Learning rules for cortical-like spontaneous replay of an internal model

Toshitake Asabuki[1,2] and Tomoki Fukai[1]*

[1] Okinawa Institute of Science and Technology Graduate University, Tancha 1919-1, Onna-son, Okinawa 904-0495, Japan

[2] Bioengineering Department, Imperial College London, London SW7 2AZ, UK.

*Correspondence to: tomoki.fukai@oist.jp

Abstract

The brain is thought to learn an internal model of the environment for improved performance in perception, decision making, and inference. Evidence suggests that spontaneous cortical activity represents such a model, or prior distribution, by cycling through stimulus-evoked activity patterns at frequencies proportional to the probabilities that these stimuli were previously experienced. However, how the brain encodes priors into spontaneous activity and utilizes them for inference tasks remains unclear. Here, we present a synaptic plasticity mechanism to generate cell assemblies encoding the statistical structure of salient sensory events and spontaneously replay these assemblies in spiking recurrent neural networks. The plasticity mechanism installs a Hebbian-like learning rule at excitatory and inhibitory synapses to minimize mismatches in the probability structure between stimulus-evoked and internally driven activities. Our model replicates the behavioral biases of monkeys performing perceptual decision making with surprising accuracy, demonstrating how spontaneous replay of previous experiences biases cognitive behaviors.

Introduction

The brain needs to know the statistical structure of an uncertain environment to generate optimal behavioral responses. Bayesian inference identifies a process that accounts for uncertain observations from the environment and likely underlies various cognitive processes of the brain[1-5]. Computational studies suggest that the brain implements Bayesian computation at various levels of its processing hierarchy, from synaptic[6-10] and local network-level[11-16] to brain-wide computations[17-19]. Prior distributions accumulate subjective experiences in the environment, exerting crucial effects on statistical inference by individuals[20-23]. Clarifying how the brain performs statistical learning from the continuous stream of sensory experiences is vital to understanding the principles of brain computing and individuals' perceptual biases.

Evidence suggests that spontaneous brain activity learns a model of the sensory environment. In the primary visual cortex of anesthetized cats, spontaneously emerging activity patterns cycle through cortical states that include neural response patterns to oriented bars[24]. In the ferret visual cortex, the similarity between spontaneous activity and activities evoked by natural scenes increased gradually with visual experiences[25]. These results imply that cortical networks implement an internal model progressively adapting at the neural level to the statistical structure of sensory stimuli and spontaneously replay the modeled activity patterns. Then, the question arises about how cortical circuits extract salient events, encode them in spontaneous activity yielding an internal model, and use it for guiding cognitive behaviors. The underlying mechanisms of all these processes are only poorly understood[26, 27].

Here, we present a principle of learning to solve the above computational tasks by a recurrent spiking neural network. This principle proposes that all excitatory and inhibitory, or afferent and recurrent synapses on each excitatory or inhibitory neuron in the network learn to predict the neuron's response. The learning rules derived from this principle self-organize a spontaneous network activity that captures the statistical structure of the population activity patterns evoked by external stimuli. Thus, the spontaneous activity eventually represents an internal model of previous probabilistic sensory experiences. To show the biological plausibility of our model, we train our network model to perform a perceptual decision-making task[21]. In the experiment, monkeys had to indicate the direction (left or right) of coherent random-dot motions by a saccadic eye movement. During learning, monkeys were exposed to the two directions at different frequencies, which strongly biased the perceptual discrimination of monkeys during the test. Without fine-tuning, our model replicates the behavioral biases of monkeys surprisingly well, predicting that spontaneous activity presumably underlies the biased decision making of monkeys.

Intriguingly, our learning rules also predict a specific structure of cortical inhibitory circuits, that is, two inhibitory connection types with different computational roles. One connection type emerges between the assemblies of excitatory neurons to impose lateral inhibition among them, and the other type within the individual cell assemblies for desynchronizing neurons. Moreover, both connection types contribute to excitation-inhibition balance in recurrent networks. We demonstrate that the emergence of the two inhibitory circuits is crucial for robustly learning an internal model.

Results

Prediction-driven synaptic plasticity for encoding an internal model

First, we illustrate the class of tasks our model solves. Consider a task in which the animal should decide whether a given stimulus coincides with or resembles any of two previously learned stimuli. Whether the animal learned these stimuli with a 50-50 chance or an 80-20 chance should affect the animal's anticipation of their occurrence and hence affect its decision (Fig. 1a). It has been suggested that spontaneous activity expresses an optimal model of the sensory environment[25]. In our example, the evoked activity patterns of the two stimuli should be spontaneously replayed with the same probabilities as these stimuli were experienced during learning:

$$< P(\text{features}|\text{input}, \text{model}) >_{P(\text{input})} = P(\text{features}|\text{model}),$$

where features = {stimulus 1, stimulus 2} and the right-hand side expresses the probabilities of replayed activities. The angular brackets indicate averaging over the stimuli. We propose a mathematical principle for learning an optimal internal model. Namely, the same learning rule chunks salient features (here, stimuli 1 and 2) in temporal input, self-organizes cell assemblies encoding these features, and replays them in the post-learning spontaneous network activity (Fig. 1b). We show that a recurrent neural network can perform all these processes by minimizing the error between the statistical structure of sensory-evoked responses and that predicted by the internal model (Fig. 1c).

We introduce our learning principle using a toy network model that does not obey Dale's law for distinguishing between excitatory and inhibitory neurons (Methods). A more realistic model with distinct excitatory and inhibitory neuron pools will be shown later. The toy model consists of input neuron groups and a recurrent neural network of output neurons, where all neurons are Poisson spiking neurons. The number of input neuron groups can vary from task to task, and

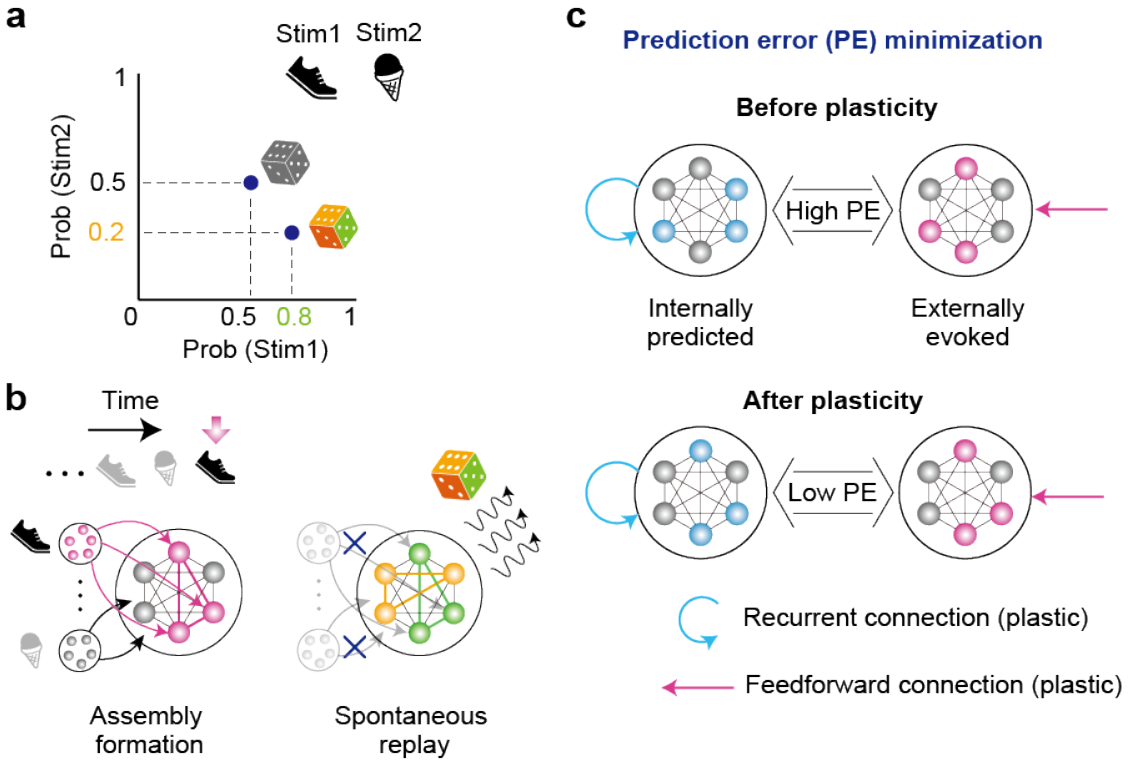


Figure 1. Spontaneous activity as a neural substrate for priors of experiences. (a) The animal may experience two sensory stimuli with different probabilities (e.g., a 50-50 chance in one case and a 20-80 chance in another. (b) According to Hebb's hypothesis, two cell assemblies are formed to memorize the two stimuli during learning (left). Doing so requires the segmentation of the stimuli repeated in the continuous and noisy stream of input from the environment. In the absence of the stimuli (during awake rest or sleep), spontaneous cortical activity replays sensory-evoked activity patterns with the probabilities presented during the learning (right). This match of the probability structure makes the spontaneous activity the optimal model of previous experiences. (c) During learning, all synaptic weights are asked to learn the probabilistic structure of sensory-evoked activities of individual neurons. This learning should enable a recurrent neural network to reproduce the learned probabilistic activity patterns spontaneously without external input.

every input neuron projects to all output neurons via modifiable afferent feedforward connection matrix $\mathbf{W}$ (Fig. 2a). Two types of all-to-all modifiable recurrent connections, $\mathbf{M}$ and $\mathbf{G}$, exist among the output neurons. Matrix $\mathbf{M}$ is a mixture of excitatory and inhibitory connections, and matrix $\mathbf{G}$ represents inhibitory-only connections. Due to a minus sign for v^G, all components of $\mathbf{G}$ are positive. All types of connections, both afferent and recurrent ones, are modifiable by self-supervised learning rules derived from a common principle: on each neuron, all synapses learn to predict the neuron's response optimally (Fig. 2b: see Methods). Therefore, unlike most previous studies, the network model does not have preconfigured stimulus-selective cell assemblies. During learning, the network should simultaneously self-organize cell assemblies and their afferent connections from specific input neuron groups. This unsupervised learning has a trivial solution in which all neurons become silent. To avoid it, the excitability of

