

Embedding stochastic dynamics of the environment in spontaneous activity by prediction-based plasticity

Reviewed Preprint

v2 • December 3, 2024

Revised by authors

Reviewed Preprint

v1 • April 12, 2024

Toshitake Asabuki , Claudia Clopath 

Department of Bioengineering, Imperial College London, London, UK • RIKEN Center for Brain Science, RIKEN ECL Research Unit, Wako, Japan • RIKEN Cluster for Pioneering Research, Wako, Japan

 https://en.wikipedia.org/wiki/Open_access Copyright information

eLife Assessment

This is an **important** study that investigates how neural networks can learn to stochastically replay presented sequences of activity according to learned transition probabilities. The authors use error-based excitatory plasticity to minimize the difference between internally predicted activity and stimulus-driven activity, and inhibitory plasticity to maintain E-I balance. The approach is **solid** but the choice of learning rules and parameters is not always justified, with some unclear aspects to the formal derivation.

<https://doi.org/10.7554/eLife.95243.2.sa3>

Abstract

The brain learns an internal model of the environment through sensory experiences, which is essential for high-level cognitive processes. Recent studies show that spontaneous activity reflects such learned internal model. Although computational studies have proposed that Hebbian plasticity can learn the switching dynamics of replayed activities, it is still challenging to learn dynamic spontaneous activity that obeys the statistical properties of sensory experience. Here, we propose a pair of biologically plausible plasticity rules for excitatory and inhibitory synapses in a recurrent spiking neural network model to embed stochastic dynamics in spontaneous activity. The proposed synaptic plasticity rule for excitatory synapses seeks to minimize the discrepancy between stimulus-evoked and internally predicted activity, while inhibitory plasticity maintains the excitatory-inhibitory balance. We show that the spontaneous reactivation of cell assemblies follows the transition statistics of the model's evoked dynamics. We also demonstrate that simulations of our model can replicate recent experimental results of spontaneous activity in songbirds, suggesting that the proposed plasticity rule might underlie the mechanism by which animals learn internal models of the environment.

Introduction

The brain is thought to use its sensory experience to learn an appropriate internal model of the environment, which can improve perception and behavioral performance. (Merfeld et al., 1999 [↗](#); Lewald and Ehrenstein, 1998 [↗](#); Bell et al., 1997 [↗](#); Yasui and Young, 1975 [↗](#); Wolpert et al., 1995 [↗](#)). Such learning is thought to be fundamental to higher-order cognitive processes such as perception, decision making, and prediction of sensory stimuli. Recent computational and experimental evidence suggests that the brain's learned internal model may be reflected in spontaneous activity. For example, in the visual cortex of awake ferrets, spontaneous activity shows spatial similarity to activity elicited by natural scenes (Berkes et al., 2011 [↗](#)). Furthermore, hippocampus generates sequential replay of place fields during rest and sleep (Wilson and McNaughton, 1994 [↗](#); Skaggs and McNaughton, 1996 [↗](#); Lee and Wilson, 2002 [↗](#)). Such hippocampal replay occurs in a highly stereotyped temporal order, with the same sequence of replayed activities often observed across multiple events (Davidson et al., 2009 [↗](#); Diba and Buzsáki, 2007 [↗](#); Gupta et al., 2010 [↗](#); Wu and Foster, 2014 [↗](#)).

Several computational studies have proposed variants of Hebbian plasticity rules for learning deterministic or even stochastic switching dynamics of replayed activities (Levy et al., 2001 [↗](#); Litwin-Kumar and Doiron, 2014 [↗](#); Triplett et al., 2018 [↗](#); Ocker and Doiron, 2019 [↗](#); Asabuki and Fukai, 2023 [↗](#)). However, it has been challenging to extend these results to generate dynamic spontaneous activity obeying appropriate transition probabilities learned through sensory experience. Finding a plasticity rule which is capable of learning structured transitions in spontaneous activity could be instrumental for understanding the mechanism underlying cognitive processes in the brain.

In this paper, we propose a local biologically-plausible plasticity rule for learning the statistical transitions between assemblies in spontaneous activity. We use a recurrent spiking neural network model consisting of distinct excitatory and inhibitory populations. The proposed synaptic plasticity rule for excitatory synapses seeks to minimize the discrepancy between stimulus-evoked and internally predicted activity, while inhibitory plasticity maintains the excitatory-inhibitory balance. We explore the potential performance of our model by learning the Markovian transition statistics of evoked network states. Our results show that the trained model exhibits spontaneous stochastic transitions of cell assemblies, even after structured external inputs are removed. We show that the transition statistics of spontaneous activity show a striking similarity to those of the evoked dynamics.

To further validate our model, we compare the model behavior with recent experimental results in songbirds (Bouchard and Brainard, 2016 [↗](#)), which show that the uncertainty of upcoming states in a bird song modulates the degree of neural predictability. Our model replicates this experimental result, suggesting that the connectivity structure learned via the proposed plasticity mechanism could plausibly underlie the songbird's learned internal model.

Results

Spontaneous replay of learnt stochastic sequences

While most studies have investigated plasticity mechanisms for learning random switching (Litwin-Kumar and Doiron, 2014 [↗](#); Triplett et al., 2018 [↗](#); Ocker and Doiron, 2019 [↗](#); Asabuki and Fukai, 2023 [↗](#)) or deterministic transitions (Chadwick et al., 2015 [↗](#)) between cell assemblies, our objective is to create a network model that spontaneously replays stochastic sequences of

assemblies following synaptic plasticity. To that end, we first design a simple task whereby stimuli undergo stochastic transitions over time, and presentation of each stimulus increases excitatory drive to neurons targeted by that pattern (**Fig. 1a** [↗](#), top). We assume that a non-overlapping subset of excitatory network neurons receive its preferred stimulus (**Fig. 1b** [↗](#)). After learning, the network should replay stochastic sequences of assemblies with transitions that are statistically consistent with evoked dynamics, without relying on external stimuli (**Fig. 1a** [↗](#), bottom).

We examined the possible learning mechanisms of stochastic neural sequences with a recurrent spiking network. Our network model consists of excitatory (E) and inhibitory (I) model neurons (**Fig. 1b** [↗](#)). Only excitatory neurons are driven by external stochastic sequences. Initially, neurons in the network have random recurrent connections.

To learn a network model to obtain transition statistics of evoked dynamics, we proposed different local plasticity mechanisms for excitatory and inhibitory synapses. We assumed that only connections onto excitatory neurons were plastic (**Fig. 1b** [↗](#)), while all others (i.e., connections onto inhibitory neurons) were fixed. In the excitatory recurrent connectivity, all synaptic weights were modified to reduce the error between internally generated and stimulus-evoked activities (**Fig. 1c** [↗](#), blue square). This plasticity rule is mathematically similar to that proposed in (Pfister et al., 2006 [↗](#); Urbanczik and Senn, 2014 [↗](#)), which minimizes the discrepancy between the somatic and dendritic activities (Asabuki and Fukai, 2020 [↗](#); Asabuki and Fukai, 2024 [↗](#)). Through this process, excitatory synapses that contribute to predicting neural activity will be strengthened, thereby increasing the similarity between spontaneous and evoked activity. Instead of predicting the firing rate of neurons, the inhibitory synapses were modified to predict the recurrent excitatory potential (**Fig. 1c** [↗](#), orange square). This inhibitory plasticity is crucial for the network to maintain excitatory-inhibitory balance (Vogels et al., 2011 [↗](#)) and generate spontaneous replay of stochastic assembly sequences, as we will see later. All feedforward connections were fixed and receptive fields were preconfigured. Finally, as in previous study (Asabuki and Fukai, 2023 [↗](#)), parameters of the response function are regulated according to the activity history of individual neurons (Methods). This regulation maintains the appropriate dynamic range of activities irrespective of the strength of external stimuli.

To examine how external stochastic sequences can influence network wiring, we trained a network model driven by stochastic external inputs. These inputs were generated by first-order Markovian chains with three 200ms long states, governed by fixed transition probabilities (**Fig. 2a** [↗](#)). During training, excitatory synapses were modified much quicker than inhibitory synapses (**Fig. 2b** [↗](#)). This difference in plasticity timescales follows from the nature of our learning rules: the wiring of excitatory synapses is reorganized by external stimuli, while inhibitory synapses only change to rebalance excitation. As such, excitatory plasticity in our model occurs before inhibitory plasticity, consistent with the experimental results (D'amour JA and Froemke, 2015 [↗](#)). Indeed, even when the learning rate of inhibitory plasticity was twice that of excitatory plasticity, inhibitory plasticity still occurred on a slower timescale than excitatory plasticity. (Supplementary Fig. 2).

We then asked how plasticity affects the neural dynamics by comparing the spontaneous activities of the network before and after learning. Here, we simulated spontaneous activity by replacing the temporally structured stimulation (i.e., the Markovian chain in **Fig. 2a** [↗](#)) with constant background input. Further, all synapses were kept fixed during spontaneous activity. Before learning, due to uniform initial connectivity, all excitatory neurons showed synchronous and spatially unstructured spontaneous activity (Supplementary Fig. 1). However, after learning, three cell assemblies emerged in the network, each of which encoded one external stimulus. Sequences of these cell assemblies were replayed stochastically in spontaneous activity (**Fig. 2c** [↗](#)), and durations of given assembly reactivations were biased toward shorter durations but distributed broadly (**Fig. 2d** [↗](#)).

