Predictive learning rules generate a cortical-like replay of probabilistic sensory experiences

Toshitake Asabuki[1,2,3]* and Tomoki Fukai[1]*

[1] Okinawa Institute of Science and Technology Graduate University, Tancha 1919-1, Onna-son, Okinawa 904-0495, Japan
[2] RIKEN Center for Brain Science, RIKEN ECL Research Unit
[3] RIKEN Cluster for Pioneering Research, Japan
*Correspondence to: tomoki.fukai@oist.jp (T.F.); toshitake.asabuki@riken.jp (T.A.)

Abstract

The brain is thought to construct an optimal internal model representing the probabilistic structure of the environment accurately. Evidence suggests that spontaneous brain activity gives such a model by cycling through activity patterns evoked by previous sensory experiences with the experienced probabilities. The brain's spontaneous activity emerges from internally-driven neural population dynamics. However, how cortical neural networks encode internal models into spontaneous activity is poorly understood. Recent computational and experimental studies suggest that a cortical neuron can implement complex computations, including predictive responses, through soma-dendrite interactions. Here, we show that a recurrent network of spiking neurons subject to the same predictive learning principle provides a novel mechanism to learn the spontaneous replay of probabilistic sensory experiences. In this network, the learning rules minimize probability mismatches between stimulus-evoked and internally driven activities in all excitatory and inhibitory neurons. This learning paradigm generates stimulus-specific cell assemblies that internally remember their activation probabilities using within-assembly recurrent connections. The plasticity of cells' intrinsic excitabilities normalizes neurons' dynamic ranges to further improve the accuracy of probability coding. Our model contrasts previous models that encode the statistical structure of sensory experiences into Markovian transition patterns among cell assemblies. We demonstrate that the spontaneous activity of our model well replicates the behavioral biases of monkeys performing perceptual decision making. Our results suggest that interactions between intracellular processes and

recurrent network dynamics are more crucial for learning cognitive behaviors than previously thought.

Introduction

The brain is believed to construct an internal statistical model of an uncertain environment from sensory information streams for predicting the external events that are likely to occur. Evidence suggests that spontaneous brain activity learns the representation of such a model through repeated experiences of sensory events. In the cat visual cortex, spontaneously emerging activity patterns cycle through cortical states that include neural response patterns to oriented bars (Kenet et al., 2003). In the ferret visual cortex, spontaneous activity gradually resembles a superposition of activity patterns evoked by natural scenes, eventually giving an optimal model of the visual experience (Berkes et al., 2011). As replay activities can provide prior information for hierarchical Bayesian computation by the brain (Ernst & Banks, 2002; Kording & Wolpert, 2004; Friston, 2010; Fiser et al., 2010; Bastos et al., 2012; Orban et al., 2016; Legaspi & Toyoizumi, 2019), clarifying how the brain learns the spontaneous replay of optimal internal models is crucial for understanding whole-brain computing. However, the neural mechanisms underlying this modeling process are only poorly understood.

Several mechanisms of the brain's probabilistic computation have been explored (Jimenez Rezende & Gerstner, 2014; Li et al., 2022). Models with reverberating activity are particularly interesting owing to their potential ability to generate spontaneous activity. For instance, spiking neuron networks with symmetric recurrent connections were proposed for Markov Chain Monte Carlo sampling of stochastic events (Buesing et al, 2011; Bill et al., 2015). Spike-timing-dependent plasticity was used to organize spontaneous sequential activity patterns, providing a predictive model of sequence input (Hartmann et al., 2015). However, previous models did not clarify how recurrent neural networks learn the spontaneous replay of the probabilistic structure of sensory experiences, for which these networks should learn the accurate probabilities of sensory stimuli and an appropriate excitation-inhibition balance simultaneously. Moreover, previous models assumed that each statistically salient stimulus in temporal input is already segregated and is delivered to a pre-assigned assembly of coding neurons, implying that the recurrent network, at least partly, knows the stochastic events to be modeled before

learning. How the brain extracts salient events for statistical modeling has not been addressed.

Here, we present a learning principle to encode the experiences' probability structure into spontaneous network activity. To this end, we extensively use the synaptic plasticity rule proposed previously based on the hypothesis that the dendrites of a cortical neuron learn to predict its somatic responses (Urbanczik & Senn, 2014; Asabuki & Fukai, 2020). We generalize the hypothetical predictive learning to a learning principle at the entire network level. Namely, in a recurrent network driven by external input, we ask all synapses on the dendrites of each excitatory or inhibitory neuron to learn to predict its somatic responses (although the dendrites will not be explicitly modeled). This enables the network model to simultaneously learn the events' probabilistic structure and the excitation-inhibition balance required to replay this structure. Further, our network model requires no pre-assigned cell assemblies since the model neuron can automatically segment statistically salient events in temporal input (Asabuki & Fukai, 2020) - a cognitive process known as "chunking" (Fujii & Graybiel, 2003; Jin & Costa, 2010; Jin et al., 2014; Schapiro et al., 2013; Zacks et al., 2001). Intriguingly, the cell assemblies generated by our model store their replay probabilities primarily in the within-assembly network structure, and intrinsic dynamical properties of membership neurons also contribute to this coding. This is in striking contrast to other network models that encode probabilities into the Markovian transition dynamics among cell assemblies (Buesing et al., 2011; Hartmann et al., 2015).

Our model trained on a perceptual decision-making task can replicate both unbiased and biased decision behaviors of monkeys without fine-tuning of parameters (Hanks et al., 2011). In addition, in a network model consisting of distinct excitatory and inhibitory neural populations, our learning rule predicts the emergence of two types of inhibitory connections with different computational roles. We show that the emergence of the two inhibitory connection types is crucial for robust learning of an optimal internal model.

Results

Replay of probabilistic sensory experiences - A toy example

We first explain the task our model solves with a toy example. 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 a 30-70 chance should affect the animal's anticipation of their occurrence and hence affect its decision.

It has been suggested that spontaneous activity expresses an optimal internal model of the sensory environment (Berkes et al., 2011). In our toy 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. According to Hebb's hypothesis, two cell assemblies should be formed to memorize the two stimuli in the toy example. Moreover, the spontaneous replay of these cell assemblies should represent the probabilities given in the right-hand side of the above equation. Below, we propose a mathematical principle of learning to achieve these requirements.

Prediction-driven synaptic plasticity for encoding an internal model

We previously proposed a learning rule for a single two-compartment neuron (Asabuki & Fukai, 2020). Briefly, our previous model learns statistically salient features repeated in input sequences by minimizing the error between somatic and dendritic response probabilities without external supervision to identify the temporal locations of these features. In this study, we extend this plasticity rule to recurrent networks by asking all neurons in a network to minimize the error in response probabilities between the internally generated and stimulus-evoked activities (Fig. 1). Our central interest is whether this learning principle generates spontaneous activity representing the statistical model of previous experiences.

