mid x^t, y^{1:t-1}\right) p\left(x^t \mid y^{1:t-1}\right) \\ &= \left(\int p\left(y^t, B \mid x^t, y^{1:t-1}\right) dB\right)\left(\sum_{x^{t-1}} \int p\left(x^t, x^{t-1}, A \mid y^{1:t-1}\right) dA\right) \\ &= \left(\int p\left(y^t \mid x^t, B\right) p\left(B \mid y^{1:t-1}\right) dB\right)\left(\sum_{x^{t-1}} p\left(x^{t-1} \mid y^{1:t-1}\right) \int p\left(x^t \mid x^{t-1}, A\right) p\left(A \mid y^{1:t-1}\right) dA\right). \end{aligned}$$

The last line holds because x^{t-1} and A are independent given $y^{1:t-1}$. Hence, if we denote $r_\mu^t \equiv p\left(x^t = \mu \mid y^{1:t}\right)$, the likelihood $r_\mu{}^t$ is given as,

$$\begin{aligned} r_\mu^t &= \left(\int p\left(B \mid y^{1:t-1}\right) \prod_i h\left(y_i^t; b_{i\mu}\right) dB\right)\left(\sum_\nu r_\nu^{t-1} \int p\left(A \mid y^{1:t-1}\right) a_{\mu\nu}\, dA\right) \\ &\approx \prod_i \left(\int p\left(b_{i\mu} \mid y^{1:t-1}\right) h\left(y_i^t; b_{i\mu}\right) db_{i\mu}\right)\left(\sum_\nu r_\nu^{t-1} \int p\left(a_{\mu\nu} \mid y^{1:t-1}\right) a_{\mu\nu}\, da_{\mu\nu}\right). \end{aligned}$$

Thus, for the given observation $y^{1:t}$, state x^t can be inferred recurrently. Here, we assumed independence of elements of the transition matrix A as $p\left(A \mid y^{1:t-1}\right) \approx \prod_\mu \prod_\nu p\left(a_{\mu\nu} \mid y^{1:t-1}\right)$ although they are mutually constrained by a boundary condition $\sum_\mu a_{\mu\nu} = 1$. Similarly, we assumed independence of elements of the observation matrix B.

The integral over $a_{\mu\nu}$ and $b_{i\mu}$ are generally not analytically calculable, but still approximately attainable by using particle filtering (Freitas et al., 2000). By taking K samples $\{a^k{}_{\mu\nu}\}$ from a proposed distribution $q_A(a)$, and by defining $\alpha_{\mu\nu}^{k,t} \equiv p\left(a_{\mu\nu}^k \mid y^{1:t}\right) / K q_A\left(a_{\mu\nu}^k\right)$, the integral can be approximated as

$$\int p\left(a_{\mu\nu} \mid y^{1:t-1}\right) a_{\mu\nu}\, da_{\mu\nu} \approx \sum_{k=1}^{K} \alpha_{\mu\nu}^{k,t-1} a_{\mu\nu}^k.$$

Similarly, by taking samples from a distribution $q_B(b)$ as $\{b^k{}_{i\mu}\} \sim q_B(b)$, the integration over b is approximated as $\int p\left(b_{i\mu} \mid y^{1:t-1}\right) h\left(y_i^t; b_{i\mu}\right) db_{i\mu} \approx \sum_k \beta_{i\mu}^{k,t-1} h\left(y_i^t; b_{i\mu}^k\right)$, where $\beta_{i\nu}^{k,t} \equiv p\left(b_{i\mu}^k \mid y^{1:t}\right) / K q_B\left(b_{i\mu}^k\right)$. Therefore, the update rule of $r_\mu{}^t$ is given as

$$r_\mu^t = \exp\left[\sum_i J_{i\mu}^t + \log\left(\sum_\nu \sum_k w_{\mu\nu}^{k,t-1} r_\nu^{t-1}\right) - I^t\right],$$

where $w_{\mu\nu}^{k,t-1} \equiv \alpha_{\mu\nu}^{k,t-1} a_{\mu\nu}^k$, $J_{i\mu}^t \equiv \log\left[\sum_k \beta_{i\mu}^{k,t-1} h\left(y_i^t; b_{i\mu}^k\right)\right]$, and $I^t \equiv \log\left(\sum_\nu \exp\left[\sum_i J_{i\nu}^t + \log\left(\sum_\rho \sum_k w_{\nu\rho}^{k,t-1} r_\rho^{t-1}\right)\right]\right)$.