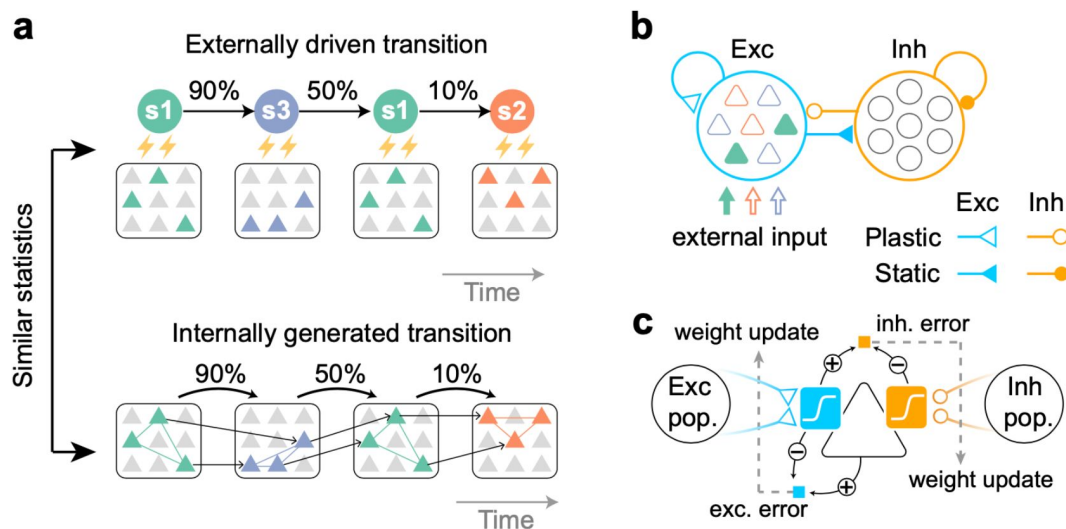


Figure 1.

Task to be learned (a, top) An example of a task used to test the model. Stimulus patterns evolve in time according to structured transition probabilities. The presentation of each stimulus pattern activates the corresponding group of neurons. Recurrent connections are learned by synaptic plasticity (a, bottom). The learned network should replay assemblies spontaneously, where the transition statistics are consistent with the evoked stimuli. (b) A network model with distinct excitatory and inhibitory populations. Only excitatory populations are driven by external inputs. Only synapses that project to excitatory neurons are assumed to be plastic. (c) A schematic of the plasticity rules proposed in (Asabuki & Clopath, 2023). Excitatory (blue) and inhibitory (orange) synapses projecting to an excitatory neuron (triangle) obey different plasticity rules. For excitatory synapses, errors between internally driven excitation (blue sigmoid) and the output of the cell provide feedback to the synapses (dashed arrow) and modulate plasticity (blue square; exc. error). All excitatory connections seek to minimize these errors. For inhibitory synapses, the error between internally driven excitation (blue sigmoid) and inhibition (orange sigmoid) must be minimized to maintain excitation-inhibition balance (orange square; inh. error).

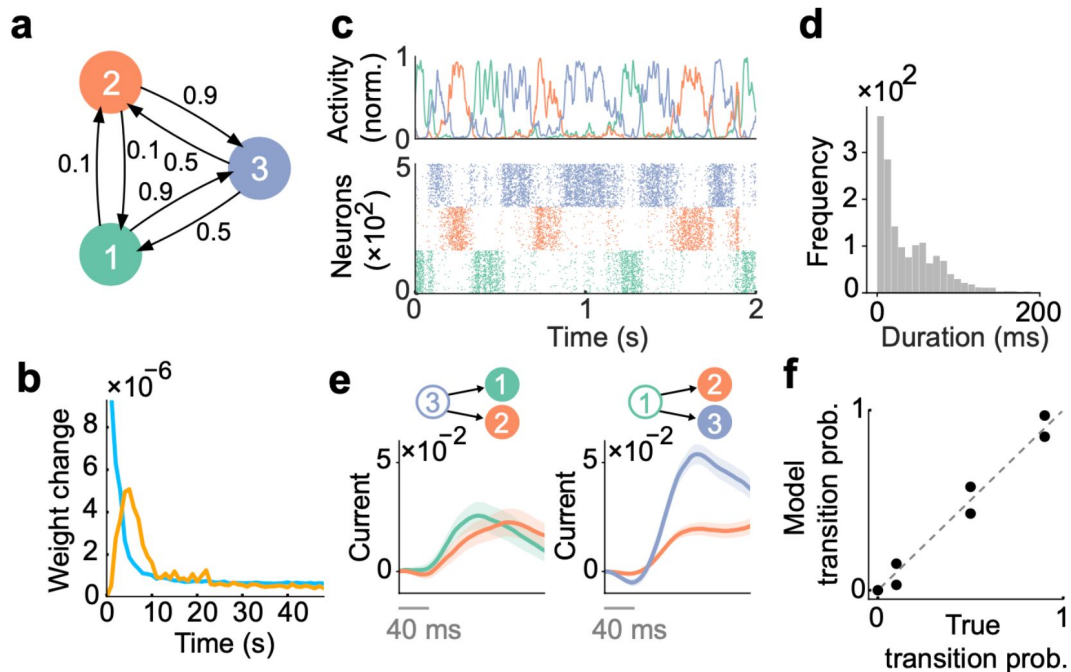


Figure 2.

Spontaneous replay of stochastic transition of assemblies. (a) First, we considered a simple stochastic transition between three stimulus patterns. (b) Dynamics of weight change via plasticity. Excitatory synapses (blue) converged quicker than inhibitory synapses (orange). (c) Example spontaneous assembly reactivations (top) and raster plot (bottom) of the learned network are shown. Colors indicate the corresponding stimulus patterns shown in a. (d) Distribution of assembly reactivations. (e, left) The network currents to assembly 1 (green) and assembly 2 (orange) immediately after the reactivation of assembly 3 ceased. Both currents were similar in magnitude. (e, right) Currents to assembly 2 (orange) and assembly 3 (blue) immediately after the reactivation of assembly 1 ceased. The current to assembly 3 was stronger than that to assembly 2. (f) Relationship between the transition statistics of stimulus patterns and that of replayed assemblies. The spontaneous activity reproduced transition statistics of external stimulus patterns.

We next asked whether or not the statistics of assembly switching were influenced by the temporal structure of the external sequence received by the network while it was learning. Since each assembly reactivation was contingent upon previous assembly, statistics of external sequence may influence the strength of synaptic currents via recurrent connectivity. To test this prediction, we first investigated how spontaneous reactivation of assembly 3 drives the subsequent assemblies (i.e., assemblies 1 and 2). Immediately after the reactivation of assembly 3 ceased, currents onto both subsequent assemblies increased gradually, without showing significant difference (**Fig.2e**, left). This is due to the fact that state 1 and 2 are structurally symmetrical in our setting (**Fig.2a**). We then asked how reactivation of assembly 1 drives the subsequent assemblies (i.e., assemblies 2 and 3). We note that the transition probabilities in the stimulus patterns were biased towards state 3 in this case (**Fig.2a**). Consistent with this bias between transition probabilities, we found that assembly 3 was driven much strongly than assembly 2 (**Fig.2e**, right). These results suggest that the temporal statistics of the trained external sequence influence the strength of synaptic currents that drive each assembly. We then quantified the similarity between the transition statistics of stimulus patterns and that of the replayed assemblies. We defined the transition probabilities between assemblies by simply counting the occurrence of switching events over all possible pairs of assemblies (Methods). Comparison between transition probabilities of stimulus patterns and that of the reactivated assemblies revealed a clear alignment of temporal statistics (**Fig.2f**). Interestingly, even when the network was trained with input states of half the duration, the distributions of the durations of assembly reactivations remain almost identical to those in the original case (Supplementary Fig. 3a). Furthermore, the transition probabilities in the replay were still consistent with the true transition probabilities (Supplementary Fig. 3b).

We then asked how robust our model is to different stimulus settings and parameters. To this end, we first asked whether varying the size of the cell assemblies would affect learning. We ran simulations with two different configurations (in the task shown in **Figure 2**). The first configuration used three assemblies with a size ratio of 1:1.5:2. After training, these assemblies exhibited transition statistics that closely matched those of the evoked activity (Supplementary Fig. 4a). In contrast, the second configuration, which used a size ratio of 1:2:3, showed worse performance compared to the 1:1.5:2 case (Supplementary Fig. 4b). These results suggest that the model can learn appropriate transition statistics as long as the size ratio of the assemblies is not drastically varied. We next asked whether the speed of plasticity, controlled by the learning rate, would affect model performance. To see this, we trained the network model in two cases, one with a fast plasticity and one with a slow plasticity. We found that the two models still showed spontaneous assembly replay whose statistics clearly matched those of the evoked dynamics (Supplementary Fig. 5b,d). Interestingly, however, we found that the duration of assembly became longer in the slow learning case than in the fast case (Supplementary Fig. 5a,c). Finally, we found that the weaker background input causes spontaneous activity with a lower replay rate, which in turn leads to a high variance of the encoded transition (Supplementary Fig. 6a,b), while stronger inputs make the assembly replay transitions more uniform (Supplementary Fig. 6c,d).