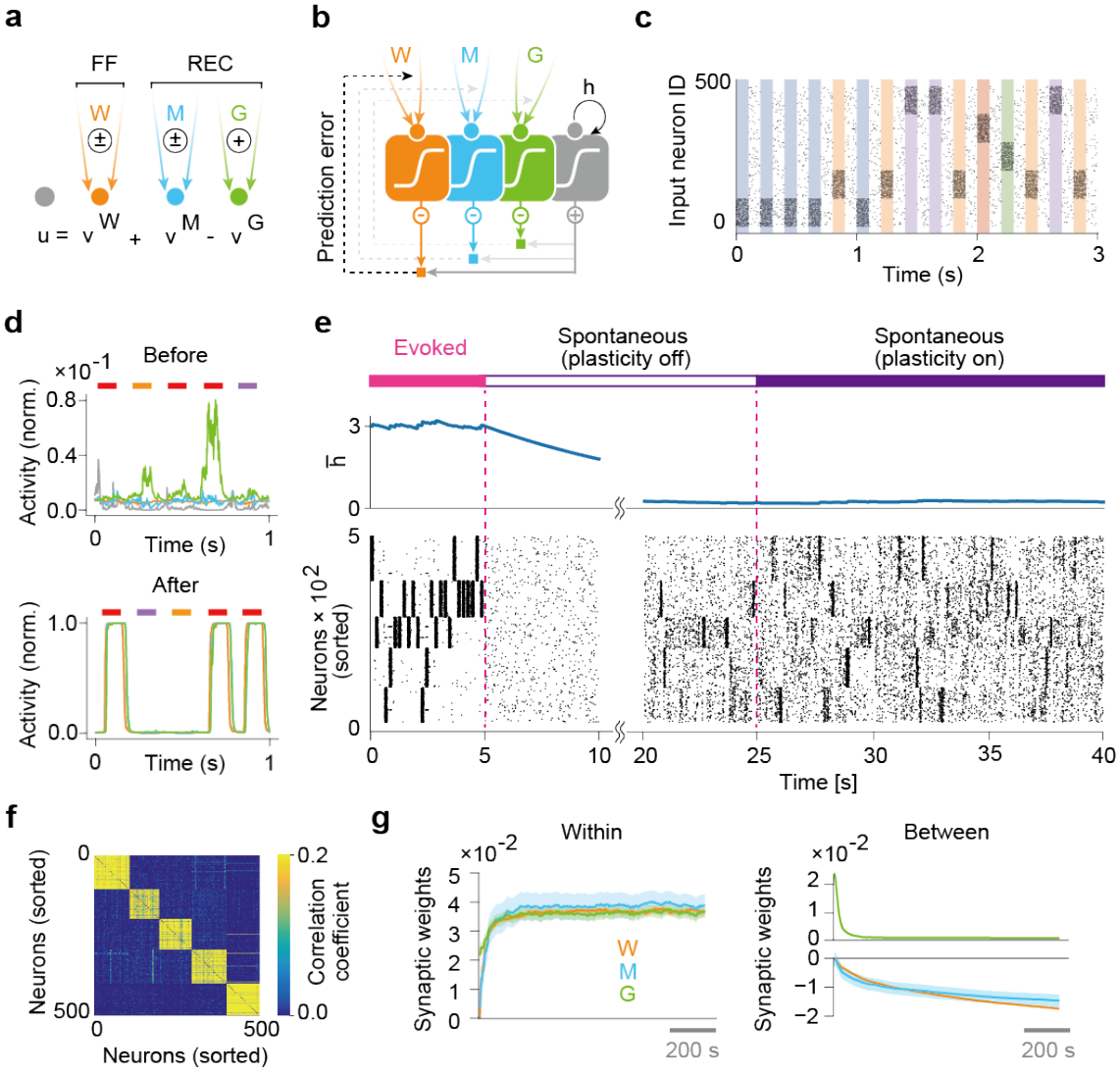


Figure 2. Formation of stimulus-selective assemblies in a recurrent network. (a) A schematic of connectivity in a network neuron is shown. A grey circle denotes the membrane potential u of a network neuron, which was calculated as a linear combination of inputs via different connections (v^W, v^M, and v^G; colored circles). Here, **W** denotes feedforward (FF), and **M** and **G** denote recurrent (REC) connections. (b) A schematic of the learning rule for a network neuron is shown. For each type of connection on a postsynaptic neuron, synaptic plasticity minimized the error between output (gray sigmoid) and synaptic prediction (colored sigmoid). The response function for the output was regulated by a dynamical variable h, which tracks the history of the membrane potential. (c) The example input spike trains during the initial three seconds are shown. Colored areas show periods of the presentation of the corresponding stimuli. All input neurons project to all network neurons. (d) Example dynamics of neuronal output and synaptic predictions are shown before (top) and after (bottom) learning. Colored bars at the top of the figures represent periods of stimulus presentations. (e) Example dynamics of the dynamical variable h (top) and the learned network activity (bottom) are shown. Neurons are sorted according to their preferred stimuli. During the spontaneous activity, afferent inputs to the network were lifted. (f) Correlation coefficients of spontaneous activities of every pair of neurons are shown. (g) Dynamics of the mean strengths of all types of connections projecting to assembly 1 are shown within (left) and between (right) assemblies. Shaded areas represent SDs over the designated network connectivity.

each output neuron (i.e., the slope and threshold of its response function) undergoes dynamical regulations according to the activity history (Eq. 10).

As the activity of a single neuron propagates to others in a recurrent network, the above learning rules also self-supervise the entire neural network to predict its activity patterns. Indeed, the proposed learning rules significantly extend the previous rule proposed for a two-compartment neuron, in which the dendritic compartment learns to predict the somatic response[28]. When the afferent input involves multiple repeated spike patterns, the dendritic compartment selectively learns one (rarely, more than one) of these patterns to achieve a statistical consistency between the somatic and dendritic responses. Consequently, the neuron model performs unsupervised segmentation of temporal input or chunking, which is essential for hierarchical sequence processing[29-31].

We simulated the learning of the network model with five input neuron groups, each of which consisted of 100 input neurons throughout this study. These neuron groups were activated randomly for 100 ms, one at a time, with an interval of 100 ms between the consecutive active epochs (Fig. 2c). Therefore, the input neuron groups were activated with equal probabilities. Each input neuron generated a Poisson spike train with a firing rate 2Hz (background activity) or 50Hz (during the active epoch). In the simulations, the elements of $\mathbf{W}$ and $\mathbf{M}$ were initialized by independent Gaussian distributions with mean zero. The elements of $\mathbf{G}$ were initially set to positive values and kept positive during the simulation.

The learning modified afferent connections $\mathbf{W}$ to segment repeated input patterns (Fig. 2d) and self-organized assemblies of output neurons selectively responding to one of the segmented patterns. We examined whether and how the spontaneous activity preserves and replays the learned cell-assembly structure without afferent input. The loss of afferent inputs initially lowered the activities of output neurons. Interestingly, however, their dynamic ranges gradually recovered owing to the excitability regulation in output neurons (see Eq. 10) and the network eventually started the spontaneous replay of the cell assemblies (Fig. 2e). It is noted that the plasticity rules of all synapses were turned off during this recovery period (20 seconds from the input termination). Once the network settled in a stable state with fixed connectivity (plasticity off), the plasticity rules were rebooted (plasticity on) without destabilizing the structured spontaneous replay. Spontaneous activities of neurons were highly correlated within each cell assembly, but uncorrelated between different cell assemblies (Fig. 2f). Consistent with this, recurrent connections $\mathbf{M}$ attained excitatory components only within, but not between, cell assemblies during learning (Fig. 2g, left). In contrast, the between-assembly components of $\mathbf{M}$

gradually and steadily became inhibitory, suppressing a simultaneous replay of different cell assemblies (Fig. 2g, right). Contrary to the inhibitory components of $\mathbf{M}$, inhibitory-only connections $\mathbf{G}$ eventually vanished between cell assemblies and only grew within cell assemblies (Fig. 2g, left), decorrelating neural activities within the cell assemblies. Thus, the network model successfully segregates, remembers, and replays stimulus-evoked activity patterns in temporal input.

Probability coding by replay events

Now, we turn to the central question of this study. We investigated whether spontaneous replay of cell assemblies can represent an optimal model of previous sensory experiences in this network. Namely, we examined whether the network spontaneously reactivates the learned cell assemblies with relative frequencies proportional to the probabilities with which the external stimuli activated the corresponding cell assemblies during learning.

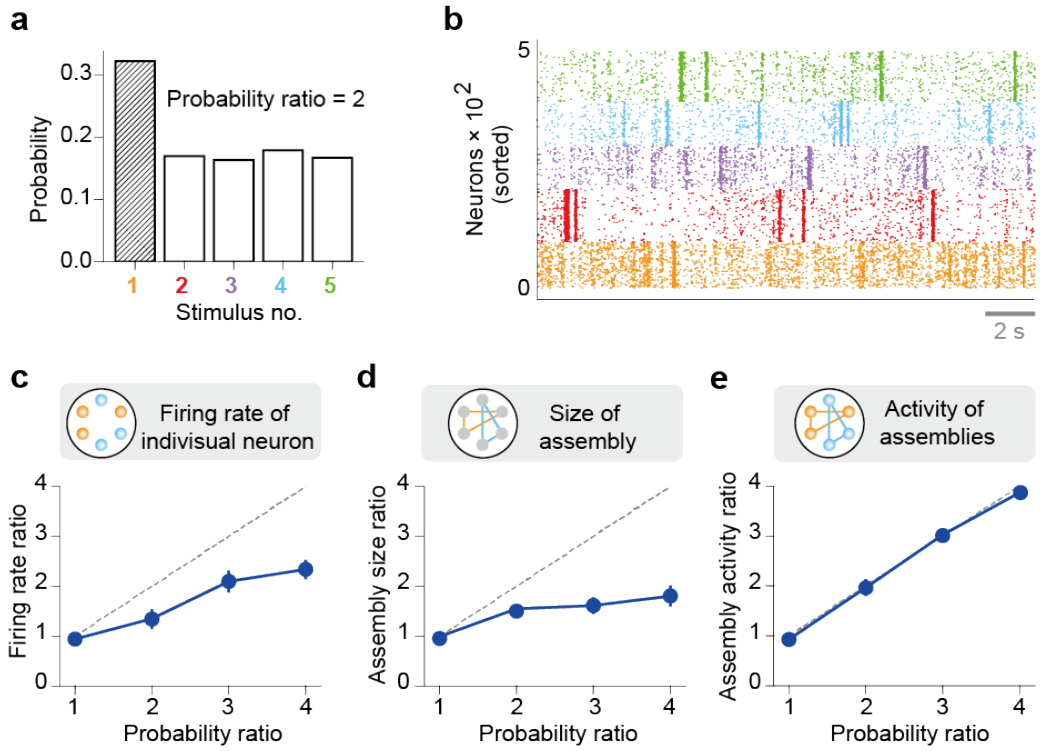


Figure 3. Priors coded in spontaneous activity. A network was trained with five probabilistic inputs. (a) Stimulus 1 appeared twice as often as the other four stimuli during learning. The example empirical probabilities of the stimuli used for learning are shown. (b) The spontaneous activity of the trained network shows distinct assembly structures. (c) The mean ratio of the population-averaged firing rate of assembly 1 to those of the other assemblies is shown for different values of the occurrence probability of stimulus 1. Vertical bars show SDs over five trials. A diagonal dashed line is a ground truth. (d) Similarly, the mean ratios of the size of assembly 1 to those of the other assemblies are shown. (e) The mean ratios of total activities of neurons in assembly 1 to those of the other assemblies are shown.