We first introduce our learning principle using a recurrent network model (nDL model) that does not obey Dale's law for distinguishing between excitatory and inhibitory neurons (Materials and Methods). A more realistic model with distinct excitatory and inhibitory neuron pools will be shown later. The nDL model consists

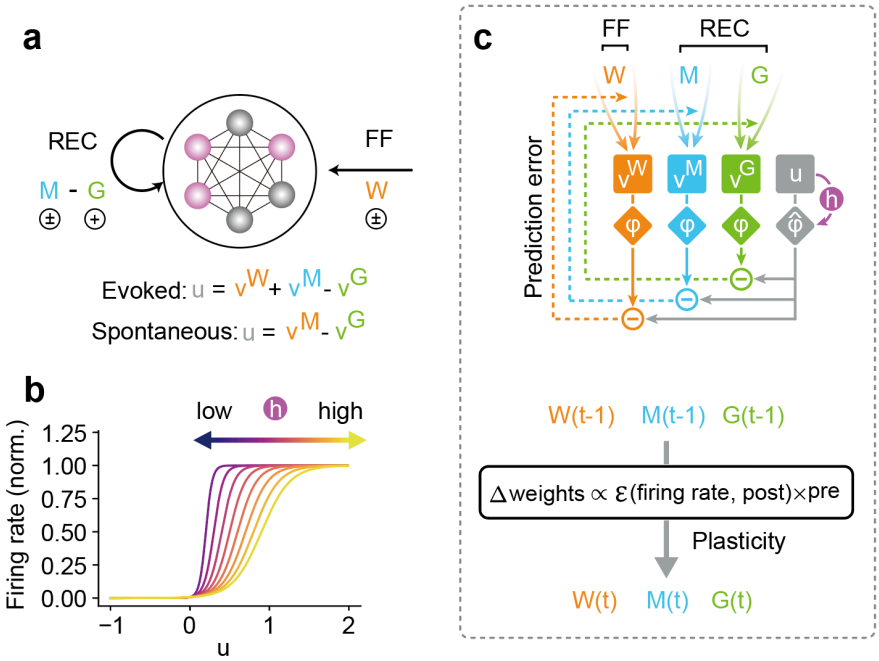

Figure 1. Unsupervised prior learning in a recurrent neural network. (a) A schematic of a network model is shown. The interconnected circles denote the model neurons, of which the activities are controlled by two types of inputs: feedforward (FF) and recurrent (REC) inputs. Colored circles indicate active neurons. Here, $\mathbf{W}$ denotes FF, and $\mathbf{M}$ and $\mathbf{G}$ denote REC connections. We considered two modes of activity (i.e., evoked and spontaneous activity). In the evoked mode, the membrane potential u of a network neuron was calculated as a linear combination of inputs across all different connections (v^W, v^M, and v^G). This evoked mode is considered during the learning phase, when all synapses attempt to predict the network activity, as we will explain in the main text. Once all synapses are sufficiently learned, all FF inputs are removed, and the network is driven spontaneously (spontaneous mode). Our interest lies in the statistical similarity of the network activity in these two modes. (b) The gain and threshold of output response function was controlled by a dynamic variable, h, which tracks the history of the membrane potential. (c) A schematic of the learning rule for a network neuron is shown (top). During learning, for each type of connection on a postsynaptic neuron, synaptic plasticity minimizes the error between output (gray diamond) and synaptic prediction (colored diamonds). Note that all types of synapses share the common plasticity rule, where weight updates are calculated as the multiplication of the error term and the presynaptic activities (bottom). Our hypothesis is that such plasticity rule allows a recurrent neural network to spontaneously replay the learned stochastic activity patterns without external input.

of Poisson spiking neurons, each receiving Poisson spike trains from all input neurons via a modifiable all-to-all afferent feedforward (FF) connection matrix $\mathbf{W}$ (Fig. 1a). These input neurons may be grouped into multiple input neuron groups responding to different sensory features. Due to the all-to-all connectivity, the afferent input has no specific predefined structure. Two types of all-to-all modifiable recurrent connections (REC), $\mathbf{M}$ and $\mathbf{G}$, exist among the 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. The firing rate of neurons are defined as a modifiable sigmoidal function

of the membrane potential (Fig.1b), which we will explain later in detail. All types of connections, both afferent and recurrent ones, are modifiable by unsupervised learning rules derived from a common principle: on each neuron, all synapses learn to predict the neuron's response optimally (Fig. 1c: see Materials and Methods). In reality, all synaptic inputs may be terminated on the dendrites, although they are not modeled explicitly.

Without a teaching signal, predictive learning may suffer a trivial solution problem in which all synapses vanish, and hence all neurons become silent (Asabuki & Fukai, 2020). To avoid it, we homeostatically regulate the dynamic range of each neuron (i.e., the slope and threshold of the response function) according to the history h of its subthreshold activity (see Eqs. 10-12). When the value of h is increased, the neuron's excitability is lowered (Fig.1b). The input-output curves of neurons are known to undergo homeostatic regulations through various mechanisms (Chance et al., 2002; Mitchell and Silver, 2003; Torres-Torrelo et al., 2014). Though no direct experimental evidence is available for our homeostatic process via h, it mathematically avoids saturating neuronal activity.

Note that the present homeostatic regulation of intrinsic excitability differs from the homeostatic synaptic scaling mechanism. The role of homeostatic synaptic scaling in generating irregular cell-assembly activity patterns was previously studied computationally (Hiratani and Fukai, 2014; Litwin-Kumar and Doiron, 2014; Zenke et al., 2015). However, unlike the present model, the previous models did not address whether and how synaptic scaling contributes to statistical modeling by recurrent neural networks. Furthermore, unlike our model, in which neurons in the recurrent layer and input neurons are initially connected in an all-to-all manner, most previous models assumed preconfigured receptive fields for recurrent-layer neurons, implying that these models had predefined stimulus-specific cell assemblies.

Cell assembly formation for learning statistically salient stimuli

We first explain how our network segments salient stimuli and forms stimulus-specific cell assemblies via network-wide predictive learning rules. To this end, we tested a simple case in which two non-overlapping input groups are intermittently and repeatedly activated with equal probabilities. The two input patterns were separated by irregular, low-frequency, unrepeated spike trains of all input

neurons (Materials and Methods). We will consider input patterns with unequal occurrence probabilities later. After several presentations of individual input patterns, each network neuron responded selectively to one of the repeated patterns (Figure 2a). This result is consistent with our previous results (Asabuki & Fukai, 2020) that the plasticity of feedforward connections segments input patterns. Indeed, feedforward synapses W on each neuron were strengthened or weakened when they mediated its preferred or non-preferred stimulus, respectively (Fig. 2b, left; Fig. 2c). Inhibitory connections G grew between neurons within the same assembly but not between assemblies (Fig. 2b, right; Fig. 2c, bottom), enhancing the decorrelation of within-assembly neural activities (Asabuki & Fukai, 2020). Recurrent connections M were modified to form stimulus-specific cell assemblies, as evidenced by the self-organization of excitatory (Fig. 2c, top) and inhibitory (Fig. 2c, bottom) recurrent connections within and between cell assemblies, respectively. The inhibitory components are necessary for suppressing the simultaneous replay of different cell assemblies, as shown later.

We then investigated whether and how spontaneous activity preserves and replays these cell assemblies in the absence of afferent input. To demonstrate this in a more complex task, we trained the network with afferent input involving five repeated patterns and then removed the input and observed post-training spontaneous network activity (Fig. 2d). The termination of afferent input initially lowered the activities of neurons, but their dynamic ranges gradually recovered with the excitability of the neural population (indicated by the population-averaged h value), and the network eventually started spontaneously replaying the learned cell assemblies. All plasticity rules were turned off during the recovery period (about 20 seconds from the input termination), after which the network settled in a stable spontaneous firing state (plasticity off). Then, the plasticity rules could be turned on (plasticity on) without drastically destroying the structure of spontaneous replay. Intriguingly, spontaneous neuronal activities were highly correlated within each cell assembly but were uncorrelated between different cell assemblies (Fig. 2e). This was because self-organized recurrent connections $\mathbf{M}$ were excitatory within each cell assembly, whereas the between-assembly recurrent connections were inhibitory, as in Fig. 2c.