We then tested whether learning performance would be affected by setting the ratio of excitatory to inhibitory neurons to 80% and 20% (Supplementary Figure 7a; left). Even in such a scenario, the network still showed structured spontaneous activity (Supplementary Figure 7a; center), with transition statistics of replayed events matching the true transition probabilities (Supplementary Figure 7a; right). We then asked whether the model with our plasticity rule applied to all synapses would reproduce the corresponding stochastic transitions. We found the network could replay the appropriate transition only under some conditions (Supplementary Fig. 7b). The replay failed when the inhibitory neurons were no longer driven by the synaptic currents reflecting the stimulus, due to a tight balance of excitatory and inhibitory currents on the inhibitory neurons. We found similarly that when each stimulus pattern activates a non-overlapping subset of neurons, the network does not exhibit the correct stochastic transition of assembly reactivation

(Supplementary Fig. 7c). Interestingly, when the activity of each neuron is triggered by multiple stimuli and has mixed selectivity, the reactivation reproduced the appropriate stochastic transitions (Supplementary Fig. 7d).

In summary, the plasticity rules in our model learn the transition statistics of evoked patterns while maintaining excitation-inhibition balance. Our results show that this prediction-based plasticity rule allows the model to learn and spontaneously replay the transition statistics of evoked patterns.

Learned excitatory synapses encode transition statistics

To further understand the mechanism underlying the statistical similarity between the evoked patterns and spontaneous activity, we next asked how the transition statistics of stimulus patterns can influence network wiring. Over the course of training, the average weights of connections in each of the 3 cell assemblies increased gradually and converged to a strong value (**Fig.3a** [↗](#) middle and **Fig.3b** [↗](#), top), indicating the formation of assemblies. On the other hand, we found that the average weights between each pair of assemblies decreased and settled at different stationary values (**Fig.3a** [↗](#), right and **Fig.3b** [↗](#), bottom). After training, we reasoned that the transition probabilities between states should be encoded exclusively via between-assembly connections, as none of the states in the Markovian chain have self-transitions. To test this prediction, we first compared the average between-assembly connection matrix (**Fig.3a** [↗](#), right) and the ground truth transition aligned well to the ground truth probabilities (**Fig.3d** [↗](#)). These results indicate that the network learns the temporal statistics of sequences by modifying the structure of inter-assembly excitatory connections.

The above analysis of excitatory weights revealed its crucial role in learning transition probabilities. Next, we examined the role of inhibitory plasticity in our model's function. To do so, we first simulated the network with fixed inhibitory weights performing the same task shown in **Figure 2** [↗](#). We found that such a model exhibited spontaneous activity with blurred assembly structures compared to the original model (Supplementary Fig.8a). Furthermore, the transition probabilities between replayed assemblies in this case did not show clear alignment with true transition (Supplementary Fig. 8b), though the excitatory weights reached values which did encode transitions (Supplementary Fig. 8c, d). These results suggest that maintenance of EI balance through inhibitory plasticity is necessary for generating structured spontaneous activity, even if excitatory connections learn transition probabilities.

Network can adapt fast to task switching

In the above results, transitions between stimulus patterns obeyed fixed transition probabilities. We then wondered how the network learning would be affected if transition structures of stimulus patterns changed over time. To test such a scenario, we considered a case where the transition matrix in a Markovian chain switches between the first half and the second half of the learning phase (Supplementary Fig. 9a). We will refer these matrices as task1- and task2-matrix, respectively, and examine whether switching of transition matrixes influences the connectivity. During the first half of learning phase, between-assembly connections converged to certain values to encode task1-matrix (Supplementary Fig. 9b, bottom, 0-500 seconds). However, such stable connectivity reorganized quickly once the imposed task was switched to task2-matrix (Supplementary Fig. 9b, bottom, 500-1,000 seconds). Note that in contrast to between-assemblies connections, within-assembly connections did not show such reorganization (Supplementary Fig. 9b, top). These results indicate that our model adapted to the second task even if distinct assembly structures were already formed during the first task.

To further understand how the model adapts to the new task, we next asked how error terms in excitatory and inhibitory plasticity (Eqs.11 and 13) change through learning. As expected, the low-pass filtered errors PE^{exc} and PE^{inh} decreased as the network trained on task1 (Supplementary

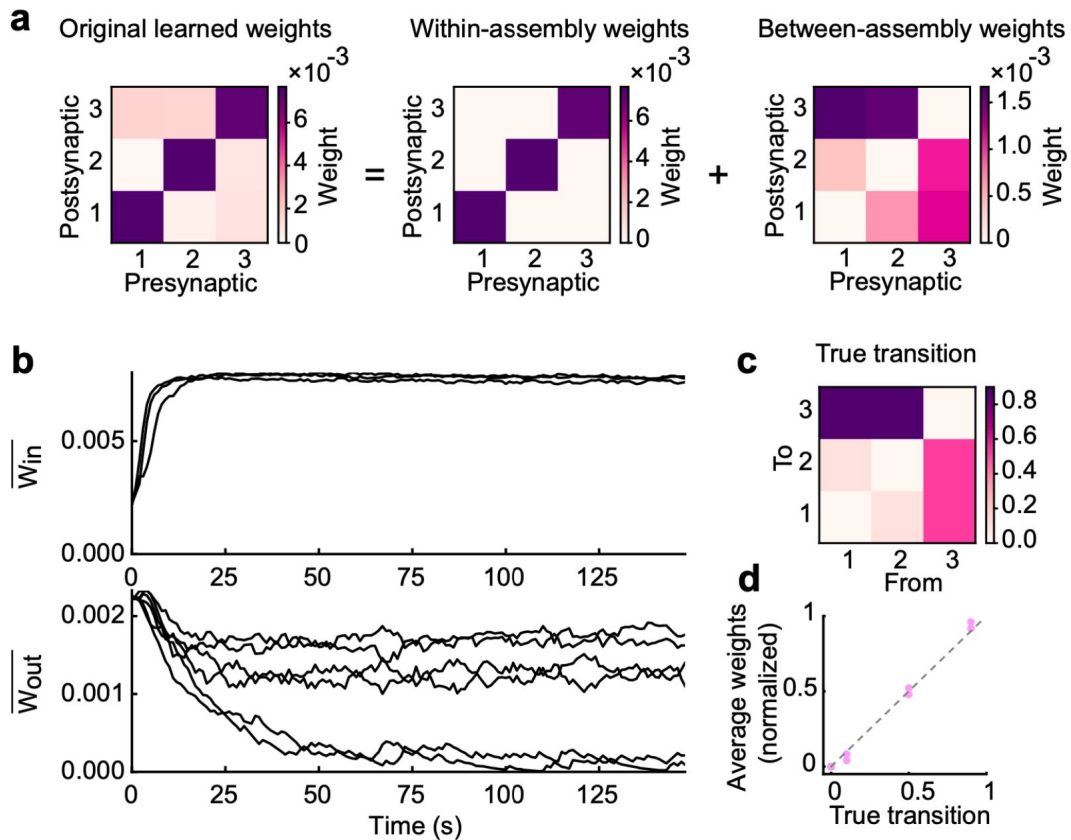


Figure 3.

Learned excitatory synapses encode transition statistics. (a) A 3 by 3 matrix of excitatory connections, learned with the task in [Fig.2a](#) (left). The matrix can be decomposed to within-(middle) and between-assembly connections (right). (b) Strength of within-(top) and that of between-assembly excitatory synapses (bottom) during learning are shown. (c) True transition matrix of stimulus patterns. (d) Relationship between the strength of excitatory synapses between assemblies and true transition probabilities between patterns.

Fig.9c, 0-500 seconds). However, once the task was switched, errors showed an abrupt increase followed by a gradual decrease as the network learned the second task (Supplementary Fig.9c, 500-1,000 seconds; Supplementary Fig.9d). Consistent with the previous result (**Figure 2b**), the peak of inhibitory error occurred delayed after that of excitatory one in each task (D'amour JA and Froemke, 2015; Vogels et al., 2011) (Supplementary Fig. 9d). In summary, our model is also capable of task switching, via the reorganization of its weight structures through continuing plasticity.

The network can learn complex stochastic sequences

So far, we have considered the capabilities of our model in regard to the relatively simple class of stochastic dynamics. In particular, the task we considered above contains only three states, and the transition structure was symmetric. In a realistic sequence, like the song of a bird, transition statistics are typically heterogeneous and more structured. To evaluate the model performance over a wide variety of structures, we now consider a transition diagram with more complex structure (**Fig.4a**). Despite its complex structure, the learned network showed spontaneous reactivations of all assemblies evoked during learning (**Fig.4b**), and the transition dynamics between these assemblies were governed by learned transition probabilities (**Fig.4c**). Indeed, the learned weight structures were consistent with the transition probabilities between states as we have seen in simpler task (Supplementary Fig.10).