We first examined a case in which stimulus 1 was presented twice as often as the other four stimuli (Fig. 3a). Hereafter, the probability ratio refers to the relative number of times stimulus 1 is presented during learning. For instance, the case shown in Fig. 2 represents the probability ratio one. As in Fig. 2, the network self-organized five cell assemblies to encode stimuli 1 to 5 and replayed all of them in subsequent spontaneous activity (Fig. 3b). We found that output neurons were activated more frequently and strongly in cell assembly 1 than in other cell assemblies. Therefore, we accessed quantitative differences in neuronal activity between different cell assemblies by varying the probability ratio. The neuronal firing rate of cell assembly 1 relative to other cell assemblies increased approximately linearly with an increase in the probability ratio (Fig. 3c). Similarly, the size of cell assembly 1 relative to other cell assemblies also increased with the probability ratio (Fig. 3d). However, neither the relative firing rate nor the relative assembly size faithfully reflects changes in the probability ratio: scaling the probability ratio with a multiplicative factor does not scale these quantities with this factor. Therefore, we further investigated whether the assembly activity ratio, the ratio in the total firing rate of cell assembly 1 to other cell assemblies (Methods), scales faithfully with the probability ratio of cell assembly 1. This was the case: the scaling was surprisingly accurate (Fig. 3e). These results show that the trained network remembers the probabilities of repetitively experienced stimuli by the population firing rates of the encoding cell assemblies.

Relationship to Bayesian inference

We analytically show that recurrent neural networks trained by the proposed learning rules generate a prior distribution statistically equivalent to the posterior distribution determined by external stimuli. In the approximate treatment below, we assume that both system noise and observation noise have constant intensities and also ignore the system noise component induced by the neuron's own activity through recurrent connections. Then, as u_i and $\mathrm{v}_i^{\mathrm{W}}$ are calculated as linear combinations of independent synaptic input, according to the central limit theorem, we may assume that the prior distribution of the membrane potentials, which is provided by recurrent synaptic input, is Gaussian,

$$p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{M}}\right) \propto \exp\left[-\frac{\left(\mathrm{u}_i-\mathrm{v}_i^{\mathrm{M}}\right)^2}{2\sigma_{\mathrm{M}}^2}\right], \qquad (1)$$

and that the likelihood distribution of afferent inputs is also Gaussian,

$$p\left(\mathrm{v}_i^{\mathrm{W}}|\mathrm{u}_i\right) \propto \exp\left[-\frac{\left(\mathrm{u}_i-\mathrm{v}_i^{\mathrm{W}}\right)^2}{2\sigma_{\mathrm{W}}^2}\right], \qquad (2)$$

where σ_{M}^2 and σ_{W}^2 are the variances of the prior and likelihood, respectively. Using these expressions, we can represent the posterior distribution of cells' membrane potentials as

$$p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{W}},\mathrm{v}_i^{\mathrm{M}}\right) \propto p\left(\mathrm{v}_i^{\mathrm{W}}|\mathrm{u}_i,\mathrm{v}_i^{\mathrm{M}}\right)p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{M}}\right) = p\left(\mathrm{v}_i^{\mathrm{W}}|\mathrm{u}_i\right)p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{M}}\right)$$
$$\propto \exp\left[-\frac{\left(\mathrm{u}_i-\mathrm{v}_i^{\mathrm{M}}\right)^2}{2\sigma_{\mathrm{M}}^2}\right]\exp\left[-\frac{\left(\mathrm{u}_i-\mathrm{v}_i^{\mathrm{W}}\right)^2}{2\sigma_{\mathrm{W}}^2}\right] \propto \exp\left[-\frac{\left(\mathrm{u}_i-\mu_{\mathrm{post}}\right)^2}{2\sigma_{\mathrm{post}}^2}\right], \quad (3)$$

with the mean $\mu_{\mathrm{post}} = \frac{\sigma_{\mathrm{W}}^2\mathrm{v}_i^{\mathrm{M}}+\sigma_{\mathrm{M}}^2\mathrm{v}_i^{\mathrm{W}}}{\sigma_{\mathrm{W}}^2+\sigma_{\mathrm{M}}^2}$ and the variance $\sigma_{\mathrm{post}}^2 = \frac{\sigma_{\mathrm{W}}^2\sigma_{\mathrm{M}}^2}{\sigma_{\mathrm{W}}^2+\sigma_{\mathrm{M}}^2}$. Here, we used the fact that the relation $\left(\mathrm{v}_i^{\mathrm{W}} \perp\!\!\!\perp \mathrm{v}_i^{\mathrm{M}}|\mathrm{u}_i\right)$ holds (i.e., $\mathrm{v}_i^{\mathrm{W}}$ and $\mathrm{v}_i^{\mathrm{M}}$ are conditionally independent given u_i). The present learning rules minimize the difference, i.e., the KL divergence, between the posterior and prior distributions, which is expressed as

$$\mathrm{KL}\left[p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{W}}\right)\middle\|p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{M}}\right)\right] := \int p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{W}}\right)\log\left[\frac{p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{W}}\right)}{p\left(\mathrm{u}_i|\mathrm{v}_i^{\mathrm{M}}\right)}\right]d\mathrm{u}_i$$
$$= \log\left(\frac{\sigma_{\mathrm{M}}}{\sigma_{\mathrm{post}}}\right) + \frac{\sigma_{\mathrm{post}}^2+\left(\mathrm{v}_i^{\mathrm{M}}-\mu_{\mathrm{post}}\right)^2}{2\sigma_{\mathrm{M}}^2} - \frac{1}{2}. \quad (4)$$

Under the assumption that the variances are constant, synaptic weights making the term $\left(\mathrm{v}_i^{\mathrm{M}}-\mu_{\mathrm{post}}\right)^2$ vanish also minimize the KL divergence. Using the definition of μ_{post}, we can further rewrite this term as

$$\left(\mathrm{v}_i^{\mathrm{M}}-\mu_{\mathrm{post}}\right)^2 = \left(\frac{\sigma_{\mathrm{M}}^2}{\sigma_{\mathrm{W}}^2+\sigma_{\mathrm{M}}^2}\right)^2\left(\mathrm{v}_i^{\mathrm{M}}-\mathrm{v}_i^{\mathrm{W}}\right)^2.$$

The above term is approximately zero in the learned network as the relation $\mathrm{v}_i^{\mathrm{W}} \approx \mathrm{v}_i^{\mathrm{M}}$ holds after learning.

Roles of inhibitory plasticity for stabilizing cell assemblies

Experimental and computational results suggested that inhibitory synapses are more robust to spontaneous activity than excitatory ones and pivotal in preserving cortical circuit functions[32]. To see the crucial role of the inhibitory plasticity of $\mathbf{G}$ in our model, we compared how the self-organized assembly structure of recurrent connections $\mathbf{M}$ evolves in the two simulation settings shown in Fig. 4a. In the control model, we turned off the plasticity of $\mathbf{G}$ for a while after the cessation of external stimuli but again switched it on, as was previously in Fig. 2. The cell-assembly structure initially dissipated but eventually reached a well-defined equilibrium structure (Fig. 4b, magenta). Consistent with this, the postsynaptic potentials mediated by