Thus, the network model successfully segregates, remembers, and replays stimulus-evoked activity patterns in temporal input. The loss of between-assembly

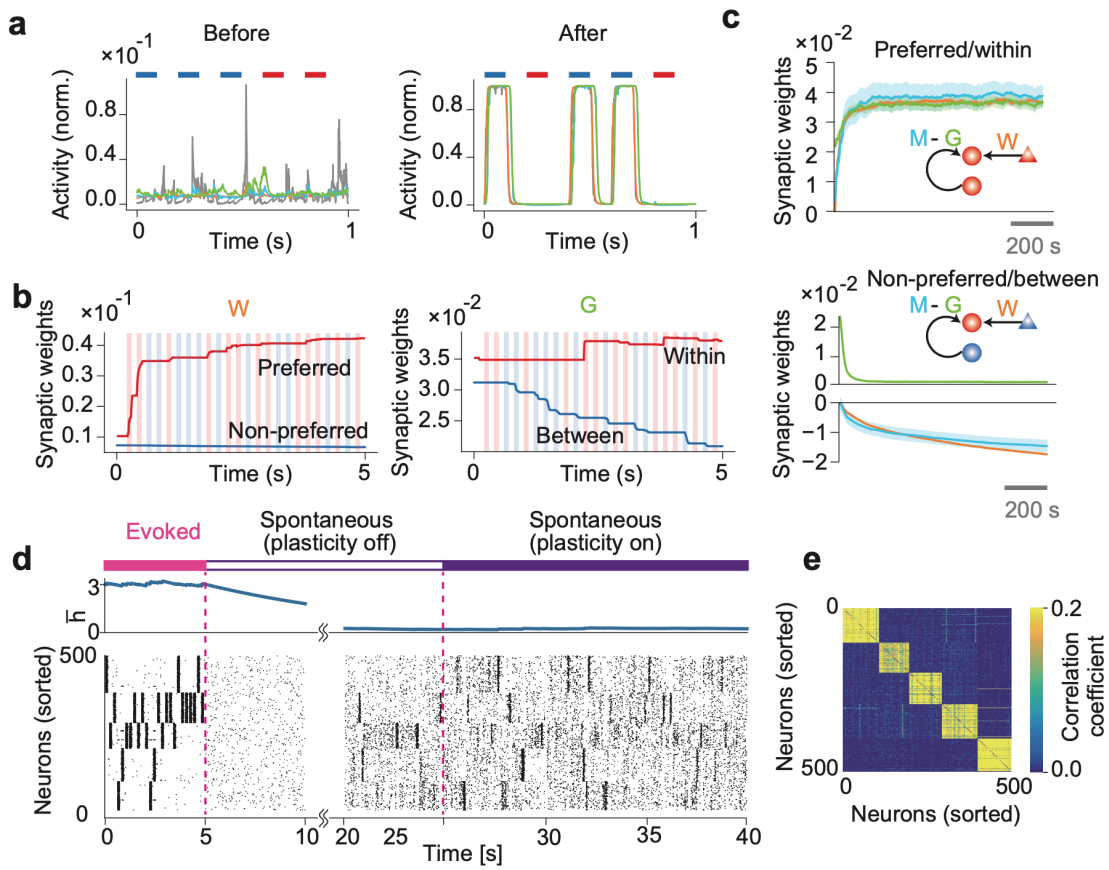

Figure 2. Formation of stimulus-selective assemblies in a recurrent network. (a) Example dynamics of neuronal output and synaptic predictions are shown before (left) and after (right) learning. Colored bars at the top of the figures represent periods of stimulus presentations. (b) Example dynamics of feedforward connection W and inhibitory connection G are shown. W-connections onto neurons organizing to encode the same or different input patterns are shown in red and blue, respectively. Similarly, the same colors are used to represent G connections within and between assemblies. (c) Dynamics of the mean connection strengths are shown on neuron in cell assembly 1. Shaded areas represent SDs. In the schematic, triangles indicate input neurons and circles indicate network neurons. The color of each neuron indicates the stimulus preference of each neuron. (d) Example dynamics of the averaged dynamical variable $\bar{h}$ (top) and the learned network activity (bottom) are shown. The dynamical variables are averaged over the entire network. Neurons are sorted according to their preferred stimuli. During the spontaneous activity, afferent inputs to the network were removed. (e) Correlation coefficients of spontaneous activities of every pair of neurons are shown.

excitatory connections is interesting as it indicates that the present spontaneous reactivation is not due to the sequential activation of cell assemblies. This can also be seen from the relatively long intervals between consecutive cell-assembly activations: spontaneous neural activity does not propagate directly from one cell assembly to another (Fig. 2d). Indeed, within-assembly excitation is the major cause of spontaneous replay in this model, which we will study later in detail.

In summary, we have proposed the predictive learning rules as a novel plasticity

mechanism for all types of synapses (i.e., feedforward and recurrent connections). We have shown that the plasticity rules in our model learn the segmentation of salient patterns in input sequences and form pattern-specific cell assemblies without preconfigured structures. We also showed that our model replays the learned assemblies even when external inputs were removed.

Replays of cell assemblies reflect a learned statistical model

We now turn to the central question of this study. We asked whether internally generated network dynamics through recurrent synapses (i.e., spontaneous replay of cell assemblies) can represent an optimal model of previous sensory experiences. Specifically, we examined whether the network spontaneously reactivates learned cell assemblies with relative frequencies proportional to the probabilities with which external stimuli activated these cell assemblies during learning. We addressed these questions in slightly more complex cases with increased numbers of external stimuli.

We first examined a case with five stimuli 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. 2d represents the probability ratio one. As in Fig. 2d, 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 (Materials and Methods), scales faithfully with the probability ratio of cell assembly 1. This was the case: the scaling was surprisingly accurate (Fig. 3e).