Recent experimental studies which examined temporal community structure (i.e., highly structured graph structure consisting of clusters of densely interconnected nodes; **Fig.4d**) found that human subjects tend to associate a given visual stimulus with other stimuli within the same “community” (Schapiro et al., 2013; Pudhiviyath et al., 2022). To investigate whether the model can learn to associate states within a stimulated community, we first trained the network with a stochastic sequence of inputs, generated by a random walk over graph with temporal community structure (**Fig.4d**). The learned model showed stochastic assembly transition during spontaneous activity (**Fig.4e**) relying on the appropriate weight structure (**Fig.4f**). Although transition occurred between all pairs of assemblies, transitions between connected states in the diagram occurred much frequently than transitions between disconnected states (**Fig.4g**). This is because plasticity formed strong excitatory connections between assemblies with nonzero transition probabilities, as shown in **Figure 4f**.

In human participants, low-dimensional representations of evoked activities in different cortical regions have been reported to show clusters consistent with the structure of communities (Schapiro et al., 2013). To test whether our model could reproduce such representation of communities, we analyzed the low-dimensional representation of evoked activities in our model by applying principal component analysis (PCA) (see Methods). Such analysis revealed that the representation of stimulus patterns were grouped together into clusters or communities of mutually predictive stimuli, consistent with the experimental results (**Fig.4h**). We found that the clustered representations still exist even if the input sequences were scrambled after learning (Supplementary Fig.11), indicating that this result does not rely on the stimulus protocol, but instead on the learned weights.

We further asked whether within- and between-community reactivations showed any differences in terms of their behavior. To this end, we perturbed an assembly corresponding to non-boundary states in the first community (states 2-4 in the transition diagram shown in **Fig.4d**) and monitored the behavior of subsequent autonomous network activities. According to the above results, we expect that within-community reactivations should occur quicker than between-community assemblies, due to strong within-community coupling. To test this hypothesis, we calculated the duration from the end of the perturbation until subsequent activity reached a certain threshold (**Fig.4i**). As expected, the transition to within-community states showed much shorter durations than to between-community case (**Fig.4j**), indicating that between-community transition occurred with much slower time scale compared to within-community case. Together,

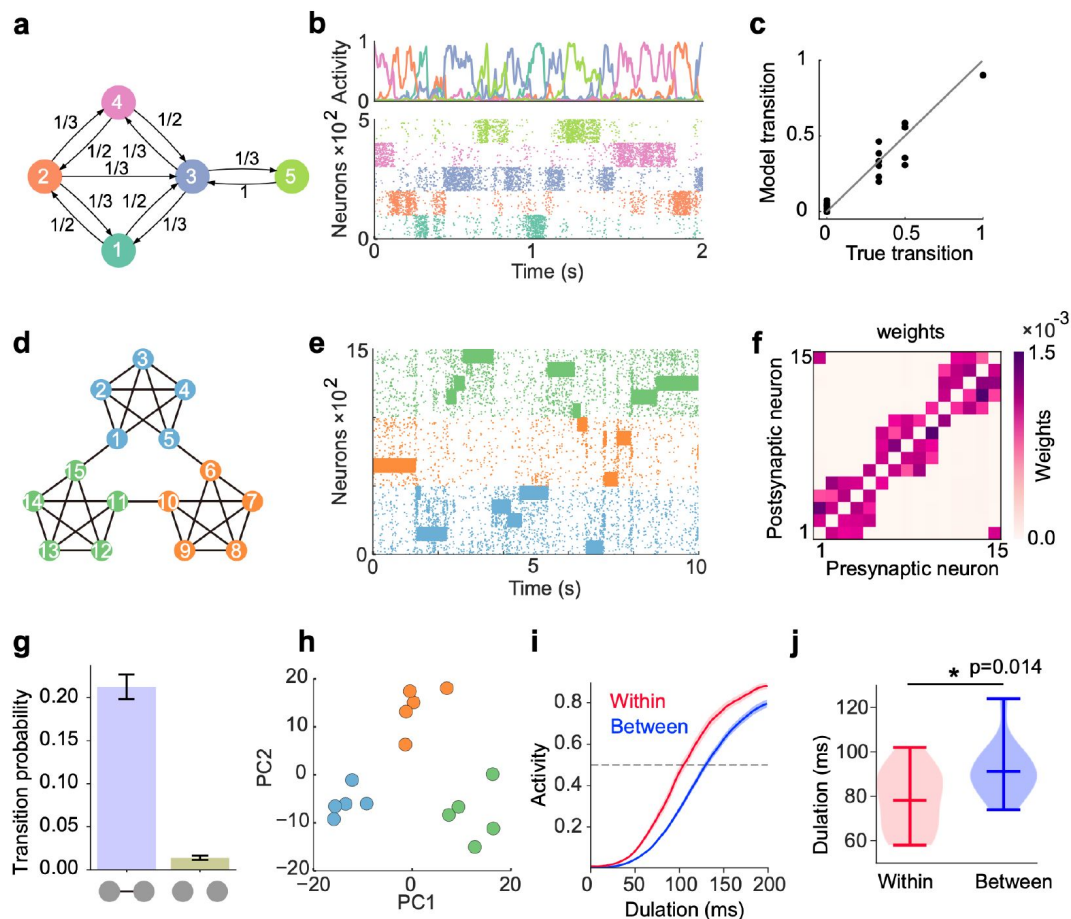


Figure 4.

Learning complex structures. (a) Transition diagram of complex task. (b) Spontaneous activity of learned network. (c) Transition statistics of assemblies reproduce true statistics. (d) Transition diagram of temporal community structure. (e) Raster plot of spontaneous activity of the network trained over structure shown in (d). (f) Structure of learned excitatory synapses encode the community structure. (g) Spontaneous transition between assemblies connected in the diagram shown in d occurs much frequent than disconnected case. (h) Low dimensional representation of evoked activity patterns shows high similarity with community structure. (i) Time courses of replayed activities transitioning within (red) and between (blue) communities. (j) Comparison of mean durations in (i). P-value was calculated by two-sided Welch's t-test.

these results indicate that our network can learn complex temporal structures in spontaneous activity and reproduce the neural representation of the temporal community structure observed in the experiment.

Network dynamics consistent with recorded neural data of songbird

Finally, we tested whether the spontaneous activity in our model resembles recorded neural activity of HVC in Bengalese finch (Bf). Bf learns songs composed of multiple stereotyped short sequences, or syllables. The transitions between these syllables can be described via Markovian process with varying levels of certainty. Intuitively, given one syllable in a bird song, precise prediction about the neural response to the next syllable can be made if the transition from that syllable is highly certain, while imprecise transitions will lead to imprecise predictions about the neural response. Indeed, recent experimental study reported that uncertainty of upcoming syllables in a Bf song modulates the degree of predictability of subsequent neural activation (poststimulus activity; PSA) in HVC (Bouchard and Brainard, 2016). We sought to test whether our model would exhibit a similar property. To this end, we analyzed the behavior of a network model that had already learned the task (shown in **Figure 4a**). The transition structure we chose is relatively simple compared to the real song of a Bf, yet captures measured features of bird songs (i.e., both structures consist of highly certain and lesscertain transitions). In the experiment, similarities were calculated between the trial-averaged PSA following a short sequence of stimuli, and the response to an isolated stimulus. To mimic this experimental design, we measured stimulus-triggered averages of our autonomous network activity as a proxy for PSA (**Fig.5a**). To examine how uncertainty of state transitions in a sequence influence predictive strength in network activity, we first calculated the Pearson correlation coefficient between PSA and responses to next states in a sequence. We will refer to such correlations as “next-state correlations”. Note that if there were multiple next-states from a given state, all correlations corresponding to that state were averaged. We further calculated the correlation between PSA and responses to other states that did not follow the given state (“other-state correlations”). Similar to the next-state correlations, other-state correlations were averaged over all disconnected states from each state. We then compared next-state correlations and other-state correlations between highly certain (**Fig.5b**, left) and less-certain (**Fig.5b**, right) transitions. Here, highly certain transitions refer to those which have a transition probability greater than 1/2. Other transitions were classified as less-certain transitions. Consistent with experimental results, next-state correlations were significantly greater than other-state correlations in the highly certain case (**Fig.5b**, left). This correlation difference was less significant in less-certain case (**Fig.5b**, right). These results indicate that transition uncertainty modulated the degree to which PSA is predictive of upcoming states.

We performed a more direct comparison of predictive strength by measuring the difference between two types of correlations (i.e., next- and other-state correlations) over multiple levels of transition uncertainty. Here, for each state, next-state correlation was subtracted by other-state correlation. Transition uncertainties were quantified by calculating the conditional entropy of transition probabilities of stimulus patterns. Note that a higher value of entropy indicates less-certain transition, and vice versa. As expected, correlation differences increased as entropy decreased (**Fig.5c**), indicating that the predictive strength of network PSA was larger for low-entropy transitions (i.e., highly certain transitions) than for high-entropy transitions (i.e., less-certain transitions). What is the underlying mechanism of such predictability differences? Although each trial of assembly perturbation lead to subsequent reactivation of one of the assemblies, trial-averaged activities (i.e., PSAs) marginalized all possible transitions in the transition diagram (**Fig.5a**). Due to this averaging process, similarities between PSA and stimulus-evoked activities increases if conditional entropy is low (i.e., certain transition), and vice versa. Overall, our results suggest that our model learns transition statistics of stimulus patterns, with transition uncertainty influencing predictive strength in the network activity.