The equation roughly corresponds to the firing dynamics of a recurrent network in which each neuron pair is connected with K number of synapses, assuming $r_\mu{}^t$ is the firing rate of neuron μ, $J^t{}_{i\mu}$ are the feed-forward inputs, and I^t is the global inhibition.

We next consider estimation of the importance weights. Because elements of matrix A are not independent, marginalization over all the other $\{a_{\mu'\nu'}\}_{(\mu',\nu') \neq (\mu,\nu)}$ is in general necessary to obtain $p(a_{\mu\nu}|y^{1:t})$ (i.e. $p\left(a_{\mu\nu} \mid y^{1:t}\right)=\int\prod_{(\mu',\nu')\neq(\mu,\nu)} p\left(A \mid y^{1:t}\right) da_{\mu'\nu'}$). However, if each $a_{\mu\nu}$ is successfully learned under the assumption of independence, constraints over $\{a_{\mu\nu}\}$ should be satisfied naturally. Thus, $p(a_{\mu\nu}|y^{1:t})$ is calculated as

$$p\left(a_{\mu\nu} \mid y^{1:t}\right) \propto p\left(a_{\mu\nu} \mid y^{1:t-1}\right)\left(1+\frac{pa_{\mu\nu}-1}{p-1} r_{\nu}^{t-1}\left(pr_{\mu}^{f,t}-1\right)\right),$$

by using following approximation:

$$\begin{aligned}
p\left(a_{\mu\nu} \mid y^{1:t}\right) &\propto p\left(a_{\mu\nu} \mid y^{1:t-1}\right) p\left(y^{t} \mid a_{\mu\nu}, y^{1:t-1}\right) \\
&= p\left(a_{\mu\nu} \mid y^{1:t-1}\right) \sum_{\lambda}\sum_{\rho} p\left(y^{t} \mid x^{t}=\rho, y^{1:t-1}\right) p\left(x^{t}=\rho \mid x^{t-1}=\lambda, a_{\mu\nu}, y^{1:t-1}\right) p\left(x^{t-1}=\lambda \mid y^{1:t-1}\right) \\
&\approx p\left(a_{\mu\nu} \mid y^{1:t-1}\right)\left(\sum_{\lambda\neq\nu}\sum_{\rho}\frac{1}{p} p\left(y^{t} \mid x^{t}=\rho, y^{1:t-1}\right) p\left(x^{t-1}=\lambda \mid y^{1:t-1}\right)\right. \\
&\qquad \left. +\sum_{\rho\neq\mu}\frac{1-a_{\mu\nu}}{p-1} p\left(y^{t} \mid x^{t}=\rho, y^{1:t-1}\right) p\left(x^{t-1}=\nu \mid y^{1:t-1}\right)+a_{\mu\nu} p\left(y^{t} \mid x^{t}=\rho, y^{1:t-1}\right) p\left(x^{t-1}=\nu \mid y^{1:t-1}\right)\right) \\
&\propto p\left(a_{\mu\nu} \mid y^{1:t-1}\right)\left(\frac{1-r_{\nu}^{t-1}}{p}+\frac{1-a_{\mu\nu}}{p-1}\left(1-r_{\mu}^{f,t}\right) r_{\nu}^{t-1}+r_{\mu}^{f,t} a_{\mu\nu} r_{\nu}^{t-1}\right) \\
&\propto p\left(a_{\mu\nu} \mid y^{1:t-1}\right)\left(1+\frac{pa_{\mu\nu}-1}{p-1} r_{\nu}^{t-1}\left(pr_{\mu}^{f,t}-1\right)\right),
\end{aligned}$$

where $r_{\mu}^{f,t} \equiv p\left(y^{t} \mid x^{t}=\mu, y^{1:t-1}\right) \Big/ \left[\sum_{\rho} p\left(y^{t} \mid x^{t}=\rho, y^{1:t-1}\right)\right]$. Therefore, the update rule of spine size is given as,

$$\tilde{\alpha}_{\mu\nu}^{k,t}=\left(1+\frac{\left(pa_{\mu\nu}^{k}-1\right) r_{\nu}^{t-1}\left(pr_{\mu}^{f,t}-1\right)}{p-1}\right)\alpha_{\mu\nu}^{k,t-1}, \quad \alpha_{\mu\nu}^{k,t}=\tilde{\alpha}_{\mu\nu}^{k,t} \Big/ \sum_{k'}\tilde{\alpha}_{\mu\nu}^{k',t}.$$