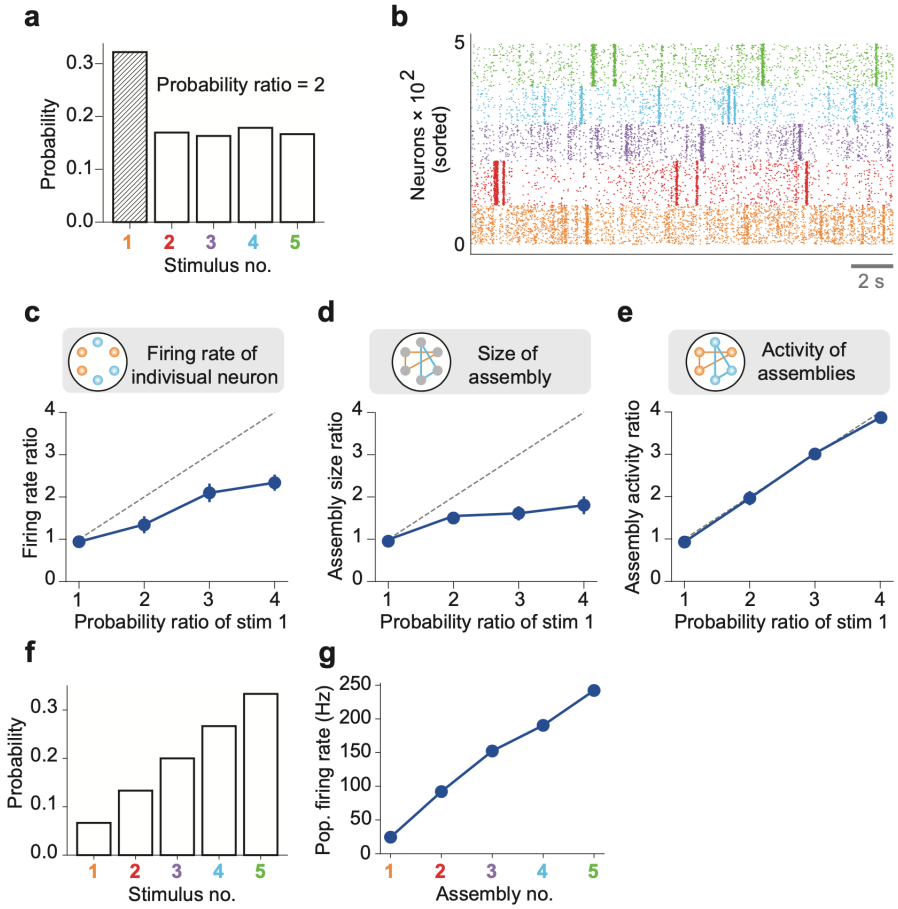

Figure 3 Priors coded in spontaneous activity. An nDL network was trained with five probabilistic inputs. (a) Stimulus 1 appeared twice as often as the other four stimuli during learning. The example empirical probabilities of the stimuli used for learning are shown. (b) The spontaneous activity of the 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 the total activities of neurons in assembly 1 to those of the other assemblies are shown. (f) Five stimuli occurring with different probabilities were used for training the nDL model. (g) The population firing rates are shown for five self-organized cell assemblies encoding the stimulus probabilities shown in (f).

To examine the ability of the nDL network further, we trained it with five stimuli occurring with various probabilities (Fig. 3f and Supplementary Fig. 1a). After learning, the spontaneous activity of the model replayed the learned cell assemblies at the desired ratios of population firing rates (Fig. 3g and Supplementary Fig. 1b).

We then asked whether our model would learn a prior distribution for more stimuli. To this end, we presented seven stimulus patterns to the same network with

graded probabilities (Supplementary Fig. 1c). The self-organized spontaneous activity exhibited cell assemblies that well learned the graded probability distribution of these stimuli (Supplementary Fig. 1d). These results demonstrate that the trained network remembers the probabilities of repetitively experienced stimuli by the spontaneous firing rates of the encoding cell assemblies and that this dynamical coding scheme has a certain degree of scalability.

So far, we have represented external stimuli with non-overlapping subgroups of input neurons. However, in biologically realistic situations, input neuron groups may share part of their membership neurons. We tested whether the proposed model could learn the probability structure of overlapping input patterns in a case where two input neuron groups shared half of their members. The two patterns were presented with probabilities of 30% and 70%, respectively (Supplementary Fig. 2a). After sufficient learning, the network model generated two assemblies that encoded the two stimuli without sharing the coding neurons (Supplementary Fig. 2b) and replayed these assemblies with frequencies proportional to the stimulus presentation probabilities (Supplementary Fig. 2c). The results look reasonable because each neuron in the network segments one of the stimulus patterns and recurrent connections within each non-overlapping assembly can encode the probability of its replay.

Altogether, these results suggest that our model spontaneously replays learned cell assemblies with relative frequencies proportional to the probability that each cell assembly was activated during the learning phase. We have shown that the population activities of assemblies, rather than the firing rates of individual neurons, encode the occurrence probabilities of stimulus patterns.

Within-assembly recurrent connections encode probabilistic sensory experiences

To understand the mechanism underlying the statistical similarity between the evoked patterns and spontaneous activity, we then investigated whether and how biases in probabilistic sensory experiences influence the strengths of recurrent connections. To this end, we compared two cases in which two input patterns (stim 1 and stim 2) occurred with equal (50% vs. 50%) and different (30% vs. 70%) probabilities during learning (Fig. 4a). From the results shown in Fig. 3, we hypothesized that within-assembly learned connections should reflect the

stimulus occurrence probabilities and hence the activation probabilities of the corresponding cell assemblies during spontaneous activity. Therefore, we calculated the total strengths of incoming recurrent synapses on each neuron within the individual cell assemblies (Fig. 4b). While the distributions of incoming synaptic strengths are similar between cell assemblies coding stimulus 1 and stimulus 2 in the 50-vs-50 case, they look different in the 30-vs-70 case (Fig. 4c).

Since incoming weights increased more significantly in the cell assembly activated by a more frequent stimulus (i.e., the assembly encoding stimulus 2 in the 30-vs-70 case), we expect that the degree of positive shifts in incoming weight distributions will reflect stimulus probabilities. To examine whether this is indeed the case, we computed the sum of total excitatory incoming weights (i.e., the sum of positive elements of M) over neurons belonging to each assembly after training. We then normalized these excitatory incoming weights over the two assemblies.