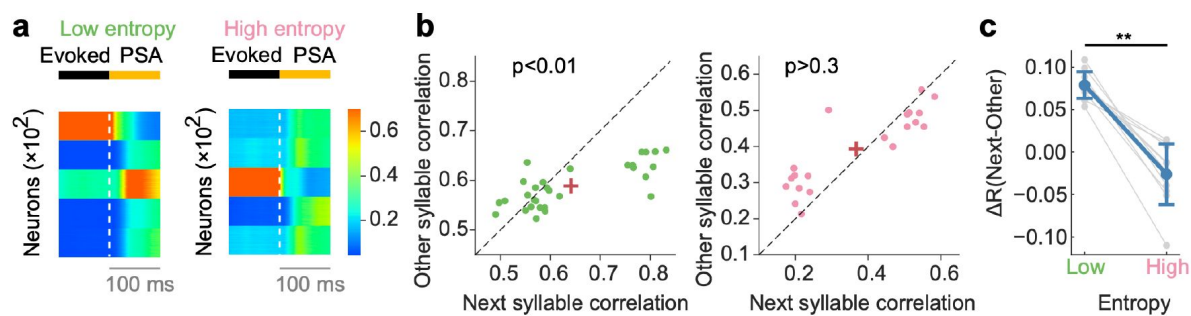


Figure 5.

Network dynamics consistent with recorded neural data of songbird (a) Example post-stimulus activity (PSA) for low-(left) and high-entropy (right) transition cases. (b) Comparison of correlation coefficients between PSA and evoked single-syllable responses for next syllables and other syllables. For low entropy transition case, the next-syllables correlations were significantly higher than other-syllables correlations ($p < 0.01$, Wilcoxon signed-rank test) (left). In contrast, such correlation coefficients showed no significant difference for high entropy transition case ($p > 0.3$, Wilcoxon signed-rank test) (right). Red crosses are mean. (c) The difference in correlation coefficients between next and other syllables (ΔR) was significantly greater for low entropy transitions than for high entropy transitions ($p < 0.01$, two-sided Welch's t-test).

Discussion

Understanding how the brain learns internal models of the environment is a challenging problem in neuroscience. In this study, we proposed synaptic plasticity rules for learning assembly transitions via sensory experiences. Our excitatory plasticity aims at minimizing the error between sensory-evoked and internally generated predictions of upcoming activity. We showed that the network learns the appropriate wiring patterns to encode the transition structure of states, and thus exhibits stochastic transitions between assemblies in spontaneous activity. We further showed that appropriate replay of stochastic transitions requires both excitatory and inhibitory plasticity. These plasticity rules showed a clear division of labor. For excitatory synapses, the connectivity learns transition probabilities during the evoked phase, and inhibitory plasticity seeks to maintain the excitatory-inhibitory balance. We showed that network excitatory plasticity alone cannot account for stochastic replay of learned activity, even if excitatory synapses learn an appropriate structure. Future experimental studies could examine our model's predictions by testing how blocking synaptic plasticity specifically in excitatory or inhibitory neuron populations distinctly impacts the transition statistics in spontaneous replay events.

Variants of the Hebbian plasticity rule have been widely used to learn the precise order of sequential reactivations. For example, a rate-based Hebbian rule has been proposed to generate trajectories along a chain of metastable attractors, each corresponding to a reactivation of a single network state (Fonollosa et al., 2015 [\[1\]](#)). Another proposed mechanism is that the transitions are governed by theta oscillations, which form a temporal backbone of the sequential reactivation of assemblies (Chadwick et al., 2015 [\[2\]](#)). Despite the successes of these Hebbian rules in learning precise order in sequences, plasticity rules that learn structured transition probabilities and replay them in spontaneous activity were still unknown.

How does our plasticity mechanism differ from the Hebbian rule? In the Hebbian rule, synaptic strength is potentiated as long as pre- and postsynaptic neurons show correlated activities. Due to this nature of the Hebbian rule, after sufficient potentiation, synapses reach a predefined upper limit, making the strength uniform among strong synapses (Kempster et al., 1999 [\[3\]](#); Song et al., 2000 [\[4\]](#); Masquelier et al., 2008 [\[5\]](#)). Such connectivity is useful when the network learns deterministic sequences, but it alone is insufficient to learn transition probabilities. In contrast, our proposed model aims at predicting the evoked activities by internally generated dynamics, so that learning ceases when the prediction error is sufficiently minimized. A similar plasticity rule has been proposed to minimize the discrepancy between stimulus-evoked and internally predicted activity, generating a stable synaptic distribution that allows pattern completion and unsupervised feature detection from noisy sensory input. (Tavazoie 2013 [\[6\]](#)). These mechanisms result in learned synaptic distributions that are not uniform as observed in STDP, but rather converge to values proportional to the transition probabilities between assemblies (as shown in **Figure 3b** [\[7\]](#)).

Our proposed plasticity mechanism could be implemented through somatodendritic interactions. Analogous to previous computational works (Urbanczik and Senn., 2014 [\[8\]](#); Asabuki and Fukai., 2020 [\[9\]](#); Asabuki et al., 2022 [\[10\]](#)), our model suggests that somatic responses may encode the stimulus-evoked neural activity states, while dendrites encode predictions based on recurrent dynamics that aim to minimize the discrepancy between somatic and dendritic activity. To directly test this hypothesis, future experimental studies could simultaneously record from both somatic and dendritic compartments to investigate how they encode evoked responses and predictive signals during learning (Francioni et al., 2022 [\[11\]](#)).

The proposed mechanism of learning stochastic transitions between cell assemblies may offer several advantages over deterministic transitions, as suggested by previous studies. One possibility is that the internal dynamics of stochastic transitions can be used as prior knowledge about the

structure of the world. In particular, the learned information about the transition statistics can be used to make probabilistic predictions about upcoming sensory events. It may also provide a flexible representation of the environment. In a deterministic case, assemblies are replayed in a fixed temporal order, which may make the network susceptible to noise or unexpected changes in the environment. In contrast, stochastic transitions may allow the network to generate rich repertoires of representations that could provide flexible computation against an uncertain environment.

In reinforcement learning (RL), balancing the tradeoff between exploration and exploitation to maximize a long-term reward signal is one of the most challenging problems. While both exploration and exploitation phases are crucial in RL, exploration is often much more difficult. This difficulty arises from the fact that exploration is especially important when the agent does not have an optimal policy. One way in which the agent might bypass or speed up this exploration phase is through prior knowledge of the environment's transition statistics. Furthermore, learning transition statistics as an internal model may be beneficial when an agent solves a task in an environment where the reward distribution is sparse. Having an internal model of the transition statistics may allow an agent to predict the expected value of the future reward for taking a particular action in a given state. However, the relationship between the reward-based plasticity rule and our proposed rule still needs further study.

Several computational models have demonstrated that Hebbian-like plasticity rule can learn appropriate Markovian statistics (Kappel et al., 2014; Barber 2002 [↗](#); Barber and Agakov 2002 [↗](#)). However, our model differs conceptually from these previous models in some respects. While Kappel et al. demonstrated that STDP in winner-take-all circuits can approximate online learning of hidden Markov models (HMMs), a key distinction from our model is that their neural representations acquire deterministic sequential activations, rather than exhibiting stochastic transitions governing Markovian dynamics. Specifically, in their model, the neural representation of state B would be different in the sequences ABC and CBA, resulting in distinct deterministic representations like ABC and C'B'A', where 'A' and 'A'' are represented by different neural states (e.g., activations of different cell assemblies). In contrast, our network learns to generate stochastically transitioning cell assemblies that replay Markovian trajectories of spontaneous activity obeying the learned transition probabilities between neural representations of states. For example, starting from reactivation from assembly 'A', there may be an 80% probability to transition to assembly 'B' and 20% to 'C'. Although Kappel et al.'s model successfully solves HMMs, their neural representations do not themselves stochastically transition between states according to the learned model. Similar to the Kappel et al.'s model, while the models proposed in Barber (2002 [↗](#)) and Barber and Agakov (2002 [↗](#)) learn the Markovian statistics, these models learned a static spatiotemporal input patterns only and how assemblies of neurons show stochastic transition in spontaneous activity has been still unclear. In contrast with these models, our model captures the probabilistic neural state trajectories, allowing spontaneous replay of experienced sequences with stochastic dynamics matching the learned environmental statistics.

Our model results were also compared to experimental results of sequence predictability in a songbird. Recent experiments have shown that the predictive uncertainty of the upcoming stimulus modulates the degree of similarity between stimulus-evoked and post-stimulus autonomous activity in the HVC of the Bengel finch (Bouchard and Brainard, 2016 [↗](#)). However, the underlying mechanism is still unknown. Here, we have shown that a stochastic state transition in spontaneous activity can explain such a dependence of activity similarity on stimulus uncertainty. Our model predicts that the PSA reflects a trial average of stochastic transitions of evoked activity from a given stimulus. Trial-averaged neural activity washes out the variability of all possible realizations of the stochastic transition. Thus, PSA of an uncertain stimulus results in a combination of multiple transitions, leading to activity less similar than that evoked by a single stimulus. Several studies have shown that Hidden Markov Models or other statistical methods

could account for the transition statistics in bird song (Kogan and Margoliash, 1998 [↗](#); Katahira et al., 2011 [↗](#)). However, our study suggests that trial averaging operations can influence the degree of similarity between stimulus-evoked and post-stimulus activity.