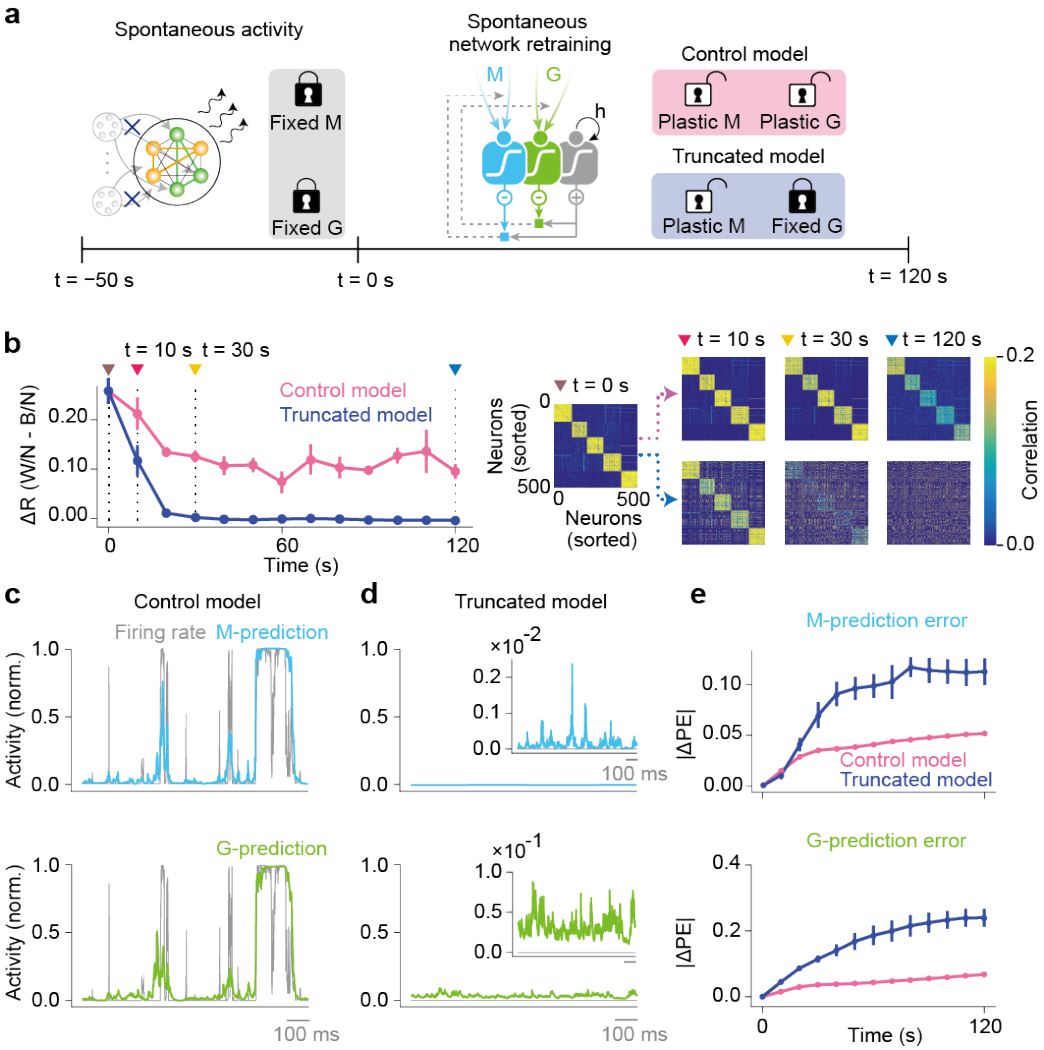


Figure 4 Crucial roles of inhibitory plasticity in prior learning. We first trained the network models with five external stimuli. (a) Then, we terminated the stimuli at -50 sec and waited until 0 sec for the recovery of network activity through the renormalization process (Eq. 10) with all plasticity rules turned off. We turned on the plasticity of $\mathbf{M}$ at time 0 sec. We kept the plasticity of $\mathbf{G}$ turned off in the truncated model (blue), while we turned on the G-plasticity in the control model (magenta). (b) *left*, The time evolution of the difference between the average within-assembly coherence ($\overline{R_{\mathrm{W/N}}}$) and the average between-assembly coherence ($\overline{R_{\mathrm{B/N}}}$) was plotted for the control (magenta) and truncated (blue) models. Larger differences imply more robust cell assemblies. Error bars indicate the SDs over five trials. *right*, Activity coherences between neurons are shown at the indicated times. (c) The time-varying normalized firing rate of a neuron (grey) and the values predicted by recurrent synaptic inputs (top) and lateral inhibition (bottom) are shown for the control model. (d) Similar plots are shown for the truncated model. (e) Changes in prediction errors in the control (magenta) and truncated (blue) models are shown for recurrent synaptic inputs (top) and lateral inhibition (bottom).

connections $\mathbf{M}$ and $\mathbf{G}$ predicted the normalized firing rate of a postsynaptic excitatory neuron in the control model (Fig. 4c). In striking contrast, the cell-assembly structure rapidly dissipated in the truncated model in which the G-plasticity was kept turned off after the cessation of external stimuli (Fig. 4b, blue). Accordingly, the postsynaptic potentials induced by $\mathbf{M}$ and $\mathbf{G}$, so was the

normalized firing rate, evolved into trivial solutions and almost vanished in the truncated model (Fig. 4d). Only the control model, but not the truncated model, could maintain prediction errors small and nearly constant after the termination of the stimuli (Fig. 4e). These results indicate that maintaining the learned representations requires the continuous tuning of within-assembly inhibition.

Learning conditional probabilities

In the previous section, we showed that the present network model successfully encodes the probabilities of stimulus experiences into spontaneous network activity. The results raise a question about how the brain uses the proposed mechanisms in other probabilistic inference tasks. Below, we show some examples of such behavioral tasks. First, we construct a neural network model that learns the conditional probabilities or the likelihoods of given pairs of stimuli. The network model encodes conditional probabilities among stimuli into the total firing rates of coactivated cell assemblies.

The neural network consists of two mutually interacting subnetworks of equal sizes sharing none of their membership neurons (Fig. 5a). Subnetwork A was randomly exposed to stimuli 1 and 2 (S1 and S2) with equal probabilities 1/2, whereas subnetwork B was to stimuli 3 and 4 (S3 and S4) with the conditional probabilities 1/3 and 2/3 if S1 was presented to subnetwork A and the conditional probabilities 2/3 and 1/3 if S2 was presented to subnetwork A. After learning, the network model self-organized four cell assemblies each of which responded preferentially to one of the four stimuli (Fig. 5b). Consistent with this, the self-organized connection matrix represented strong within-assembly connections within each cell assembly and weak between-assembly connections (Fig. 5c). Note that between-assembly connections were inhibitory between assemblies encoding mutually exclusive stimuli, i.e., S1 and S2 and S3 and S4, as they should be. Now, we turned off S3 and S4 to subnetwork B and only applied S1 or S2 to subnetwork A each at one time. Applying the same stimulus (i.e., S1 or S2) to subnetwork A activated either S3- or S4-coding cell assembly in subnetwork B in a probabilistic manner (Fig. 5d). The cell assemblies evoked in subnetwork B by S1 or S2 to subnetwork A varied the total firing rates approximately in proportion to the conditional probabilities (e.g., $P(S3|S1)=1/3$ vs. $P(S4|S1)=2/3$) used during learning (Fig. 5e). Note that S3- and S4-coding cell assemblies could become simultaneously active to represent the desired activation probabilities (e.g., a vertical arrow in Fig. 5d).

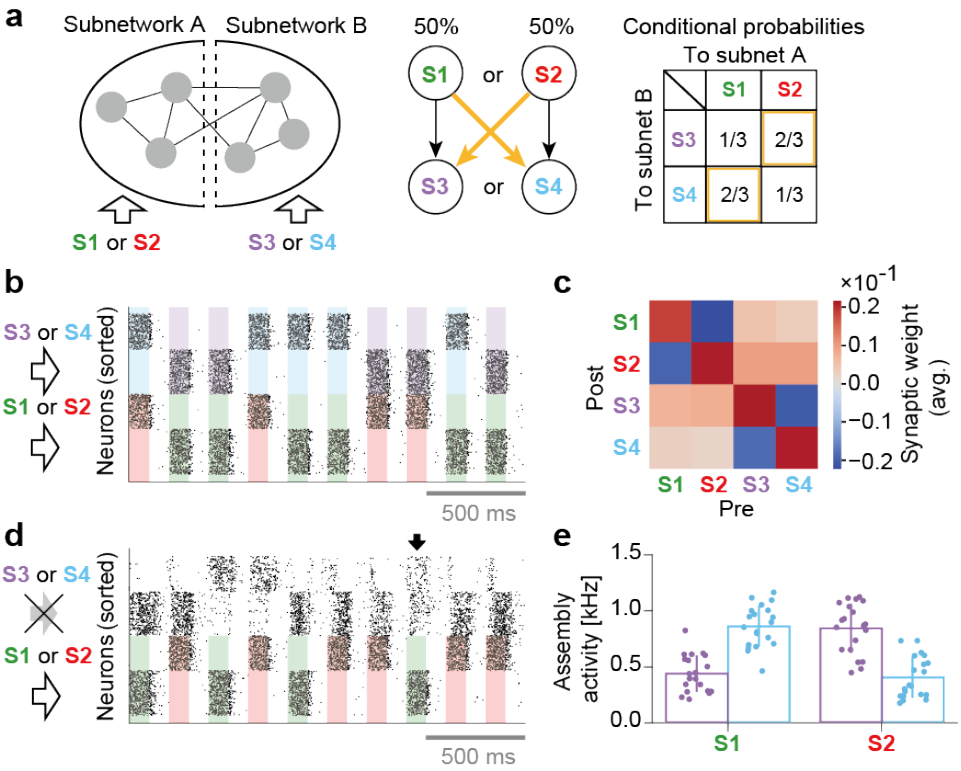


Figure 5. Learning of multivariate priors with assemblies. (a) Network neurons were separated into two populations receiving different groups of feedforward inputs (left). Subnetwork A received stimuli 1 (S1) and 2 (S2), each presented one at a time with probability 1/2. Subnetwork B received stimuli 3 (S3) or 4 (S4) exclusively when subnetwork A was also stimulated. S3 or S4 was sampled at each presentation according to the probability distribution conditioned on the stimulus presented to subnetwork A (middle and right). (b) Raster plot of evoked activity after training. Each subnetwork formed two assemblies responding to different preferred stimuli. Shaded areas with four colors indicate the duration of stimuli given to the two subnetworks. (c) The connection matrix self-organized among the cell assemblies is shown. (d) The activities of the four assemblies in the presence of S1 and S2 but not S3 and S4 are shown. Despite the absence of stimuli, subnetwork B replayed the assemblies encoding S3 and S4 when subnetwork A was activated by S1 or S2. (e) Activities of assemblies 3 and 4 in subnetwork B varied with the stimulus presented to subnetwork A. Each data point corresponds to one of 20 independent stimulus presentations. Error bars represent SDs.

Replication of biased perceptual decision making in monkeys

Prior knowledge about the environment often biases our percept of the external world. For instance, if we know that two possible stimuli exist and that stimulus A appears more often than stimulus B, we tend to feel that a given stimulus is more likely to be stimulus A than stimulus B. Previously, a similar bias was quantitatively studied in monkeys performing a perceptual decision making task[21]. In the experiment, monkeys had to judge the direction (right or left) of the coherent motion of moving dots on a display. When both directions of coherent motion appeared randomly during learning, the monkey showed unbiased choice behaviors. However, if the frequencies of the two motion directions were different, the monkey's choice was biased toward the direction of a more frequent motion stimulus.