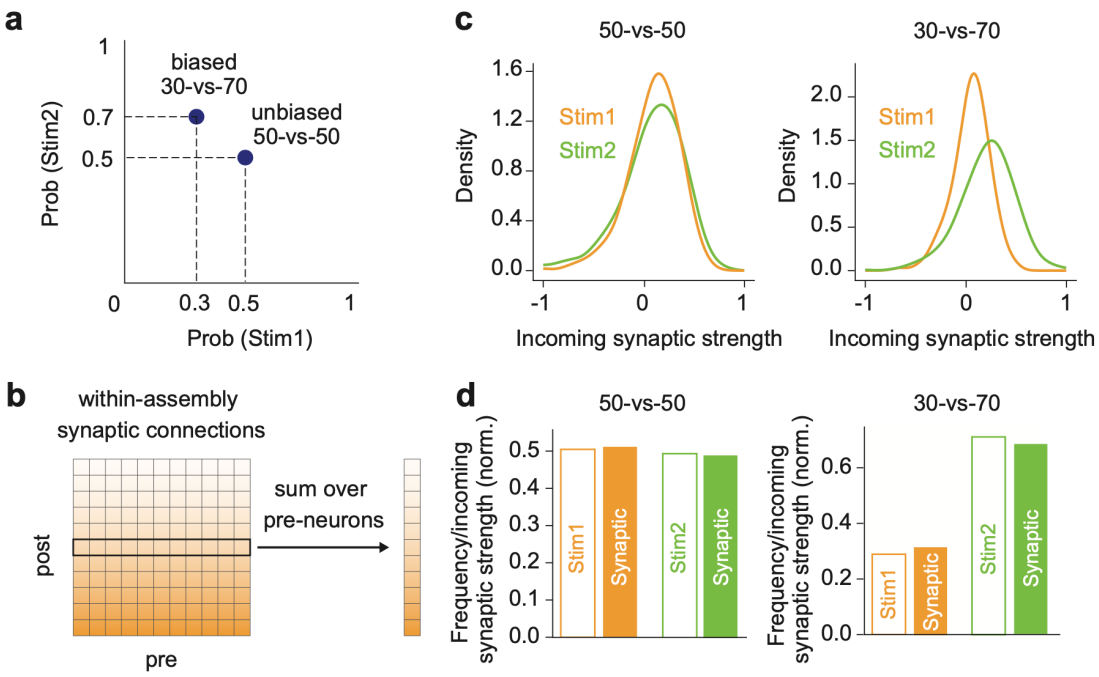

Figure 4. Probability encoding by learned within-assembly synapses. (a) Two input stimuli were presented in two protocols: uniform (50% vs. 50%) or biased (30% vs. 70%). (b) The total incoming synaptic strength on each neuron was calculated within each cell assembly. (c) *left,* The distributions of incoming synaptic strength are shown for the learned assemblies in the 50-vs-50 case. *right*, Same as in the left figure, but in the 30-vs-70 case. (d) *left,* The empirical probabilities of stimuli 1 and 2 and the normalized excitatory incoming weights within assemblies are compared in the 50-vs-50 case. *right*, Same as in the left figure, but in the 30-vs-70 case.

Interestingly, we found that the normalized excitatory incoming weights for the two assemblies well approximates the empirical probabilities of the two stimuli in both the 50-vs-50 and 30-vs-70 cases (Figure 4d). These analyses revealed that recurrent connections learned within assemblies encode biases in probabilistic sensory experiences.

Roles of inhibitory plasticity for stabilizing cell assemblies

Experimental and computational results suggest that inhibitory synapses are more robust to spontaneous activity than excitatory synapses and are crucial for maintaining cortical circuit function (Mongillo et al., 2018). To see the crucial role of the inhibitory plasticity of G for cell assembly formation, we compared the spontaneously driven activities in the learned network between two cases, plastic inhibitory connection G versus fixed G, in the 30-vs-70 case. The results show that only a single, highly active assembly self-organizes for fixed inhibitory synapses (Supplementary Fig. 3a). In contrast, such winner-take-all dynamics do not emerge from plastic inhibitory synapses (Supplementary Fig. 3b), suggesting the crucial role of inhibitory plasticity in stabilizing spontaneous activity.

To further clarify the functional role of inhibitory plasticity in regulating spontaneous activity, we compared how the self-organized assembly structure of recurrent connections $\mathbf{M}$ evolves in the two simulation settings shown in Supplementary 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 (Supplementary Fig. 4b, magenta). Consistent with this, the postsynaptic potentials mediated by connections $\mathbf{M}$ and $\mathbf{G}$ predicted the normalized firing rate of a postsynaptic excitatory neuron in the control model (Supplementary 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 (Supplementary 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 (Supplementary 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 (Supplementary Fig. 4e). These results indicate that maintaining the

learned representations requires the continuous tuning of within-assembly inhibition.

The role of homeostatic regulation of neural activities

As indicated by the weak couplings between cell assemblies, the present mechanism of probability learning differs from the conventional sequence learning mechanisms. Consistent with this, the network trained repetitively by a fixed sequence of patterned inputs does not exhibit stereotyped sequential transitions among cell assemblies (due to the lack of strong inter-assembly excitatory connections; Supplementary Fig. 5). Indeed, the probability-encoding spontaneous activity emerges in the present model mainly from the intra-assembly dynamics driven by strong within-assembly reverberating synaptic input. However, homeostatic variable h also plays a role in maintaining a stable spontaneous network activity after learning (see Fig. 2d; activity pattern from 5 to 10 sec). This is achieved by the time evolution of h, which maintains the firing rate of each neuron in a suitable range by adjusting the threshold and gain of the somatic sigmoidal response function (Fig. 1b).

Therefore, we explored the role of the homeostatic variable in learning an accurate internal model of the sensory environment. In each neuron, the variable h is updated whenever the membrane potential undergoes an abrupt increase (Eq. 10). Therefore, the time evolution of h monitors the approximate duration of active epochs of the neuron and, consequently, its membership cell assembly (Supplementary Fig. 6a). The total duration of active epochs determines the activation probability of each cell assembly. Furthermore, when the instantaneous value of h is high, the neuron's excitability is lowered (namely, the gain and threshold of the response function are decreased or increased, respectively: see Eqs. 11 and 12). The gain modulation enables the neuron to homeostatically regulate its susceptibility to patterned inputs and hence is crucial for measuring the duration of its active epochs. Actually, a model with fixed value of h showed spontaneous replay with less accurate estimates of the true probability distribution (Supplementary Fig. 6b: cf. Fig. 3f). In addition, the elimination of between-assembly excitatory connections did not significantly affect the replay probabilities, as the variable h is driven by strong within-assembly recurrent inputs after learning (Supplementary Fig. 6c).

As demonstrated above, the mechanism underlying our learning rule is significantly different from the previously proposed rule. Most previous models perform stochastic sampling based on Markov chains of network dynamics, which requires precise wiring patterns between assemblies. In contrast, in our model, individual assemblies sample independently via homeostatically regulated activities, and within-assembly connectivity alone is sufficient to perform such sampling.

Learning conditioned prior distributions

The predictive coding hypothesizes that top-down input from higher cortical areas provides prior knowledge about computations in lower cortical areas. This implies in the brain's hierarchical computation that the top-down input conditions the prior distributions in local cortical areas to those relevant to the given context. The proposed learning rules can account for how a conditioned input from other cortical areas conditions the prior distribution in a local cortical circuit.

The neural network consists of two mutually interacting non-overlapping subnetworks of equal sizes, where the subnetworks may represent different cortical areas (Supplementary Fig. 7a). 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 (Supplementary Fig. 7b). Consistent with this, the self-organized connection matrix represented strong within-assembly connections within each cell assembly and weak between-assembly connections (Supplementary Fig. 7c). 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 (Supplementary Fig. 7d). 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 (Supplementary Fig. 7e). Note that S3- and S4-coding cell assemblies could become simultaneously active to represent the desired activation

probabilities (e.g., a vertical arrow in Supplementary Fig. 7d). Together, these results indicate that our network can learn prior distributions conditioned by additional inputs through different pathways.

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 (Hanks et al., 2011). 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.

We constructed a network model shown in Fig. 5a 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 (Materials and Methods). During learning, we trained this model with external stimuli having input coherence Coh of either -0.5 or +0.5 (Materials and 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