Although we have shown that the proposed model can learn Markovian transitions, several studies suggest that animals often exhibit behaviors with non-Markovian or hierarchical statistics (Seeds et al., 2014 [↗](#); Berman et al., 2016 [↗](#); Jovanic et al., 2016 [↗](#); Jin and Costa, 2015 [↗](#); Geddes et al., 2018 [↗](#); Markowitz et al., 2018 [↗](#); Kato et al., 2015 [↗](#); Kaplan et al., 2020 [↗](#)). In principle, our learning rule cannot be applied to learning non-Markovian transitions, since it only learns local transitions between states (Brea et al., 2013 [↗](#)). Therefore, to learn higher-order stochastic transitions, recurrent neural networks like ours may need to integrate higher-order inputs with longer time scales. Another limitation of our model is that it cannot learn transition statistics if the states are separated in time. Both of these problems could be solved by considering working memory (WM) (Baddeley, 1992 [↗](#); Miller and Cohen, 2001 [↗](#)) in an activity-dependent (Funahashi et al., 1989 [↗](#); Goldman-Rakic, 1995 [↗](#); Fuster and Alexander, 1971 [↗](#); Amit and Brunel, 1997 [↗](#)) or activity-silent manner (Mongillo et al., 2008 [↗](#); Barak and Tsodyks, 2014 [↗](#); Zucker and Regehr, 2002 [↗](#); Erickson et al., 2010 [↗](#)). Clarifying the relationship between the proposed prediction-based plasticity rule and plasticity rules that support memory traces, such as short-term plasticity, will warrant future computational studies.

Our work sheds light on the learning mechanism of the brain's internal model, which is a crucial step towards a better understanding of the role of spontaneous activity as an internal generative model of stochastic processes in complex environments.

Methods

The source codes used in this study are available at: https://github.com/TAsabuki/stochastic_transition [↗](#)

Neural network model

Our recurrent neural networks consist of N_E excitatory and N_I inhibitory neurons. During learning, the membrane potentials of neuron at time t with external current I_i^{ext} were calculated as follows:

$$u_i^E(t) = \sum_{j=1}^{N_E} W_{ij}^{EE} x_j^E(t) - \sum_{k=1}^{N_I} W_{ik}^{EI} x_k^I(t) + I_i^{\text{ext}}(t), \quad (1)$$

$$u_i^I(t) = \sum_{j=1}^{N_E} W_{ij}^{IE} x_j^E(t) - \sum_{k=1}^{N_I} W_{ik}^{II} x_k^I(t), \quad (2)$$

where u_i^E and u_i^I are the membrane potential of i -th excitatory and inhibitory neuron, respectively (see **Table 1** [↗](#) for the list of variables and functions). The strength of external input I_i^{ext} takes the value 1 if stimulus pattern targets neuron i was presented and 0 otherwise. This structured external input was replaced to constant inputs I_i^{const} of value 0.3 during spontaneous activity. We will describe the details of stimulus patterns later. W_{ij}^{ab} ($a, b = E, I$) is a recurrent connection weight from j -th neuron in population b to i -th neuron in population a . All neurons were connected with a coupling probability of $p = 0.5$. Initial value of synaptic weights W_{ij}^{ab} were uniformly set to $0.1/\sqrt{pN_b}$ if $a = E$ and $1/\sqrt{pN_b}$ if $a = I$. x_i^a is a postsynaptic potential evoked by i -th neuron in population a , which will be described later.

u_i^E, u_i^I	Membrane potentials
x_j^E, x_k^I	Postsynaptic potentials
X_i^a	Poisson spike train generated by network neurons
$W_{ij}^{EE}, W_{ik}^{EI}, W_{ij}^{IE}, W_{ik}^{II}$	Recurrent connections
I_i^{ext}	External current elicited by stimulus presentation
I_i^E, I_i^I	Synaptic currents generated by network neurons
f_i^E, f_i^I	Instantaneous firing rates
y_i^E, y_i^I	Recurrent predictions
h_i	Memory trace
φ	Dynamic sigmoidal function
$\hat{\varphi}$	Static sigmoidal function
$\text{PE}^{\text{exc}}, \text{PE}^{\text{inh}}$	Filtered prediction errors

Table 1.

Definition of variables and functions.

Spiking of each neuron model in population E was modeled as an inhomogeneous Poisson process with instantaneous firing rate f_i^E with a dynamic sigmoidal response function ϕ with parameters of slope β and threshold θ as:

$$f_i^E = \phi(u_i^E; h_i) \equiv \phi_0 [1 + \exp[g\beta(h_i)(-u_i^E + g\theta(h_i))]]^{-1}, \quad (3)$$

where ϕ_0 is the maximum instantaneous firing rate of 50 Hz and $g = 2$. The slope β and threshold θ of sigmoidal function of population E was regulated by the memory trace h_i as:

$$\beta(h_i) = h_i^{-1} \beta_0 \quad (4)$$

$$\theta(h_i) = h_i \theta_0, \quad (5)$$

where the values of constant parameters are $\beta_0 = 5$ and $\theta_0 = 1$. The memory trace tracks the maximum value of the short history of membrane potential u_i^E as

$$\begin{aligned} \dot{h}_i &= -\tau_h^{-1} h_i, & \text{if } h_i > u_i^E, \\ h_i &\leftarrow u_i^E, & \text{otherwise,} \end{aligned} \quad (6)$$

where $\tau_h = 10$ s is a time scale of memory trace. In the previous study (Asabuki and Fukai, 2023), such dynamic response function was introduced to prevent trivial solutions during the learning of recurrent and feedforward connections. In the current model, we assumed that each stimulus presentation drives a specific subset of network neurons with a fixed input strength, which avoids convergence to trivial solutions. Nevertheless, the dynamic sigmoid function could facilitate stable replay by regulating neuron activity to prevent saturation.

Inhibitory neurons' firing rate were assumed to be calculated with static sigmoidal function as:

$$f_i^I = \hat{\phi}(u_i^I) \equiv \phi_0 [1 + \exp[\beta_0(-u_i^I + \theta_0)]]^{-1}, \quad (7)$$

Where the maximum instantaneous firing rate ϕ_0 was assumed to be same with that of excitatory neurons (i.e., 50 Hz). The parameters β_0 and θ_0 are the constant values already appeared in Eqs. (4) and (5).

Neuron i in population a generates a Poisson spike train at the instantaneous firing rate of f_i^a . Let us describe the generated Poisson spike trains as:

$$X_i^a(t) = \sum_{t' \in t_i^a} \delta(t - t'), \quad (8)$$

where δ is the Dirac's delta function and t_i^a is the set of time of the spikes of the neuron. The postsynaptic potential evoked by the neuron (i.e., x_i^a) was then calculated as:

$$\tau_s \dot{x}_i^a = -x_i^a + \frac{1}{\tau} X_i^a \quad (9)$$

$$\dot{x}_i^a = -\frac{x_i^a}{\tau} + x_0 I_i^a, \quad (10)$$

where $\tau_s = 5$ ms, $\tau = 15$ ms, and $x_0 = 25$.

The learning rules

All excitatory synaptic connections onto excitatory neurons were modified to minimize the following cost function:

$$\mathcal{L}_E = \sum_{i=1}^{N_E} [f_i^E - y_i^E]^2, \quad (11)$$

where y_i^E is a recurrent prediction of a firing rate, defined as:

$$y_i^E = \hat{\phi}(\sum_{j=1}^{N_E} W_{ij}^{EE} \cdot x_j^E), \quad (12)$$

where the function $\hat{\phi}(\cdot)$ is the static sigmoid function defined in Eq.7. The above cost function evaluates to what extent the recurrent potential predicts the activity of postsynaptic neurons. Taking the gradient of the cost function in Eq.11, we derived the plasticity rule for the excitatory plasticity as :

$$\begin{aligned} \Delta W_{ij}^{EE} &\propto -\frac{\partial \mathcal{L}_E}{\partial W_{ij}^{EE}} \\ &\propto [f_i^E - y_i^E] \times \frac{\partial y_i^E}{\partial W_{ij}^{EE}} \\ &\propto y_i^E(1 - y_i^E) \cdot [f_i^E - y_i^E] \cdot x_j^E. \end{aligned} \quad (13)$$

While the term the term $y_i^E(1 - y_i^E)$, which arose from the derivative of y_i^E , avoids saturation of neural activity, we show numerically this can be ruled out in the learning rule. Hence, the resultant plasticity rule for the excitatory synapses can be written as:

$$\Delta W_{ij}^{EE} = \epsilon [f_i^E - y_i^E] \cdot x_j^E, \quad (14)$$

where ϵ is a learning rate and was set to $\epsilon = 10^{-4}$, unless otherwise specified.