By defining the total synaptic weight as $w_{\mu\nu}^{t} \equiv \sum_{k}\alpha_{\mu\nu}^{k,t} a_{\mu\nu}^{k}$, this rule can be rewritten as,

$$\alpha_{\mu\nu}^{k,t}=\frac{1+f_{a}\left(r_{\nu}^{t-1}, r_{\mu}^{f,t}; a_{\mu\nu}^{k}\right)}{1+f_{a}\left(r_{\nu}^{t-1}, r_{\mu}^{f,t}; w_{\mu\nu}^{t-1}\right)}\alpha_{\mu\nu}^{k,t-1}, \quad f_{a}\left(r_{pre}, r_{post}; a^{k}\right) \equiv \left(pa^{k}-1\right) r_{pre}\left(pr_{post}-1\right) / (p-1).$$

Similarly, importance weights $\{\beta_{i\mu}^{k,t}\}$ of observation matrix B are derived as

$$\tilde{\beta}_{i\mu}^{k,t} \propto \beta_{i\mu}^{k,t-1}\left(1+\frac{\left(h\left(y_{i}^{t}; b_{i\mu}^{k}\right)-\bar{h}_{i}\left(y_{i}^{t}\right)\right)\left(\hat{r}_{\mu}^{t}-\pi_{\mu}\right)}{\bar{h}_{i}\left(y_{i}^{t}\right)\cdot\left(1-\pi_{\mu}\right)}\right), \quad \beta_{i\mu}^{k,t}=\tilde{\beta}_{i\mu}^{k,t} \Big/ \sum_{k'}\tilde{\beta}_{i\mu}^{k',t},$$

where $\pi_{\mu} \equiv p\left(x^{t}=\mu\right)$, and $\bar{h}_{i}\left(y_{i}^{t}\right) \equiv p\left(y^{t}=y_{i}^{t}\right)$. Here, for the sparseness constraint, instead of r_{μ}^{t}, we used a discretized version $\hat{r}_{\mu}^{t} \equiv \left[\hat{x}^{t}=\mu\right]_{+}$ for the learning rule, where the estimated state $\hat{x}^{t}$ is sampled from a probabilistic distribution $\{r_{\mu}^{t}\}$.

2.2 Online approximation of Baum-Welch formula

For the comparison, we also implemented three different rules. The standard rule for discrete HMM is Baum-Welch formula(Rabiner, 1989), which is described as

$$a_{\mu\nu}^{(n)} = \frac{\sum_{t=1}^{T-1} p\left(x^t = \nu, x^{t+1} = \mu \mid y^{1:T}, \theta^{(n-1)}\right)}{\sum_{t=1}^{T-1} p\left(x^t = \nu \mid y^{1:T}, \theta^{(n-1)}\right)},\quad b_{j\mu}^{(n)} = \frac{\sum_{t=1}^{T-1} p\left(x^t = \mu, y_j^t = 1 \mid y^{1:T}, \theta^{(n-1)}\right)}{\sum_{t=1}^{T-1} p\left(x^t = \mu \mid y^{1:T}, \theta^{(n-1)}\right)},$$

where $\theta^{(n)}$ is the set of parameters at *n*-th estimation. This standard machine learning method is an off-line learning rule, meaning that entire observation sequence is required for each update, thus not suitable for neural implementation.

By taking online approximation of the Baum-Welch formula, an online learning rule is obtained(Mongillo and Deneve, 2008). This rule can be extended to our problem setting straightforwardly as described below. Let us define

$$\psi_{\mu\nu}^{\lambda} \equiv \frac{1}{T}\sum_{t=1}^{T} P\left(x^t = \mu, x^{t-1} = \nu, x^T = \lambda \mid y^{1:T}\right),\quad \phi_{\mu j}^{\lambda} \equiv \frac{1}{T}\sum_{t=1}^{T}\left[y_j^t = 1\right]_+ P\left(x^t = \mu, x^T = \lambda \mid y^{1:T}\right).$$