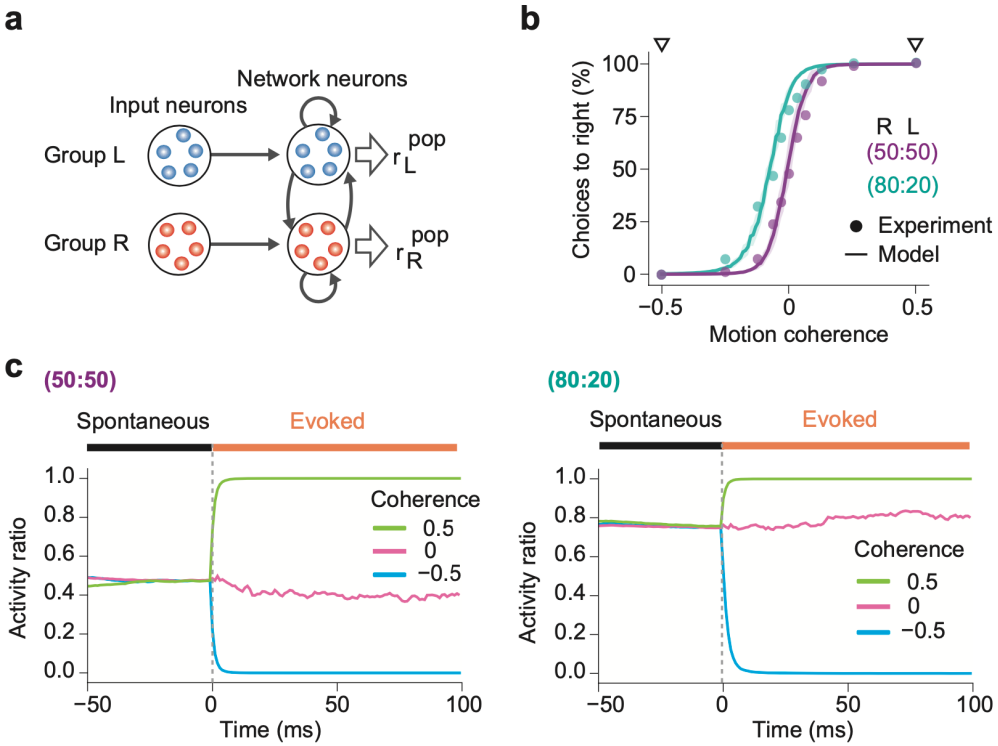

Figure 5. Simulations of biased perception of visual motion coherence. (a) The network model simulated perceptual decision-making of coherence in random dot motion patterns. In the network shown here, network neurons have already learned two assemblies encoding leftward or rightward movements from input 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 (Materials and 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.

having various degrees of input coherence. The resultant psychometric curves almost perfectly coincide with those obtained in the experiment (Fig. 5b). 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 finetuning.

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 relatively simple cases, i.e., Coh = -0.5, 0, and +0.5. Figure 5c shows the activity ratio AR between the R-encoding cell assembly and the entire network (Materials and 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. Our results show that the spontaneous activity generated by the proposed mechanism can account for the precise relationship between motion coherence and perceptual biases in decision making by monkeys.

An elaborate network model with distinct excitatory and inhibitory neuron pools

The predictive learning rule performed well in training the nDL 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 (DL model) consisting of distinct excitatory and inhibitory neuron pools, obeying Dale's law (Supplementary Fig. 8a). The nDL model suggested the essential roles of inhibitory plasticity in maintaining excitation-inhibition balance and generating an appropriate number of cell assemblies. To achieve these functions, inhibitory neurons in the DL model project to excitatory and other inhibitory neurons via two synaptic paths (Supplementary Fig. 8b). In path1, inhibitory connections alone predict the postsynaptic activity, whereas inhibitory and excitatory connections jointly predict the activity of the postsynaptic neuron in path2 (Materials and Methods). All synapses in the DL model are subject to the predictive learning rule. We trained the DL model with three input neuron groups while varying their activation probabilities. As in the nDL model, the DL model self-organized three cell assemblies activated selectively by the three input neuron groups (Supplementary Fig. 9a). Furthermore, in the absence of external stimuli, the DL model spontaneously replayed these assemblies with the assembly activity ratios in proportion to the occurrence probabilities of the corresponding stimuli during learning

(Supplementary Fig. 8c).

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 (Supplementary Fig. 8d, top). In contrast, in path 2, inhibitory connections were stronger on excitatory neurons in different assemblies than those in the same assemblies (Supplementary Fig. 8d, bottom). On both excitatory and inhibitory neurons, the total inhibition (i.e., path 1 + path 2) was balanced with excitation (Supplementary Fig. 8e). Supplementary Figure 8f summarizes the connectivity structure of the DL 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 (Supplementary Fig. 8g). 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. Supplementary Figure 8h 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 (Fukai & Tanaka, 1997). 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. 9b, c).

In summary, we have shown the roles of distinct recurrent inhibitory connections. Using a network consisting of excitatory and inhibitory populations, we have shown that distinct inhibitory circuits are necessary to generate within- and

between-assembly competition crucial to maintain the stability of learned multiple assemblies.

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 (Friston, 2010; Bastos et al., 2012; Keller & Mrsic-Flogel, 2018). 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.

Our model adopted a single predictive learning principle for the plasticity of excitatory and inhibitory synapses to learn the replay of probabilistic experiences. 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. The latter model predicts a division of labor between two inhibitory paths. Intriguingly, the inhibitory path 2 of this model resembles interpyramidal inhibitory connections driven directly by nearby pyramidal cells (Ren et al., 2007). 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 (Vogels et al., 2013; Sprekeler, 2017).Various models have been proposed to account for neural mechanisms of Bayesian computation by the brain (Tully et al., 2014; Kappel et al., 2015; Hiratani & Fukai, 2018; Hiratani & Latham, 2020; Aitchison et al., 2021; Ma et al., 2006; Deneve, 2008; Nessler et al., 2013; Hiratani & Fukai, 2016; Huang & Rao, 2016; Isomura et al., 2022; Friston, 2010; Bastos et al., 2012; Keller & Mrsic-Flogel, 2018). Typically, these models embed prior knowledge on sensory experiences into the wiring patterns of afferent (and sometimes also recurrent) synaptic inputs such that these

inputs can evoke the learned activity patterns associated with 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). To our knowledge, our model is the first to perform all these steps for encoding an optimal model of the environment into spontaneous network activity.

The present mechanism of memory formation differs from the previous ones that self-organize cell assemblies through Hebbian learning rules (Vogels et al., 2011; Hiratani and Fukai, 2014; Litwin-Kumar and Doiron, 2014; Zenke et al., 2015; Triplett et al., 2018; Montangie 2020). Firstly, these mechanisms did not aim for explicit statistical modeling of the environment. Secondly, the previous studies suggested that the orchestration of multiple plasticity rules, including inhibitory plasticity and homeostatic synaptic scaling, enables the maintenance of cell assemblies (however, see Manz et al., 2023). For instance, in spike-timing-dependent plasticity (STDP), slight changes in the relative times of pre and postsynaptic spikes can change the polarity of synaptic modifications, implying that STDP requires a mechanism to keep synaptic weights finite (Kempter et al., 1999; Song et al., 2000; Masquelier et al., 2008). In contrast, our learning rule, which induces either long-term potentiation or depression according to the sign of the prediction error calculated independently within each postsynaptic neuron, does not suffer such instability.

Our model predicts a novel intracellular process that regulates the neuron's dynamic range according to the history of its subthreshold dynamics. This process plays two important roles in the statistical modeling of our model. First, it avoids the trivial solution (i.e., the zero-weight solution) of our unsupervised predictive learning by homeostatically regulating neurons' intrinsic excitability. Second, the intracellular process cooperates with reverberating synaptic inputs within each cell assembly to generate spontaneous replay activity. Owing to the intracellular homeostasis, our model can sample from the learned distribution without relying on the recurrences among assemblies. This mechanism contrasts with the previous sampling-based models that rely on the transition dynamics between cell

assemblies (Buesing et al., 2011; Bill et al., 2015). How neural systems implement the proposed homeostasis is an open question.