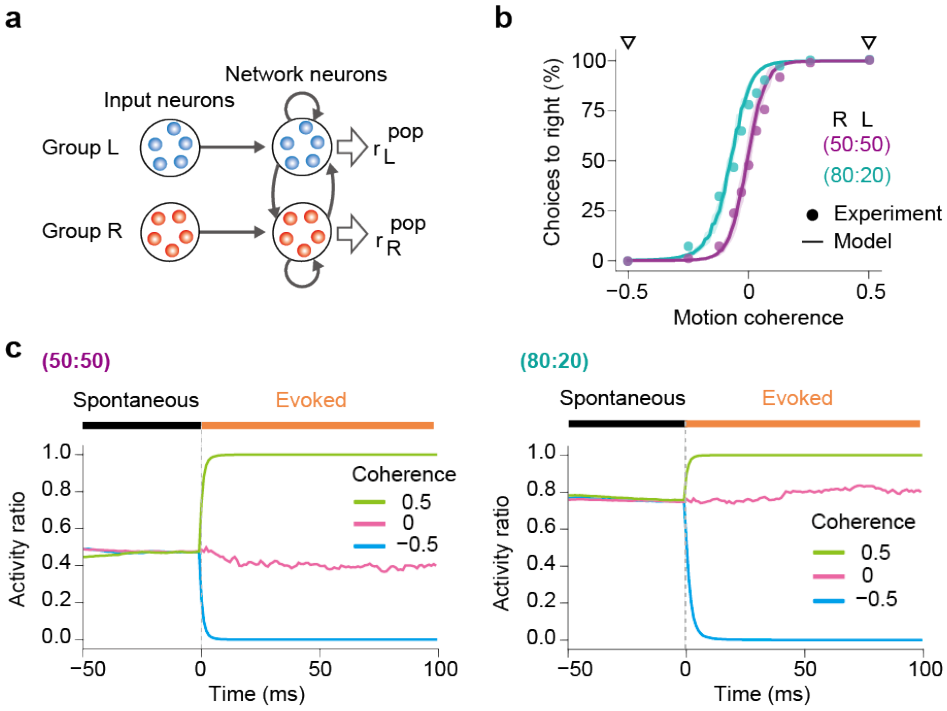


Figure 6. Simulations of biased perception of visual motion coherence. (a) The network model simulated perceptual decision-making of coherence in random dot motion patterns. In the network shown here, network neurons have already learned two assemblies encoding leftward or rightward movements from input neuron groups L and R. The firing rates of input neuron groups were modulated according to the coherence level Coh of random dot motion patterns (Methods). (b) The choice probabilities of monkeys (circles) and the network model (solid lines) are plotted against the motion coherence in two learning protocols with different prior probabilities. The experimental data were taken from Hanks et al. (2011). In the 50:50 protocol, moving dots in the "R" (Coh = 0.5) and "L" (Coh = -0.5) directions were presented randomly with equal probabilities, while in the 80:20 protocol, the "R" and "L" directions were trained with 80% and 20% probabilities, respectively. Shaded areas represent SDs over 20 independent simulations. The computational and experimental results show surprising coincidence without curve fitting. (c) Spontaneous and evoked activities of the trained networks are shown for the 50:50 (left) and 80:20 (right) protocols. Evoked responses were calculated for three levels of coherence: Coh = -50%, 0%, and 50%. In both protocols, the activity ratio in spontaneous activity matches the prior probability and gives the baseline for evoked responses. In the 80:20 protocol, the biased priors of "R" and "L" motion stimuli shift the activity ratio in spontaneous activity to an "R"-dominant regime.

We constructed a network model shown in Fig. 6a to examine whether the present mechanism of spontaneous replay could account for the behavioral bias. The model comprises a recurrent network similar to that used in Fig. 2 and two input neuron groups, L and R, encoding leftward or rightward coherent dot movements, respectively. We modulated the firing rates of these input neurons in proportion to the coherence of moving dots (Methods). During learning, we trained this model with external stimuli having input coherence Coh of either -0.5 or +0.5 (Methods), where all dots move leftward in the former or rightward in the latter. In so doing, we mimicked the two protocols used in the behavioral experiment of monkeys: in the 50:50

protocol, two stimuli with Coh = ±0.5 were presented randomly with equal probabilities, while in the 80:20 protocol, stimuli with Coh = +0.5 and -0.5 were delivered with probabilities of 80% and 20%, respectively. In the 80:20 protocol, stimuli were highly biased toward a coherent rightward motion.

The network model could explain the biased choices of monkeys surprisingly well. In either training protocol, the recurrent network self-organized two cell assemblies responding selectively to one of the R and L input neuron groups. Then, we examined whether the responses of the self-organized network are consistent with experimental observations by stimulating it with external inputs having various degrees of input coherence. The resultant psychometric curves almost perfectly coincide with those obtained in the experiment (Fig. 6b). We note that the psychometric curves of the model do not significantly depend on the specific choices of parameter values as far as the network learned stable spontaneous activity. We did not perform any curve fitting to experimental data, implying that the psychometric curves are free from parameter fine-tuning.

Biases in the psychometric curves emerged from biased firing rates of spontaneous activity of the self-organized cell assemblies. To show this, we investigated how the activities of the two self-organized cell assemblies change before and after the onset of test stimuli in three cases, i.e., Coh = -0.5, 0, and +0.5. Figure 6c shows the activity ratio AR between the R-encoding cell assembly and the entire network (Methods) in pre-stimulus spontaneous and post-stimulus evoked activity. When the network was trained in a non-biased fashion (i.e., in the 50:50 protocol), the activity ratio was close to 0.5 in spontaneous activity, implying that the two cell assemblies had similar activity levels. In contrast, when the network was trained in a biased fashion (i.e., in the 80:20 protocol), the activity ratio in spontaneous activity was close to 0.8, implying that the total spontaneous firing rate of R-encoding cell assembly was four times higher than that of L-encoding cell assembly.

An elaborate network model with distinct excitatory and inhibitory neuron pools

The predictive learning rule performed well in training the toy model to learn the probabilistic structure of the stimulus-evoked activity patterns. However, whether the same learning rule works in a more realistic neural network is yet to be investigated. To examine this, we constructed an elaborate network model consisting of distinct excitatory and inhibitory neuron pools, obeying Dale's law (Fig. 7a). The toy model suggested the essential roles of inhibitory plasticity in maintaining excitation-inhibition balance and generating an appropriate number of

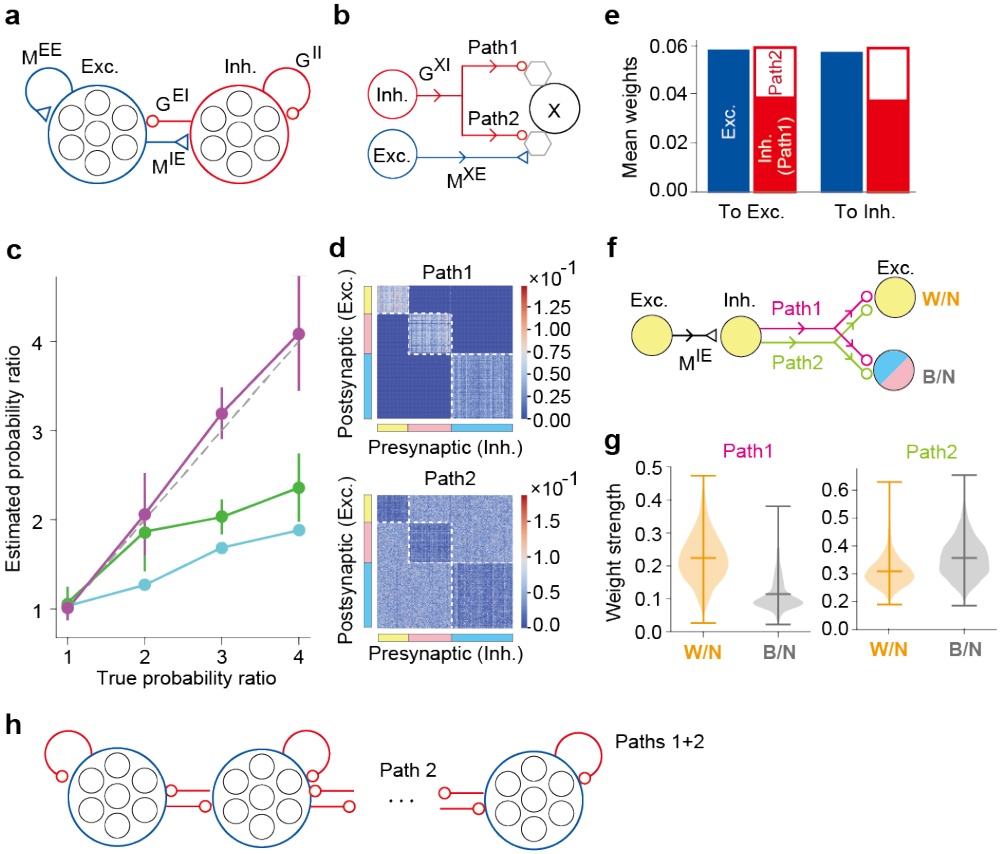
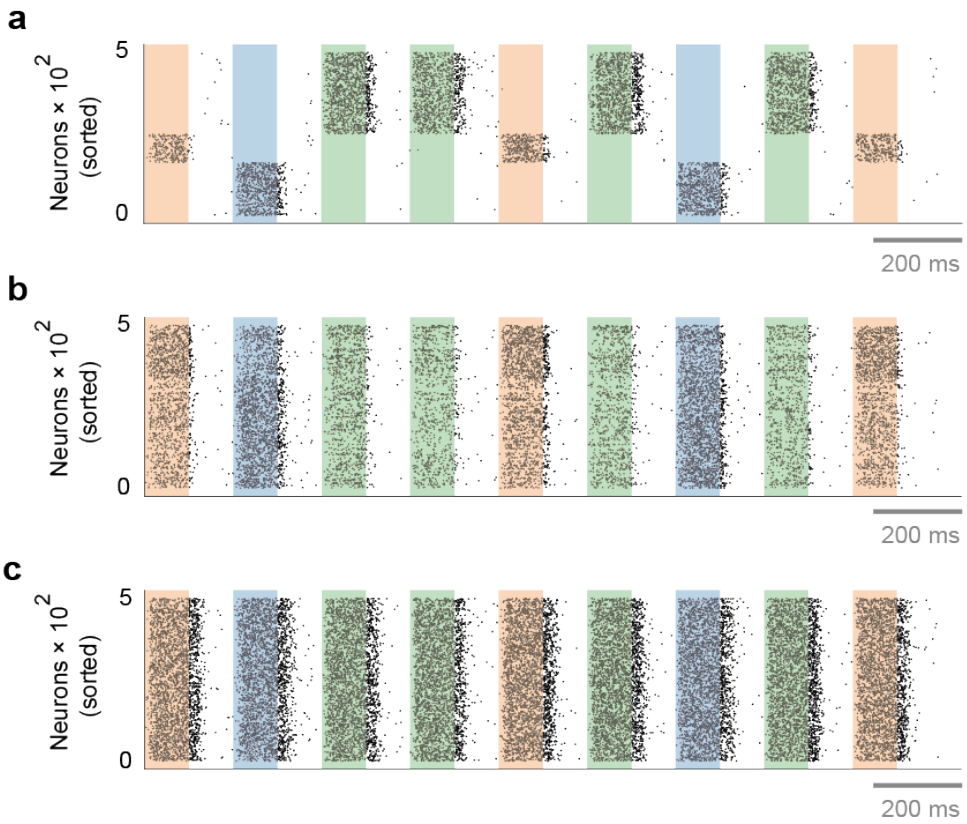