We note that for Poisson spiking neurons, the derived learning rule is equivalent to the one that minimizes the Kullback-Leibler divergence between the distributions of output firing and the dendritic prediction, in our case, the recurrent prediction (Asabuki and Fukai; 2020). Thus, the rule suggests that the recurrent prediction learns the statistical model of the evoked activity, which in turn allows the network to reproduce the learned transition statistics.

Similarly, we defined the cost function for the inhibitory plasticity as:

$$\mathcal{L}_I = \sum_{i=1}^{N_E} [y_i^E - y_i^I]^2, \quad (15)$$

where y_i^I was the total inhibitory input onto postsynaptic neuron:

$$y_i^I = \hat{\phi}(\sum_{j=1}^{N_I} W_{ij}^{EI} \cdot x_j^I). \quad (16)$$

Again, by taking the gradient with respect to W_{ij}^{EI} derive the following inhibitory plasticity rule to maintain excitatory-inhibitory balance in all excitatory neurons:

$$\begin{aligned} \Delta W_{ij}^{EI} &\propto -\frac{\partial \mathcal{L}_I}{\partial W_{ij}^{EI}} \\ &\propto [y_i^E - y_i^I] \times \frac{\partial y_i^I}{\partial W_{ij}^{EI}} \\ &\propto y_i^I(1 - y_i^I) \cdot [y_i^E - y_i^I] \cdot x_j^I. \end{aligned} \quad (17)$$

In all simulations in this paper, we dropped the term $y_i^I(1 - y_i^I)$ and modified the inhibitory synapses according to the following rule:

$$\Delta W_{ij}^{EI} = \epsilon [y_i^E - y_i^I] x_j^I. \quad (18)$$

Simulation details

The parameters used in the simulations are summarized in [Table 2](#). 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.

Table 2.

Parameter settings.

p	Connection probability	0.5
g	Gain parameter in sigmoid function	2
N_E, N_I	Network size	500, 500 (1500,1500 in Figs 5e-j; 800, 200 in Supplemen- tary Fig.7)
ϵ	Learning rate	10^{-4}
τ_s	Synaptic time constant	5 ms
τ	Membrane time constant	15 ms
β_0, θ_0	Parameters for sigmoid	5, 1
τ_h	Time constant of memory trace	10 s
φ_0	Maximal firing rate	50 Hz
x_0	Scaling factor of synaptic current	25
τ_{avg}	Time constant for low-pass filtering the error	30 s
I_i^{const}	Constant external current during spontaneous activity	0.3

Stimulation protocols

In all simulations, each stimulus patterns had a duration of 200 ms and were presented without inter-pattern interval. We assumed each neuron in a network was stimulated by one of stimulus patterns and targeted assemblies were not overlapped. Presentation of each pattern triggers excitatory current to its targeted neurons of strength 1 and zero otherwise. During spontaneous activity, stimulus patterns were replaced with constant background input I_i^{const} for all excitatory neurons. In **Figure 5**, we assumed all excitatory neurons receive both structured and constant background inputs over whole period.

Calculation of transition probabilities in spontaneous activity

In **Figs. 2f**, **4c**, and **4g**, we first calculated the population average of the instantaneous firing rates of all neurons in each assembly, during spontaneous activity. We will term such activities as assembly activities. We then defined the assembly reactivations by events that the assembly activities exceeded the threshold of which the value 50% of maximum value of each assembly activities. Transition probabilities between assemblies across all possible pairs were then calculated by counting the occurrences of reactivation of the subsequent assembly within 100 ms of the end time of reactivation of the preceding assembly. In **Fig. 2d**, durations of each assembly reactivation event were defined as a period during each assembly activation exceeded threshold.

Calculation of weight changes

In **Fig. 2b**, the weight changes were calculated every 2 s for excitatory and inhibitory synapses as:

$$\Delta W^{EE}(t) := \sqrt{\sum_{i,j} [W_{ij}^{EE}(t) - W_{ij}^{EE}(t - dt)]^2} / N_E^2 \quad (19)$$

$$\Delta W^{EI}(t) := \sqrt{\sum_{i,j} [W_{ij}^{EI}(t) - W_{ij}^{EI}(t - dt)]^2} / (N_E N_I), \quad (20)$$

where $W_{ij}^{Ea}(t)$ ($a = E, I$) is a synapse at time t and dt is a simulation time step of 1 ms.

Calculation of error dynamics in task switching

In Supplementary Figures 3c and 3d, two types of prediction errors for excitatory and inhibitory plasticity were calculated as follows. First, we obtained the lowpass filtered errors $\mathcal{E}_i^{\text{exc}}$ and $\mathcal{E}_i^{\text{inh}}$ calculated by instantaneous error values in the plasticity rules (i.e., **Eqs. 14** and **18**) as:

$$\tau_{\text{avg}} \dot{\mathcal{E}}_i^{\text{exc}} = -\mathcal{E}_i^{\text{exc}} + [f_i^E(t) - y_i^E(t)] \quad (21)$$

$$\tau_{\text{avg}} \dot{\mathcal{E}}_i^{\text{inh}} = -\mathcal{E}_i^{\text{inh}} + [y_i^E(t) - y_i^I(t)], \quad (22)$$

where $\tau_{\text{avg}} = 30$ s is a time constant for low-pass filter and i is a neuron index. We then calculated the averaged errors PE^{exc} and PE^{inh} as:

$$\text{PE}^{\text{exc}}(t) = \frac{1}{N_E} \sum_{i=1}^{N_E} |\mathcal{E}_i^{\text{exc}}(t)| \quad (23)$$

$$\text{PE}^{\text{inh}}(t) = \frac{1}{N_E} \sum_{i=1}^{N_E} |\mathcal{E}_i^{\text{inh}}(t)|, \quad (24)$$

where $|\cdot|$ is an absolute value.

Analysis of the low-dimensional representation in network

In [Fig. 4h](#), we first obtained matrix of network responses $U = (r_1, \dots, r_{15})$, where r_i ($i = 1, \dots, 15$) is a trial-averaged response of a whole network to one of 15 stimulus patterns shown in [Fig. 4d](#). Trial averaging was performed over multiple presentations of each stimulus. We then applied the principal component analysis (PCA) to matrix U and visualized the low dimensional representation of multiple stimulus in the learned network.

Correlation measure for comparison with a songbird

In [Fig 5](#), we calculated stimulus-triggered averages of autonomous network activity to obtain poststimulus activity (PSA) of a network model. In [Figs 5b](#) and [5c](#), correlation between PSA and evoked activity triggered by one stimulus pattern was calculated neuron-wise and then averaged over all neurons.

Additional information

Author Contributions

T.A. and C.C. conceived the study and wrote the paper.

T.A. performed the simulations and data analyses.

Competing Interest Statement

The authors declare no competing interests.

Acknowledgements

This work was supported by BBSRC (BB/N013956/1), Wellcome Trust (200790/Z/16/Z), the Simons Foundation (564408) and EPSRC(EP/R035806/1). The authors also thank Ian Cone for his comments on the manuscript and technical assistance.

References

1. Amit DJ, Brunel N. (1997) **Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex** *Cereb Cortex* **7**:237–52 <https://doi.org/10.1093/cercor/7.3.237>
2. Asabuki T, Fukai T. (2023) **Learning rules for cortical-like spontaneous replay of an internal model** *bioRxiv* <https://doi.org/10.1101/2023.02.17.528958>
3. Asabuki T, Fukai T. (2020) **Somatodendritic consistency check for temporal feature segmentation** *Nat Commun* **11** <https://doi.org/10.1038/s41467-020-15367-w>
4. Asabuki T, Kokate P, Fukai T. (2022) **Neural circuit mechanisms of hierarchical sequence learning tested on large-scale recording data** *PLOS Computational Biology* **18**
5. Baddeley A. (1992) **Working memory** *Science* **255**:556–9 <https://doi.org/10.1126/science.1736359>
6. Barak O, Tsodyks M. (2014) **Working models of working memory** *Curr Opin Neurobiol* **25**:20–4 <https://doi.org/10.1016/j.conb.2013.10.008>
7. Barber D, Agakov F. (2002) **Correlated sequence learning in a network of spiking neurons using maximum likelihood** Institute for Adaptive and Neural Computation
8. Barber D. (2002) **Learning in spiking neural assemblies** *Advances in neural information processing systems* **15**
9. Bell C, Bodznick D, Montgomery J, Bastian J. (1997) **The generation and subtraction of sensory expectations within cerebellum-like structures** *Brain Behav Evol* **50**:17–31 <https://doi.org/10.1159/000113352>
10. Berkes P., Orban G., Lengyel M., Fiser J. (2011) **Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment** *Science* **331**:83–87
11. Berman GJ, Bialek W, Shaevitz JW (2016) **Predictability and hierarchy in Drosophila behavior** *Proceedings of the National Academy of Sciences of the United States of America* **113**:11943–11948 <https://doi.org/10.1073/pnas.1607601113>
12. Bouchard KE, Brainard MS (2016) **Auditory-induced neural dynamics in sensory-motor circuitry predict learned temporal and sequential statistics of birdsong** *Proc Natl Acad Sci U S A* **113**:9641–6 <https://doi.org/10.1073/pnas.1606725113>
13. Brea J, Senn W, Pfister JP (2013) **Matching Recall and Storage in Sequence Learning with Spiking Neural Networks** *Journal of Neuroscience* **33**:9565–9575 <https://doi.org/10.1523/JNEUROSCI.4098-12.2013>
14. Chadwick A, van Rossum MC, Nolan MF (2015) **Independent theta phase coding accounts for CA1 population sequences and enables flexible remapping** *Elife* **4** <https://doi.org/10.7554/eLife.03542>