The proposed mechanism can account for the behavioral biases observed in perceptual decision making (Hanks et al., 2011). This behavioral experiment quantitatively clarified how the difference in the probability between sensory experiences during learning biases the alternative choice behavior of monkeys. In our model, two cell assemblies encoding the different stimuli are replayed at the total firing rates proportional to the corresponding occurrence probabilities. Our results suggest that the difference in spontaneous firing rates of cell assemblies is sufficient to explain the behavioral biases of monkeys. However, other mechanisms, such as biased top-down input, cannot be excluded.

What could be the advantages of coding prior distributions into spontaneous activity over other ways of probability coding? First, spontaneous replay activities in lower cortical areas may provide training data for modeling by higher cortical areas, promoting hierarchical statistical modeling in predictive coding. This is analogous to the situation where hippocampal engram cells are replayed to reinforce the activity patterns of cortical engrams for memory consolidation during sleep (Tonegawa et al., 2018; Ghandour et al., 2019; Klinzing et al., 2019; Takehara-Nishiuchi, 2021). Memory reinforcement by activity replay has also been studied in machine intelligence (Dayan et al., 1995; Goodfellow et al., 2014; Luczak et al., 2022). Second, spontaneous replay of internal models may support knowledge generalization during sleep. It was recently reported that a transitive inference task requires post-learning sleep (Kareem et al., 2021). In this task, mice had to infer a correct reward delivery rule in a novel behavioral situation from the outcomes of past experiences. The mice failed to generalize the learned rules if the activity of the anterior cingulate cortex was suppressed during post-learning sleep, suggesting that dynamic interactions among rule-coding cortical neurons in spontaneous activity are crucial for rule generalization. Clarifying how spontaneous brain activity generalizes the learned internal models is an intriguing open question.

Methods

Neural network model

Below, we first describe the model architecture and learning rule for the nDL 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 $\hat{\varphi}(u)$, where $\hat{\varphi}(\cdot)$ is a dynamics sigmoidal function, which we will explain later. 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

$$f_i(t) = \hat{\varphi}(u_i(t); h_i), \qquad (8)$$

in terms of a dynamical sigmoidal response function $\hat{\varphi}$:

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

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

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

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:

$$\Delta \mathrm{W}_{ij} = \eta \mathcal{E}\left(f_i, \mathrm{v}_i^{\mathrm{W}}\right) x_j, \tag{14a}$$

$$\Delta \mathrm{M}_{ij} = \eta \mathcal{E}\left(f_i, \mathrm{v}_i^{\mathrm{M}}\right) y_j, \tag{14b}$$

$$\Delta \mathrm{G}_{ij} = \eta \mathcal{E}\left(f_i, \mathrm{v}_i^{\mathrm{G}}\right) y_j, \tag{14c}$$

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

$$\mathcal{E}(f_i, \mathrm{v}_i) = \varphi_0{}^{-1}(1 - f_i/\varphi_0)[f_i - \varphi(\mathrm{v}_i)], \tag{15}$$

with a static sigmoidal function:

$$\varphi(\mathrm{v}_i) = \varphi_0\left[1 + \exp[g\beta_0(-u_i + g\theta_0)]\right]^{-1}.$$

Throughout this study, the learning rate $\eta = 10^{-4}$, 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.

Stimulation protocols

Feedforward input to the recurrent network consisted of K Poisson spike trains with a background firing rate of 2Hz. The input randomly presented n non-overlapping patterns of 100 spike trains (the duration 100 ms and the mean frequency 50 Hz), one at a time, with pattern-to-pattern intervals of 100 ms. Therefore, the number of input neurons and patterns satisfy the relationship of $K = 100 \times n$. For

simplicity, we simulated the constant-interval case, but using irregular intervals does not change the essential results. The value of n varies from task to task, and the values for each figure are as follows: $n = 5$ (Fig. 2c-e, Fig. 3); $n = 2$ (Fig. 2a-b, Figs. 4-5).

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:

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

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

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

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

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

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

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

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. (2011), 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. 5c, we calculated the activity ratio (AR) as

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

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. 5b, we defined "choices to right" as

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

A network model with distinct excitatory and inhibitory neuron populations

Here, we explain the architecture of the model used in Supplementary Fig.8. 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:

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

$$- \underbrace{\sum_{m=1}^{N_I} \mathrm{G}_{im}^{\mathrm{XI\,(path1)}} y_m^I(t)}_{=:\mathrm{v}_i^{\mathrm{M}(1)}}, \tag{22}$$

where $\{\mathrm{W}_{ik}^{\mathrm{X}}\}$ is afferent synaptic weights, which are a mixture of excitatory and inhibitory connections as in the nDL 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:

$$\Delta \mathrm{W}_{ij}^{\mathrm{X}} = \eta \mathcal{E}\left(f_i, \mathrm{v}_i^{\mathrm{W}}\right) x_j, \tag{23a}$$

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

$$\Delta \mathrm{G}_{ij}^{\mathrm{XI\,(path2)}} = \eta \mathcal{E}\left(f_i, \mathrm{v}_i^{\mathrm{M}(2)}\right) y_j^I, \tag{23c}$$

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

To satisfy the Dale's law, we truncated all weights of recurrent connections to non-

negative values during learning.

In Supplementary Fig. 8g, we measured the lateral inhibition between excitatory neurons via path 1 by calculating:

$$\left[\mathrm{W}_{i,j}^{\mathrm{LI}}\right] = \sum_{k=1}^{N_E} \mathrm{G}_{ik}^{\mathrm{EI\ (path1)}} \mathrm{M}_{kj}^{\mathrm{IE}}. \qquad (24)$$

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

Simulation details

All simulations were performed in customized Python3 code written by TA with numpy 1.17.3 and scipy 0.18. Differential equations were numerically integrated using a Euler method with integration time steps of 1 ms.

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

Competing Interest Statement: The authors declare no competing interests.

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 and 19H04994) to T.F.

References

1. 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).
2. Asabuki, T. & Fukai, T. Somatodendritic consistency check for temporal feature segmentation. *Nat. Commun.* **11**, 1554 (2020).
3. Ashourian, P. & Loewenstein, Y. Bayesian inference underlies the contraction bias in delayed comparison tasks. *PLoS One* **6**, e19551 (2011).
4. 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).

5. 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).
6. 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).
7. Bill, J., Buesing, L., Habenschuss, S., Nessler, B., Maass, W., & Legenstein, R. Distributed bayesian computation and self-organized learning in sheets of spiking neurons with local lateral inhibition. *PloS one*, **10**, e0134356 (2015).
8. Buesing, L., Bill, J., Nessler, B., & Maass, W. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS computational biology*, **7**, e1002211 (2011).
9. Chance, F. S., Abbott, L. F., & Reyes, A. D. Gain modulation from background synaptic input. *Neuron*, **35**, 773–782 (2002).
10. Dayan, P., Hinton, G.E., Neal, R.M. & Zemel, R.S. The Helmholtz machine. *Neural Comput.* **7**, 889-904 (1995).
11. 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).
12. Deneve, S. Bayesian spiking neurons I: inference. *Neural Comput.* **20**, 91-117 (2008).
13. Ernst, M.O. & Banks, M.S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429-433 (2002).
14. 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).
15. Fonollosa, J., Neftci, E., & Rabinovich, M. Learning of chunking sequences in cognition and behavior. *PLoS computational biology*, **11**, e1004592 (2015).
16. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127-138 (2010).
17. 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).
18. Fujii, N. & Graybiel, A. M. Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science* **301**, 1246–1249 (2003).