Figure 7. An elaborate model of excitatory and inhibitory cell assemblies. (a) The elaborate model consists of distinct excitatory and inhibitory neuron pools, obeying Dale's law. (b) Each inhibitory neuron projects to another neuron X through two inhibitory paths, path 1 and path 2, where the index X refers to an excitatory or an inhibitory postsynaptic neuron. Hexagons represent minimal units for prediction and learning in the neuron model and may correspond to dendrites, which were not modeled explicitly. (c) The probability ratios estimated by numerical simulations are plotted for the assembly activity ratios (purple), firing rate ratios (cyan), and assembly size ratios (green) as functions of the true probability ratio of external stimuli. Error bars indicate SEs calculated over five simulation trials with different initial states of neurons and synaptic weights in each parameter setting. (d) Inhibitory connection matrices are shown for path 1 and path 2. (e) The mean weights of self-organized synapses on excitatory and inhibitory postsynaptic neurons are shown. (f) Within-assembly and between-assembly connectivity patterns of excitatory and inhibitory neurons are shown. Colors indicate three cell assemblies self-organized. (g) The strengths of lateral inhibitions within- (W/N) and between-assemblies (B/N) are shown for paths 1 and 2. Horizontal bars show the medians and quartiles. (h) The resultant connectivity pattern suggests an effective competitive network between excitatory assemblies with self- (within-assembly) and lateral (between-assembly) inhibition.

cell assemblies. To achieve these functions, inhibitory neurons in the elaborate model project to excitatory and other inhibitory neurons via two synaptic paths (Fig. 7b). In path1, inhibitory connections alone predict the postsynaptic activity, whereas inhibitory and excitatory connections jointly predict the activity of the postsynaptic neuron in path2 (Methods). All synapses in the elaborate model are subject to the predictive learning rule. We trained the elaborate model with three input neuron groups while varying their activation probabilities. As

in the toy model, the elaborate model self-organized three cell assemblies activated selectively by the three input neuron groups (Supplementary Fig. 1a). Furthermore, in the absence of external stimuli, the elaborate model spontaneously replayed these assemblies with the assembly activity ratios in proportion to the occurrence probabilities of the corresponding stimuli during learning (Fig. 7c).



Supplementary Figure 1. The coexistence of the two inhibitory paths crucial for learning. (a) A typical spike raster of stimulus-evoked responses is presented for the elaborate network model shown in Fig. 7. (b) A spike raster of stimulus-evoked responses is shown for simulations of the elaborate model without inhibitory path 2. Inhibitory connections were modifiable in path 1. (c) A similar spike raster is presented for simulations of the elaborate model without inhibitory path 1. Inhibitory connections were modifiable in path 2. The results shown in b and c demonstrate that the network model fails to self-organize the cell assemblies encoding the different stimuli when it lacks one of the two inhibitory paths.

The two inhibitory paths divided their labors in a somewhat complex manner. To see this, we investigated the connectivity structures learned by these paths. In path 1, inhibitory connections were primarily found on excitatory neurons in the same assemblies (Fig. 7d, top). In contrast, in path 2, inhibitory connections were stronger on excitatory neurons in different assemblies than those in the same assemblies (Fig. 7d, bottom). On both excitatory and inhibitory neurons, the total inhibition (i.e., path 1 + path 2) was balanced with excitation (Fig. 7e). Figure 7f summarizes the connectivity structure of the elaborate model. Excitatory neurons in a cell assembly project to inhibitory neurons in the same assembly. Then, these inhibitory neurons project back to

excitatory neurons in the same or different assemblies via paths 1 and 2. Interestingly, lateral inhibition through path 1 is more potent between excitatory neurons within each cell assembly than between different assemblies (Fig. 7g). In contrast, path 2 mediates equally strong within-assembly and between-assembly inhibition.

We can understand the necessity of the two inhibitory paths based on the dynamical properties of competitive neural networks. Figure 7h displays the effective competitive network of excitatory cell assemblies suggested by the above results. Both paths 1 and 2 contribute to within-assembly inhibition among excitatory neurons, whereas between-assembly inhibition (i.e., lateral inhibition) mainly comes from path 2. In a competitive network, the lateral inhibition to self-inhibition strength ratio determines the number of winners having non-vanishing activities: the higher the ratio is, the smaller the number of winners is[33]. Therefore, self-organizing the same number of excitatory cell assemblies as that of external stimuli requires tuning the balance between the within-assembly and between-assembly inhibitions. This tuning during learning is likely easier when the network has two independently learnable inhibitory circuits. Indeed, a network model with only one inhibitory path rarely succeeded in encoding and replaying all stimuli used in learning (Supplementary Fig. 1b, c).

Discussion

Having proper generative models is crucial for accurately predicting statistical events. The brain is thought to improve the prediction accuracy of inference by learning internal generative models of the environment. These models are presumably generated through multiple mechanisms. For instance, the predictive coding hypothesizes that top-down cortical inputs provide lower sensory areas with prior information about sensory experiences[17-19]. However, experimental evidence also suggests that spontaneous activity represents an optimal model of the environment in sensory cortices. This study proposed a biologically plausible mechanism to learn such a model, or priors for experiences, with the brain's internal dynamics. To our knowledge, our model is the first to encode the optimal model of an environment into spontaneous activity.

Our model adopted a single principle for the plasticity of excitatory and inhibitory synapses to learn and replay the probabilistic structure of the environment. On each neuron, excitatory and inhibitory synaptic weights undergo plastic changes to improve their independent predictions on the cell's firing. This was done by minimizing the mismatch between the prior distribution and posterior distribution of the membrane potentials (Eq. 4). This simple learning rule showed

excellent performance in a simplified network model and in a more realistic model obeying Dale's law. In both models, inhibitory synaptic plasticity plays a crucial role in learning an accurate internal model by maintaining excitation-inhibition balance and decorrelating cell-assembly activities[34, 35]. The latter model predicts a division of labor between two inhibitory paths.

The mismatch between the prior and posterior distributions is called Bayesian surprise[36]. Therefore, it is said that the proposed learning rule asks a recurrent neural network to minimize the Bayesian surprise of its own activity for a given sensory input. Surprise signals have been implicated in learning cognitive behaviors[37, 38]. A well-known example is mismatch detection, in which listeners promptly perceive a deviant stimulus preceded by repetitions of standard stimuli[36, 39, 40]. Our model suggests that Bayesian surprise also plays a role in learning internal models of the environment in spontaneous brain activity.

Furthermore, we have shown that the proposed mechanism can account for the behavioral biases observed in perceptual decision making. This behavioral experiment quantitatively clarified how the difference in the probability between alternative sensory experiences during learning biases the choice behavior of monkeys during tests. However, the underlying neural mechanism was not explored. In its spontaneous activity, our model replays cell assemblies encoding the different stimuli at total firing rates proportional to the difference in their occurrence probabilities during learning. The difference in spontaneous firing rates explained the behavioral biases of monkeys accurately without any fine-tuning of model parameters. Our results strongly suggest that spontaneous cortical activity underlies the observed behavioral biases, although other mechanisms, such as biased top-down input, cannot be excluded.

As mentioned previously, various models have been proposed to account for neural mechanisms of Bayesian computation by the brain[6-19]. Typically, these models embed prior knowledge on previous sensory experiences into the wiring patterns of afferent (and sometimes also recurrent) synaptic inputs such that the corresponding stimuli can evoke the learned activity patterns to retrieve the prior knowledge. The present model differs from the previous models in several aspects: i) First, the model segments repeated stimuli to be remembered in an unsupervised fashion; ii) Then it generates cell assemblies encoding the segmented stimuli; (iii) Finally, it replays these cell assemblies spontaneously with learned probabilities. Note that the same learning rules enable the network to perform all necessary computations for (i) to (iii).

What could be the advantages of coding prior distributions into spontaneous activity over other ways of probability coding? We could consider some advantages. The replay of hippocampal engram cells is crucial for memory consolidation to cortical memory storage during sleep. Although the underlying mechanisms of hippocampal replay and memory consolidation have not been fully understood, the repeated activation of hippocampal engrams is thought to reinforce the activity patterns of cortical engrams[41-44]. Similarly, the spontaneous replay of learned cell assemblies in lower cortical regions may provide "training data" for learning priors in higher cortical areas, hence promoting hierarchical Bayesian modeling in predictive coding. Memory reinforcement by activity replay has also been studied in machine intelligence[45-47].

Another possible advantage of the spontaneous replay of internal models is that it may support the generalization of acquired knowledge. It was recently reported that a transitive inference task requires post-learning sleep in the anterior cingulate cortex (ACC)[48]. In this task, mice had to infer correct behavioral responses to inexperienced combinations of reward delivery rules from the outcomes of past experiences. The mice failed to generalize the learned rules if the activity of ACC was suppressed during post-learning sleep. The results suggested that the offline activity of ACC neurons, which likely involves the spontaneous replay of rule-encoding neurons, is crucial for rule generalization. Clarifying how the brain uses spontaneous activity to generalize the learned internal models is an intriguing open question.