Because

$$P\left(x^t = \mu, x^{t-1} = \nu, x^T = \lambda \mid y^{1:T}\right) = \sum_{\rho} \gamma_{\lambda\rho}\left(y^T\right) P\left(x^{T-1} = \rho, x^t = \mu, x^{t-1} = \nu \mid y^{1:T-1}\right),$$

where $\gamma_{\lambda\rho}\left(y^T\right) \equiv \dfrac{P\left(y^T \mid x^T = \lambda\right) P\left(x^T = \lambda \mid x^{T-1} = \rho\right)}{P\left(y^T \mid y^{1:T-1}\right)}$, $\psi_{\mu\nu}^{\lambda}(T)$ satisfies a recursive formula:

$$\psi_{\mu\nu}^{\rho}(T) = \sum_{\rho} \gamma_{\lambda\rho}\left(y^T\right)\left(\psi_{\mu\nu}^{\rho}(T-1) + \tfrac{1}{T}\left[\delta_{\nu\rho}\delta_{\mu\lambda} q_{\rho}(T-1) - \psi_{\mu\nu}^{\rho}(T-1)\right]\right) \quad \text{with} \quad q_{\rho}(T-1) \equiv P\left(x^{T-1} = \rho \mid y^{1:T-1}\right).$$

Similarly, $\phi_{\mu j}^{\lambda}(T)$ is recursively calculated by

$$\phi_{\mu j}^{\lambda}(T) = \sum_{\rho} \gamma_{\lambda\rho}\left(y^T\right)\left(\phi_{\mu j}^{\rho}(T-1) + \tfrac{1}{T}\left[\left[y_j^t = 1\right]_+ \delta_{\mu\lambda} q_{\rho}(T-1) - \phi_{\mu j}^{\rho}(T-1)\right]\right).$$

In addition, $\gamma_{\lambda\rho}\left(y^T\right)$ and $q_{\rho}(T)$ are given as

$$\gamma_{\lambda\rho}\left(y^T\right) = \frac{\left(\prod_j h\left(y_j^T; b_{j\lambda}\right)\right) a_{\lambda\rho}}{\sum_{\lambda'}\sum_{\rho'}\left(\prod_j h\left(y_j^T; b_{j\lambda'}\right)\right) a_{\lambda'\rho'} q_{\rho'}(T-1)},\quad q_{\lambda}(T) = \sum_{\rho} \gamma_{\lambda\rho}\left(y^T\right) q_{\rho}(T-1).$$

Therefore, with a learning rate parameter $\eta(t)$, an online EM algorithm is given as

$$\psi_{\mu\nu}^{\rho}(T) = \sum_{\rho} \gamma_{\lambda\rho}\left(y^T; A(T-1), B(T-1)\right)\left(\psi_{\mu\nu}^{\rho}(T-1) + \eta(T)\left[\delta_{\nu\rho}\delta_{\mu\lambda} q_{\rho}(T-1) - \psi_{\mu\nu}^{\rho}(T-1)\right]\right),$$

$$\phi_{\mu j}^{\lambda}(T) = \sum_{\rho} \gamma_{\lambda\rho}\left(y^T; A(T-1), B(T-1)\right)\left(\phi_{\mu j}^{\rho}(T-1) + \eta(T)\left[\left[y_j^t = 1\right]_+ \delta_{\mu\lambda} q_{\rho}(T-1) - \phi_{\mu j}^{\rho}(T-1)\right]\right),$$

$$a_{\mu\nu}(T)=\frac{\sum_{\lambda}\psi^{\lambda}_{\mu\nu}(T)}{\sum_{\mu'}\sum_{\lambda}\psi^{\lambda}_{\mu\nu}(T)},\quad b_{j\mu}(T)=\frac{\sum_{\lambda}\phi^{\lambda}_{\mu j}(T)}{\sum_{\nu'}\sum_{\lambda}\psi^{\lambda}_{\mu\nu'}(T)}.$$

In the simulation, we normalized $\{\psi^{\rho}_{\mu\nu}\}$ and $\{\phi^{\lambda}_{\mu j}\}$ to ensure stability of learning. To this end, we introduced auxiliary variables of $\phi^{\lambda}_{\mu j}(T)$ as

$$\bar{\phi}^{\lambda}_{\mu j}(T)=\sum_{\rho}\gamma_{\lambda\rho}\left(y^{T};A(T-1),B(T-1)\right)\left(\bar{\phi}^{\rho}_{\mu j}(T-1)+\eta(T)\left[\left[y^{t}_{j}=0\right]_{+}\delta_{\mu\lambda}q_{\rho}(T-1)-\bar{\phi}^{\rho}_{\mu j}(T-1)\right]\right).$$