19. 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).
20. 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).
21. Goodfellow, I.*, et al.* Generative Adversarial Nets. in *Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014)* 2672–2680 (2014).
22. Hachen, I., Reinartz, S., Brasselet, R., Stroligo, A. & Diamond, M.E. Dynamics of history-dependent perceptual judgment. *Nat. Commun.* **12**, 6036 (2021).
23. 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).
24. Hartmann, C., Lazar, A., Nessler, B., & Triesch, J. Where's the noise? Key features of spontaneous activity and neural variability arise through learning in a deterministic network. *PLoS computational Computational bBiology*, **11**, e1004640 (2015).
25. Hiratani, N. & Fukai, T. Hebbian wiring plasticity generates efficient network structures for robust inference with synaptic weight plasticity. *Front. Neural Circuits* **10**, 41 (2016).
26. Hiratani, N. & Fukai, T. Interplay between short- and long-term plasticity in cell-assembly formation. *PLoS One* **9**, e101535 (2014).
27. Hiratani, N. & Fukai, T. Redundancy in synaptic connections enables neurons to learn optimally. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E6871-E6879 (2018).
28. Hiratani, N. & Latham, P.E. Rapid Bayesian learning in the mammalian olfactory system. *Nat. Commun.* **11**, 3845 (2020).
29. Huang, Y. & Rao, R.P. Bayesian Inference and Online learning in Poisson neuronal networks. *Neural Comput.* **28**, 1503-1526 (2016).
30. Isomura, T., Shimazaki, H. & Friston, K.J. Canonical neural networks perform active inference. *Commun. Biol.* **5**, 55 (2022).

31. Jimenez Rezende D. & Gerstner W. Stochastic variational learning in recurrent spiking networks. Front Comput Neurosci. 8:38 doi: 10.3389/fncom.2014.00038 (2014).
32. Jin., X. & Costa, R. M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457–462 (2010).
33. Jin., X., Tecuapetla, F. & Costa, R. M. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat. Neurosci.* **17**, 423–430 (2014).
34. Kappel, D., Habenschuss, S., Legenstein, R. & Maass, W. Network plasticity as Bayesian inference. *PLoS Comput. Biol.* **11**, e1004485 (2015).
35. 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.
36. Keller, G.B. & Mrsic-Flogel, T.D. Predictive processing: a canonical cortical computation. *Neuron* **100**, 424-435 (2018).
37. Kempter R., Gerstner W. & Van Hemmen, J. L. Hebbian learning and spiking neurons. *Physical Review E*. **59**, 4498 (1999).
38. Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A. & Arieli, A. Spontaneously emerging cortical representations of visual attributes. *Nature* **425**, 954-956 (2003).
39. Kiebel, S. J., von Kriegstein, K., Daunizeau, J., & Friston, K. J. Recognizing sequences of sequences. *PLoS computational biology*, **5**, e1000464 (2009).
40. Klinzing, J.G., Niethard, N. & Born, J. Mechanisms of systems memory consolidation during sleep. *Nat. Neurosci.* **22**, 1598-1610 (2019).
41. Kording, K.P. & Wolpert, D.M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244-247 (2004).
42. Legaspi, R. & Toyoizumi, T. A Bayesian psychophysics model of sense of agency. *Nat. Commun.* **10**, 4250 (2019).
43. 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).
44. Litwin-Kumar, A. & Doiron, B. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat Commun.* **5**, 1–12 (2014).
45. Luczak, A., McNaughton, B.L. & Kubo, Y. Neurons learn by predicting future activity. *Nat. Mach. Intell.* **4**, 62-72 (2022).

46. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432-1438 (2006).
47. Manz, P., Memmesheimer, R.M. Purely STDP-based assembly dynamics: Stability, learning, overlaps, drift and aging. *PLoS Comput Biol.* **19**, e1011006 (2023).
48. Masquelier T., Guyonneau R., & Thorpe S. J. Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. *PloS one*. **3**, e1377 (2008).
49. Mitchell, S. J., & Silver, R. A. Shunting inhibition modulates neuronal gain during synaptic excitation. *Neuron*, **38**, 433–445 (2003).
50. Mongillo G, Rumpel S, Loewenstein Y. Inhibitory connectivity defines the realm of excitatory plasticity. Nat Neurosci. 21:1463-1470 (2018).
51. Montangie, L., Miehl, C. & Gjorgjieva, J. Autonomous emergence of connectivity assemblies via spike triplet interactions. *PLoS Comput Biol*.**16**, e1007835 (2020).
52. 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).
53. Orban, G., Berkes, P., Fiser, J. & Lengyel, M. Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron* **92**, 530-543 (2016).
54. 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).
55. Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. Bayesian learning of visual chunks by human observers. Proceedings of the National Academy of Sciences, **105**, 2745-2750 (2008).
56. Rabinovich, M. I., Varona, P., Tristan, I., & Afraimovich, V. S. Chunking dynamics: heteroclinics in mind. *Frontiers in computational neuroscience*, **8** (2014).
57. Ren, M., Yoshimura, Y., Takada, N., Horibe, S., & Komatsu, Y. (2007). Specialized inhibitory synaptic actions between nearby neocortical pyramidal neurons. Science, 316(5825), 758-761.
58. Reynolds, J. R., Zacks, J. M., & Braver, T. S. A computational model of event segmentation from perceptual prediction. *Cognitive science*, **31**, 613-643 (2007).

59. Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. *Nat. Neurosci*. **16**, 486–492 (2013).
60. Song S., Miller, K. D. & Abbott, L. F. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience*. **3**, 919-926 (2000).
61. Sprekeler, H. Functional consequences of inhibitory plasticity: homeostasis, the excitation-inhibition balance and beyond. *Curr. Opin. Neurobiol.* **43**, 198-203 (2017).
62. Takehara-Nishiuchi, K. Neurobiology of systems memory consolidation. *Eur. J. Neurosci.* **54**, 6850-6863 (2021).
63. 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).
64. Torres-Torrelo, J., Torres, B., & Carrascal, L. Modulation of the input-output function by GABAA receptor-mediated currents in rat oculomotor nucleus motoneurons. *The Journal of physiology*, **592**, 5047–5064 (2014).
65. Triplett, M.A., Avitan, L. & Goodhill, G.J. Emergence of spontaneous assembly activity in developing neural networks without afferent input. *PLoS Comput Biol*. **14**, e1006421 (2018).
66. Tully, P.J., Hennig, M.H. & Lansner, A. Synaptic and nonsynaptic plasticity approximating probabilistic inference. *Front. Synaptic Neurosci.* **6**, 8 (2014).
67. Urbanczik, R. & Senn, W. Learning by the dendritic prediction of somatic spiking. *Neuron*. **81**, 521–528 (2014).
68. Vogels T. P., Sprekeler H., Zenke F., Clopath C., & Gerstner W. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*. **334**, 1569-1573 (2011).
69. Vogels, T.P.*, et al.* Inhibitory synaptic plasticity: spike timing-dependence and putative network function. *Front. Neural Circuits* **7**, 119 (2013).
70. Zacks, J. M. et al. Human brain activity time-locked to perceptual event boundaries. *Nat. Neurosci.* **18**, 449–455 (2001).
71. Zenke F, Agnes EJ, Gerstner W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat Commun*. **6**, 1–13 (2015).