Methods

Neural network model

Below, we first describe the model architecture and learning rule for a toy model (i.e., single population violating the Dale's law). Details of simulation of distinct excitatory and inhibitory populations will be explained later. Unless otherwise stated, recurrent neural networks used in this study consist of $N(=500)$ Poisson neurons, which generate spikes according to a non-stationary Poisson process with rate $\varphi(u)$, where $\varphi(\cdot)$ is a sigmoidal function. The membrane potential u of neuron i at time t is given as follows:

$$u_i(t) = \sum_{k=1}^{K} \mathrm{W}_{ik} x_k(t) + \sum_{k=1}^{N} (\mathrm{M}_{ik} - \mathrm{G}_{ik}) y_k(t), \qquad (5)$$

where K is the number of input neurons. In some simulations, the network model had more than one input neuron group although the number of input neuron groups is not explicitly shown in Eq. 5. Three matrices $\mathbf{W} \in \mathbb{R}^{N\times K}$, $\mathbf{M} \in \mathbb{R}^{N\times N}$, and $\mathbf{G} \in \mathbb{R}^{N\times N}$ represent the weights of afferent synaptic connections, recurrent synaptic connections and inhibitory-only connections,

respectively, on neurons in the recurrent network. These synaptic connections are all-to-all. In terms of the kernel function

$$\varepsilon(s) = \exp(-\mathrm{s}/\tau) \cdot \Theta(\mathrm{s}), \qquad (6)$$

recurrent input and afferent input to neuron i are calculated as

$$x_i(t) = \sum_{t' \in t_{\mathrm{aff}}^{\mathrm{f}}} \varepsilon(t - t'), \qquad (7a)$$

$$y_i(t) = \sum_{t' \in t_{\mathrm{rec}}^{\mathrm{f}}} \varepsilon(t - t'), \qquad (7b)$$

where τ stands for the membrane time constant, $t_{\mathrm{aff}}^{\mathrm{f}}$ and $t_{\mathrm{rec}}^{\mathrm{f}}$ for the time sets of afferent and recurrent presynaptic spikes, and $\Theta(\cdot)$ for the Heaviside function. Throughout this study, τ = 15 ms.

The instantaneous firing rate $f_i(t)$ of each neuron is given as

$$f_i(t) = \varphi\big(u_i(t); h_i(t)\big), \qquad (8)$$

in terms of sigmoidal response function φ:

$$\varphi(u_i; h_i) = \varphi_0\big[1 + \exp[g\beta(h_i)(-u_i + g\theta(h_i))]\big]^{-1}, \qquad (9)$$

with a constant value of $g = 3$. Here, the dynamical variable h is determined by the history of the membrane potential:

$$\begin{aligned} \dot{h}_i &= -{\tau_h}^{-1} h_i, && \text{if } h_i > u_i, \\ h_i &\leftarrow u_i, && \text{otherwise.} \end{aligned} \qquad (10)$$

The maximum instantaneous firing rate φ_0 is 50 Hz and $\tau_h = 10$ s. Through Eq. 10, h_i tracks the maximum value of the membrane potential u_i in a time window of approximately the length τ_h in the immediate past. The value of h is utilized to regulate the gain β and threshold θ of the sigmoidal response function as follows:

$$\beta(h_i) = {h_i}^{-1}\beta_0 \ , \qquad (11)$$

$$\theta(h_i) = h_i\theta_0 \ , \qquad (12)$$

where the values of constant parameters are $\beta_0 = 5$, and $\theta_0 = 1$. Neuron i generates a Poisson spike train at the instantaneous firing rate of $f_i(t)$.

Learning rules

To predict the firing rate of the postsynaptic neuron, the different types of synapses obey similar learning rules in the present network model. Given the postsynaptic potentials as

$$\mathrm{v}_i^{\mathrm{W}} = \sum_{j=1}^{K} \mathrm{W}_{ij} \cdot x_j, \qquad (13a)$$

$$\mathrm{v}_i^{\mathrm{M}} = \sum_{j=1}^{N} \mathrm{M}_{ij} \cdot y_{j,} \qquad (13b)$$

$$\mathrm{v}_i^{\mathrm{G}} = \sum_{j=1}^{N} \mathrm{G}_{ij} \cdot y_j, \qquad (13c)$$

the weights of the corresponding synapses are modified according to the following equations:

$$\Delta\mathrm{W}_{ij} = \epsilon\psi\left(f_i, \mathrm{v}_i^{\mathrm{W}}\right)x_j, \qquad (14a)$$

$$\Delta\mathrm{M}_{ij} = \epsilon\psi\left(f_i, \mathrm{v}_i^{\mathrm{M}}\right)y_j, \qquad (14b)$$

$$\Delta\mathrm{G}_{ij} = \epsilon\psi\left(f_i, \mathrm{v}_i^{\mathrm{G}}\right)y_j, \qquad (14c)$$

where the error term $\psi(f_i, \mathrm{v}_i)$ is defined as

$$\psi(f_i, \mathrm{v}_i) = {\varphi_0}^{-1}(1 - f_i/\varphi_0)[f_i - \varphi(\mathrm{v}_i; c_0)], \quad (15)$$

with a constant value of $c_0 = 1$. Throughout this study, the learning rate $\epsilon = 10^{-4}$, for which the typical time length required for the convergence of learning is 1,000 s.

Initial values of $\mathbf{W}$ and $\mathbf{M}$ are sampled from gaussian distributions with the mean 0 and variances $0.1/\sqrt{K}$ and $0.1/\sqrt{N}$, respectively. During learning, the elements of $\mathbf{W}$ and $\mathbf{M}$ can take both positive and negative values. After sufficient learning, the postsynaptic potentials $\mathrm{v}_i^{\mathrm{W}}$ and $\mathrm{v}_i^{\mathrm{M}}$ on neuron on neuron i converge to a common value of v_i^*. Therefore, $\varphi\left(v_i^{\mathrm{W}}\right) \approx \varphi\left(\mathrm{v}_i^{\mathrm{M}}\right) \approx \varphi(\mathrm{v}_i^*) \approx f_i$, implying that the postsynaptic potentials of afferent and recurrent synaptic inputs to neuron i can both predict its output f_i after learning. The initial values of $\mathbf{G}$

are uniformly set to $1/\sqrt{N}$, and its elements are truncated to non-negative values during learning. This implies that $\mathrm{v}_i^{\mathrm{G}}$ does not become negative. After learning, $\varphi(\mathrm{v}_i^{\mathrm{G}}) \approx f_i$ is satisfied. Although some elements of $\mathbf{M}$ may give recurrent inhibitory connections, modifiable connections in $\mathbf{G}$ are necessary to encode all external inputs into specific cell assemblies.

Measures for cell assembly activities

Here, we explain the measures used in Fig. 3. We calculated the firing rate ratio of cell assembly 1 in Fig. 3c as follows:

$$\text{Firing rate ratio} = \frac{r_i^{(1)}}{\left(\sum_{j=2}^{5} \frac{1}{N_j} \sum_{i=1}^{N_j} r_i^{(j)}\right)/4} , \qquad (16)$$

using the average firing rate $r_i^{(j)}$ of the i-th neuron in cell assembly j and the number N_j of neurons belonging to the cell assembly. Similarly, we defined the assembly size ratio of cell assembly 1 as

$$\text{Assembly size ratio} = \frac{N_1}{\left(\sum_{j=2}^{5} N_j\right)/4} . \qquad (17)$$

in Fig.3d and assembly activity ratio of cell assembly 1 as

$$\text{Assembly activity ratio} = \frac{r_{\mathrm{pop}}^{(1)}}{\left(\sum_{i=2}^{5} r_{\mathrm{pop}}^{(i)}\right)/4} , \qquad (18)$$

in Fig. 3e. Here, $\mathrm{r}_{\mathrm{pop}}^{(i)}$ represents the population neural activity of cell assembly i:

$$r_{\mathrm{pop}}^{(j)} \equiv \sum_{i=1}^{N_j} r_i^{(j)} . \qquad (19)$$

Simulations of perceptual decision making

In each learning trial, we trained the network with either leftward or rightward dot movement represented by the corresponding input neurons firing at $r_{\max} = 50$ Hz. In test trials, we defined input coherence as $\mathrm{Coh} = \rho_R - 0.5$ according to Hanks et al.[21], where ρ_R is the ratio of R input neurons to the sum of R and L input neurons in firing rate. The value of Coh ranges between -0.5 (all dots moving leftward) and +0.5 (all dots moving rightward). Then, in test trials for input coherence Coh, we generated Poisson spike trains of R and L input neurons at the rates $(\mathrm{Coh} + 0.5)r_{\max}$ and $(-\mathrm{Coh} + 0.5)r_{\max}$, respectively.

In Fig. 6c, we calculated the activity ratio (AR) as

$$\mathrm{AR} = \frac{r_{\mathrm{R}}^{\mathrm{pop}}}{r_{\mathrm{R}}^{\mathrm{pop}} + r_{\mathrm{L}}^{\mathrm{pop}}} , \qquad (20)$$

where $r_{\mathrm{R}}^{\mathrm{pop}}$ and $r_{\mathrm{L}}^{\mathrm{pop}}$ represent the average population firing rates of R-encoding and L-encoding cell assemblies, respectively. In Fig. 6b, we defined "choices to right" as

$$\text{Choices to right} = \mathrm{AR} \times 100\ (\%). \qquad (21)$$

A network model with distinct excitatory and inhibitory neuron populations

Here, we explain the architecture of the model used in Fig. 7. The network consists of $N_E(=500)$ excitatory and $N_I(=500)$ inhibitory neurons. The membrane potential of a neuron i of a population X (= E or I) at time t is given as follows:

$$u_i^{\mathrm{X}}(t) = \underbrace{\sum_{k=1}^{K} \mathrm{W}_{ik}^{\mathrm{X}} x_k(t)}_{=:\mathrm{v}_i^{\mathrm{W}}} + \underbrace{\left[\sum_{l=1}^{N_E} \mathrm{M}_{il}^{\mathrm{XE}} y_l^E(t) - \sum_{m=1}^{N_I} \mathrm{G}_{im}^{\mathrm{XI\,(path2)}} y_m^I(t)\right]}_{=:\mathrm{v}_i^{\mathrm{M}(2)}} - \underbrace{\sum_{m=1}^{N_I} \mathrm{G}_{im}^{\mathrm{XI\,(path1)}} y_m^I(t)}_{=:\mathrm{v}_i^{\mathrm{M}(1)}}, \qquad (22)$$

where $\{\mathrm{W}_{ik}^{\mathrm{X}}\}$ is afferent synaptic weights, which are a mixture of excitatory and inhibitory connections as in the toy model. The weights of recurrent excitatory synapses are $\{\mathrm{M}_{il}^{\mathrm{XE}}\}$. Here, we considered two types of recurrent inhibitory connections (i.e., path 1 and path 2), denoted by $\mathrm{G}_{im}^{\mathrm{XI\,(path1)}}$ and $\mathrm{G}_{im}^{\mathrm{XI\,(path2)}}$, respectively. Using the same definitions of kernel function of synaptic inputs and the error term as in Eq. 15, we modified these weights according to the following equations:

$$\Delta \mathrm{W}_{ij}^{\mathrm{X}} = \epsilon \psi\left(f_i, \mathrm{v}_i^{\mathrm{W}}\right) x_j, \qquad (23a)$$

$$\Delta \mathrm{M}_{ij}^{\mathrm{XE}} = \epsilon \psi\left(f_i, \mathrm{v}_i^{\mathrm{M}(2)}\right) y_j^E, \qquad (23b)$$

$$\Delta \mathrm{G}_{ij}^{\mathrm{XI\,(path2)}} = \epsilon \psi\left(f_i, \mathrm{v}_i^{\mathrm{M}(2)}\right) y_j^I, \qquad (23c)$$

$$\Delta \mathrm{G}_{ij}^{\mathrm{XI\,(path1)}} = \epsilon \psi\left(f_i, \mathrm{v}_i^{\mathrm{M}(1)}\right) y_j^I. \qquad (23d)$$

To satisfy the Dale's law, we truncated all weights of recurrent connections to non-negative values during learning.