Normalization was performed as $\phi^{\lambda}_{\mu\nu}\leftarrow\phi^{\lambda}_{\mu\nu}\Big/\sum_{\mu',\nu',\lambda'}\left(\phi^{\lambda'}_{\mu'\nu'}+\bar{\phi}^{\lambda'}_{\mu'\nu'}\right)$, $\bar{\phi}^{\lambda}_{\mu\nu}\leftarrow\bar{\phi}^{\lambda}_{\mu\nu}\Big/\sum_{\mu',\nu',\lambda'}\left(\phi^{\lambda'}_{\mu'\nu'}+\bar{\phi}^{\lambda'}_{\mu'\nu'}\right)$ and $\psi^{\rho}_{\mu\nu}\leftarrow\psi^{\rho}_{\mu\nu}\Big/\sum_{\mu',\nu',\rho'}\psi^{\rho'}_{\mu'\nu'}$.

2.3 Stochastic gradient descent rule

In addition to the above learning rules, we implemented a stochastic gradient descent (SGD) rule by considering gradient descent on the likelihood of input $y^{1:t}$ as

$$\tilde{r}^{t}_{\mu}=\left(\prod_{j}h\left(y^{t}_{j};b^{t-1}_{j\mu}\right)\right)\left(\sum_{\nu}a^{t-1}_{\mu\nu}r^{t}_{\nu}\right),\quad r^{t}_{\mu}=\tilde{r}^{t}_{\mu}\Big/\sum_{\mu'}\tilde{r}^{t}_{\mu'},$$

$$a^{t}_{\mu\nu}=a^{t-1}_{\mu\nu}+\eta_{sgd}\left(r^{t}_{\mu}-a^{t-1}_{\mu\nu}\right)r^{t-1}_{\nu},\quad \mathrm{b}^{t}_{j\mu}=\mathrm{b}^{t-1}_{j\mu}+\eta_{sgd}\left(y^{t}_{i}-\mathrm{b}^{t-1}_{j\mu}\right)r^{t}_{\mu}.$$

2.4 Details of simulations

In the simulation, we set p=5, N=30, and we generated the hidden transition matrix $\{a_{\mu\nu}\}$ as $a_{\mu\nu}$=0.5 if $\mu=\nu$, $a_{\mu\nu}$=0.4 if $\mu=\nu+1$ (mod. p), $a_{\mu\nu}$=0.333 otherwise. Observation matrix was randomly generated by $b_{j\mu}=b_{\min}+\left(b_{\max}-b_{\min}\right)\zeta_{j\mu}$, where b_{min}=0.1, b_{max}=0.9, and $\zeta_{j\mu}$ is a random variable uniformly sampled from [0,1). For the probabilistic distribution $h(y;b)$, we used a Bernoulli distribution with mean b.

In the multisynaptic learning rule, initial values of $\{a^{k}_{\mu\nu}\}$ and $\{b^{k}_{j\mu}\}$ were uniformly sampled from [0,1), and [b_{min}, b_{max}) respectively. Rewiring of $\{a^{k}_{\mu\nu}\}$ and $\{b^{k}_{j\mu}\}$ were performed in the same manner with the thresholds $\alpha_{th}=10^{-4}$, $\beta_{th}=10^{-6}$. In Figure 5E, we set the weight of new spine as $\alpha_{new}=1.1\times10^{-4}$, and $\beta_{new}=1.1\times10^{-6}$ to avoid repetitive rewiring, and we additionally defined a threshold for effective connectivity $\alpha_{eff}=10^{-2}$ to discount silent synapses.

In SGD learning and online EM learning rules, the initial values of estimated parameters $\{a_{\mu\nu}\}$ and $\{b_{j\mu}\}$ were uniformly sampled from $\left[(1-\Delta a_{init})/p,(1+\Delta a_{init})/p\right)$ and $\left[(1-\Delta b_{init})b_{avg},(1+\Delta b_{init})b_{avg}\right)$, where $\Delta a_{init}=0.05$, $\Delta b_{init}=0.1$ and $b_{avg}=(b_{max}+b_{min})/2$. Additionally, in online EM learning rules, $\left\{\psi^{\rho}_{\mu\nu}\right\}$ and $\left\{\phi^{\lambda}_{\mu j}\right\}$ were initialized with random values uniformly sampled from $\left[0.9b_{avg}/p^2,1.1b_{avg}/p^2\right)$ and $\left[0.9/p^3,1.1/p^3\right)$ respectively.