In Fig. 7g, we measured the lateral inhibition between excitatory neurons via path 1 by calculating:

$$\left[\mathrm{W}_{i,j}^{\mathrm{LI}}\right] = \sum_{k=1}^{N_E} \mathrm{G}_{ik}^{\mathrm{EI\,(path1)}} \mathrm{M}_{kj}^{\mathrm{IE}}.$$

Lateral inhibition via path 2 was calculated in a similar fashion.

Data availability

All numerical datasets necessary to replicate the results shown in this article can easily be generated by numerical simulations with the software code provided below. No datasets were generated during this study.

Code availability

All codes were written in Python3 with numpy 1.17.3 and scipy 0.18.1. Example program codes used for the present numerical simulations and data analysis will be available after publication at https://github.com/ToshitakeAsabuki/PriorNet_codes.

References

1. Ernst, M.O. & Banks, M.S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429-433 (2002).

2. Kording, K.P. & Wolpert, D.M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244-247 (2004).

3. Fiser, J., Berkes, P., Orban, G. & Lengyel, M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* **14**, 119-130 (2010).

4. Orban, G., Berkes, P., Fiser, J. & Lengyel, M. Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron* **92**, 530-543 (2016).

5. Legaspi, R. & Toyoizumi, T. A Bayesian psychophysics model of sense of agency. *Nat. Commun.* **10**, 4250 (2019).

6. Tully, P.J., Hennig, M.H. & Lansner, A. Synaptic and nonsynaptic plasticity approximating probabilistic inference. *Front. Synaptic Neurosci.* **6**, 8 (2014).

7. Kappel, D., Habenschuss, S., Legenstein, R. & Maass, W. Network plasticity as Bayesian inference. *PLoS Comput. Biol.* **11**, e1004485 (2015).

8. Hiratani, N. & Fukai, T. Redundancy in synaptic connections enables neurons to learn optimally. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E6871-E6879 (2018).

9. Hiratani, N. & Latham, P.E. Rapid Bayesian learning in the mammalian olfactory system. *Nat. Commun.* **11**, 3845 (2020).

10. Aitchison, L., Jegminat, J., Menendez, J.A., Pfister, J.P., Pouget, A. & Latham, P.E. *et al.* Synaptic plasticity as Bayesian inference. *Nat. Neurosci.* **24**, 565-571 (2021).

11. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432-1438 (2006).

12. Deneve, S. Bayesian spiking neurons I: inference. *Neural Comput.* **20**, 91-117 (2008).

13. Nessler, B., Pfeiffer, M., Buesing, L. & Maass, W. Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Comput. Biol.* **9**, e1003037 (2013).

14. Hiratani, N. & Fukai, T. Hebbian wiring plasticity generates efficient network structures for robust inference with synaptic weight plasticity. *Front. Neural Circuits* **10**, 41 (2016).

15. Huang, Y. & Rao, R.P. Bayesian Inference and Online learning in Poisson neuronal networks. *Neural Comput.* **28**, 1503-1526 (2016).

16. Isomura, T., Shimazaki, H. & Friston, K.J. Canonical neural networks perform active inference. *Commun. Biol.* **5**, 55 (2022).

17. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127-138 (2010).

18. Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., & Friston, K.J. Canonical microcircuits for predictive coding. *Neuron* **76**, 695-711 (2012).

19. Keller, G.B. & Mrsic-Flogel, T.D. Predictive processing: a canonical cortical computation. *Neuron* **100**, 424-435 (2018).

20. Ashourian, P. & Loewenstein, Y. Bayesian inference underlies the contraction bias in delayed comparison tasks. *PLoS One* **6**, e19551 (2011).

21. Hanks, T.D., Mazurek, M.E., Kiani, R., Hopp, E. & Shadlen, M.N. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* **31**, 6339-6352 (2011).

22. Fritsche, M., Spaak, E. & de Lange, F.P. A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *Elife* **9,** e55389 (2020).

23. Hachen, I., Reinartz, S., Brasselet, R., Stroligo, A. & Diamond, M.E. Dynamics of history-dependent perceptual judgment. *Nat. Commun.* **12**, 6036 (2021).

24. Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A. & Arieli, A. Spontaneously emerging cortical representations of visual attributes. *Nature* **425**, 954-956 (2003).

25. Berkes, P., Orban, G., Lengyel, M. & Fiser, J. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* **331**, 83-87 (2011).

26. Jimenez Rezende D, Gerstner W. Stochastic variational learning in recurrent spiking networks. Front Comput Neurosci. 8:38 doi: 10.3389/fncom.2014.00038 (2014).

27. Li, A.A., Wang, F., Wu, S. & Zhang, X. Emergence of probabilistic representation in the neural network of primary visual cortex. *iScience* **25**, 103975 (2022).

28. Asabuki, T. & Fukai, T. Somatodendritic consistency check for temporal feature segmentation. *Nat. Commun.* **11**, 1554 (2020).

29. Orban, G., Fiser, J., Aslin, R.N. & Lengyel, M. Bayesian learning of visual chunks by human observers. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2745-2750 (2008).

30. Dehaene, S., Meyniel, F., Wacongne, C., Wang, L. & Pallier, C. The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* **88**, 2-19 (2015).

31. Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., Norman, K.A. Discovering event structure in continuous narrative perception and memory. *Neuron* **95**, 709-721 e705 (2017).

32. Mongillo G, Rumpel S, Loewenstein Y. Inhibitory connectivity defines the realm of excitatory plasticity. Nat Neurosci. 21:1463-1470 (2018).

33. Fukai T, Tanaka S. A simple neural network exhibiting selective activation of neuronal ensembles: from winner-take-all to winners-share-all. Neural Comput. 9:77-97 (1997).

34. Vogels, T.P.*, et al.* Inhibitory synaptic plasticity: spike timing-dependence and putative network function. *Front. Neural Circuits* **7**, 119 (2013).

35. Sprekeler, H. Functional consequences of inhibitory plasticity: homeostasis, the excitation-inhibition balance and beyond. *Curr. Opin. Neurobiol.* **43**, 198-203 (2017).

36. Ostwald, D.*, et al.* Evidence for neural encoding of Bayesian surprise in human somatosensation. *Neuroimage* **62**, 177-188 (2012).

37. Visalli A, Capizzi M, Ambrosini E, Mazzonetto I, Vallesi A. Bayesian modeling of temporal expectations in the human brain. *Neuroimage* **202**:116097 (2019).

38. Gijsen S, Grundei M, Lange RT, Ostwald D, Blankenburg F. Neural surprise in somatosensory Bayesian learning. *PLoS Comput Biol*. **17**:e1008068 (2021).

39. Cacciaglia, R., Costa-Faidella, J., Zarnowiec, K., Grimm, S. & Escera, C. Auditory predictions shape the neural responses to stimulus repetition and sensory change. *Neuroimage* **186**, 200-210 (2019).

40. Lecaignard, F., Bertrand, O., Caclin, A. & Mattout, J. Neurocomputational Underpinnings of Expected Surprise. *J. Neurosci.* **42**, 474-486 (2022).

41. Tonegawa, S., Morrissey, M.D. & Kitamura, T. The role of engram cells in the systems consolidation of memory. *Nat. Rev. Neurosci.* **19**, 485-498 (2018).

42. Ghandour K, Ohkawa N, Fung CCA, Asai H, Saitoh Y, Takekawa T, Okubo-Suzuki R, Soya S, Nishizono H, Matsuo M, Osanai M, Sato M, Ohkura M, Nakai J, Hayashi Y, Sakurai T, Kitamura T, Fukai T, Inokuchi K. Orchestrated ensemble activities constitute a hippocampal memory engram. *Nat Commun*. **10**, 2637 (2019).

43. Klinzing, J.G., Niethard, N. & Born, J. Mechanisms of systems memory consolidation during sleep. *Nat. Neurosci.* **22**, 1598-1610 (2019).

44. Takehara-Nishiuchi, K. Neurobiology of systems memory consolidation. *Eur. J. Neurosci.* **54**, 6850-6863 (2021).

45. Dayan, P., Hinton, G.E., Neal, R.M. & Zemel, R.S. The Helmholtz machine. *Neural Comput.* **7**, 889-904 (1995).

46. Goodfellow, I.*, et al.* Generative Adversarial Nets. in *Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014)* 2672–2680 (2014).

47. Luczak, A., McNaughton, B.L. & Kubo, Y. Neurons learn by predicting future activity. *Nat. Mach. Intell.* **4**, 62-72 (2022).

48. Kareem Abdou, Kiriko Choko, Mohamed H. Aly, Reiko Okubo-Suzuki, Shin ichi Muramatsu, Kaoru Inokuchi. Inspiring cognitive inference in a cortical network during REM sleep. bioRxiv 2021.04.08.439095.

Acknowledgments

The authors express their sincere thanks to Yukiko Goda for her valuable comments on our manuscript. This work was supported by KAKENHI (nos. 18H05213, 18H05213 and 19H04994) to T.F.

Author contributions

T.A. and T.F. conceived the study and wrote the paper. T.A. performed the simulations and data analyses.

Competing interests

The authors declare no competing interests.