15. D'amour JA, Froemke RC (2015) **Inhibitory and excitatory spike-timing-dependent plasticity in the auditory cortex** *Neuron* **86**:514–28 <https://doi.org/10.1016/j.neuron.2015.03.014>
16. Davidson TJ, Kloosterman F, Wilson MA (2009) **Hippocampal replay of extended experience** *Neuron* **63**:497–507 <https://doi.org/10.1016/j.neuron.2009.07.027>
17. Diba K, Buzsáki G. (2007) **Forward and reverse hippocampal place-cell sequences during ripples** *Nat Neurosci* **10**:1241–2 <https://doi.org/10.1038/nn1961>
18. Erickson MA, Maramba LA, Lisman J. (2010) **A single brief burst induces GluR1-dependent associative short-term potentiation: a potential mechanism for short-term memory** *J Cogn Neurosci* **22**:2530–40 <https://doi.org/10.1162/jocn.2009.21375>
19. Fonollosa J, Neftci E, Rabinovich M. (2015) **Learning of Chunking Sequences in Cognition and Behavior** *PLoS Comput Biol* **11** <https://doi.org/10.1371/journal.pcbi.1004592>
20. Francioni V, Tang VD, Brown NJ, Toloza EH, Harnett M. (2023) **Vectorized instructive signals in cortical dendrites during a brain-computer interface task** *bioRxiv*
21. Froemke RC, Merzenich MM, Schreiner CE (2007) **A synaptic memory trace for cortical receptive field plasticity** *Nature* **450**:425–429
22. Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) **Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex** *J Neurophysiol* **61**:331–49 <https://doi.org/10.1152/jn.1989.61.2.331>
23. Fuster JM, Alexander GE (1971) **Neuron activity related to short-term memory** *Science* **173**:652–4 <https://doi.org/10.1126/science.173.3997.652>
24. Gabriel Koch Ocker, Doiron Brent (2019) **Training and Spontaneous Reinforcement of Neuronal Assemblies by Spike Timing Plasticity** *Cerebral Cortex* **29**:937–951 <https://doi.org/10.1093/cercor/bhy001>
25. Geddes CE, Li H, Jin X. (2018) **Optogenetic Editing Reveals the Hierarchical Organization of Learned Action Sequences** *Cell* **174**:32–43 <https://doi.org/10.1016/j.cell.2018.06.012>
26. Goldman-Rakic PS (1995) **Cellular basis of working memory** *Neuron* **14**:477–85 [https://doi.org/10.1016/0896-6273\(95\)90304-6](https://doi.org/10.1016/0896-6273(95)90304-6)
27. Gupta AS, van der Meer MA, Touretzky DS, Redish AD (2010) **Hippocampal replay is not a simple function of experience** *Neuron* **65**:695–705 <https://doi.org/10.1016/j.neuron.2010.01.034>
28. Jin X, Costa RM (2015) **Shaping action sequences in basal ganglia circuits** *Current Opinion in Neurobiology* **33**:188–196 <https://doi.org/10.1016/j.conb.2015.06.011>
29. Jovanic T, Schneider-Mizell CM, Shao M, Masson JB, Denisov G, Fetter RD, et al. (2016) **Competitive Disinhibition Mediates Behavioral Choice and Sequences in Drosophila** *Cell* **167**:858–870 <https://doi.org/10.1016/j.cell.2016.09.009>
30. Kaplan HS, Salazar Thula O, Khoss N, Zimmer M. (2020) **Nested Neuronal Dynamics Orchestrate a Behavioral Hierarchy across Timescales** *Neuron* **105**:562–576 <https://doi.org/10.1016/j.neuron.2019.10.037>

31. Katahira K, Suzuki K, Okanoya K, Okada M. (2011) **Complex sequencing rules of birdsong can be explained by simple hidden Markov processes** *PLoS One* **6** <https://doi.org/10.1371/journal.pone.0024516>
32. Kato S, Kaplan HS, Schrödel T, Skora S, Lindsay TH, Yemini E, et al. (2015) **Global Brain Dynamics Embed the Motor Command Sequence of *Caenorhabditis elegans*** *Cell* **163**:656–669 <https://doi.org/10.1016/j.cell.2015.09.034>
33. Kempster R, Gerstner W, van Hemmen JL (1999) **Hebbian learning and spiking neurons** *Physical Review E* **59**:4498–4514
34. Kogan JA, Margoliash D. (1998) **Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study** *J Acoust Soc Am* **103**:2185–96 <https://doi.org/10.1121/1.421364>
35. Lee A.K., Wilson M.A. (2002) **Memory of sequential experience in the hippocampus during slow wave sleep** *Neuron* **36**:1183–1194
36. Levy N, Horn D, Meilijson I, Ruppin E. (2001) **Distributed synchrony in a cell assembly of spiking neurons** *Neural Netw* **14**:815–24 [https://doi.org/10.1016/s0893-6080\(01\)00044-2](https://doi.org/10.1016/s0893-6080(01)00044-2)
37. Lewald J, Ehrenstein WH (1998) **Influence of head-to-trunk position on sound lateralization** *Exp Brain Res* **121**:230–8 <https://doi.org/10.1007/s002210050456>
38. Litwin-Kumar A., Doiron B. (2014) **Formation and maintenance of neuronal assemblies through synaptic plasticity** *Nat Commun* **5** <https://doi.org/10.1038/ncomms6319>
39. Markowitz JE, Gillis WF, Beron CC, Neufeld SQ, Robertson K, Bhagat ND, et al. (2018) **The Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection** *Cell* **174**:44–58 <https://doi.org/10.1016/j.cell.2018.04.019>
40. Masquelier T, Guyonneau R, Thorpe SJ (2008) **Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains** *PLoS One* **3**
41. Merfeld DM, Zupan L, Peterka RJ (1999) **Humans use internal models to estimate gravity and linear acceleration** *Nature* **398**:615–8 <https://doi.org/10.1038/19303>
42. Miller EK, Cohen JD (2001) **An integrative theory of prefrontal cortex function** *Annu Rev Neurosci* **24**:167–202 <https://doi.org/10.1146/annurev.neuro.24.1.167>
43. Mongillo G, Barak O, Tsodyks M. (2008) **Synaptic theory of working memory** *Science* **319**:1543–6 <https://doi.org/10.1126/science.1150769>
44. Pfister JP, Toyoizumi T, Barber D, Gerstner W. (2006) **Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning** *Neural Comput* **18**:1318–48 <https://doi.org/10.1162/neco.2006.18.6.1318>
45. Pudhiyidath A, Morton NW, Viveros Duran R, Schapiro AC, Momennejad I, Hinojosa-Rowland DM, Molitor RJ, Preston AR (2022) **Representations of Temporal Community Structure in Hippocampus and Precuneus Predict Inductive Reasoning Decisions** *J Cogn Neurosci* **34**:1736–1760 https://doi.org/10.1162/jocn_a_01864

46. Schapiro AC, Rogers TT, Cordova NI, Turk-Browne NB, Botvinick MM (2013) **Neural representations of events arise from temporal community structure** *Nat Neurosci* **16**:486–92 <https://doi.org/10.1038/nn.3331>
47. Seeds AM, Ravbar P, Chung P, Hampel S, Midgley FM, Mensh BD, et al. (2014) **A suppression hierarchy among competing motor programs drives sequential grooming in *Drosophila*** *eLife* **3** <https://doi.org/10.7554/eLife.02951>
48. Skaggs W.E., McNaughton B.L. (1996) **Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience** *Science* **271**:1870–1873
49. Song S, Miller KD, Abbott LF (2000) **Competitive Hebbian learning through spike-timing dependent synaptic plasticity** *Nature Neuroscience* **3**:919–926
50. Tavazoie S. (2013) **Synaptic state matching: a dynamical architecture for predictive internal representation and feature detection** *Plos one* **8**
51. Triplett MA, Avitan L, Goodhill GJ (2018) **Emergence of spontaneous assembly activity in developing neural networks without afferent input** *PLOS Computational Biology* **14** <https://doi.org/10.1371/journal.pcbi.1006421>
52. Urbanczik R, Senn W. (2014) **Learning by the dendritic prediction of somatic spiking** *Neuron* **81**:521–8 <https://doi.org/10.1016/j.neuron.2013.11.030>
53. Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. (2011) **Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks** *Science* **334**:1569–73 <https://doi.org/10.1126/science.1211095>
54. Wilson M.A., McNaughton B.L. (1994) **Reactivation of hippocampal ensemble memories during sleep** *Science* **265**:676–679
55. Wolpert DM, Ghahramani Z, Jordan MI (1995) **An internal model for sensorimotor integration** *Science* **269**:1880–2 <https://doi.org/10.1126/science.7569931>
56. Wu X, Foster DJ (2014) **Hippocampal replay captures the unique topological structure of a novel environment** *J Neurosci* **34**:6459–69 <https://doi.org/10.1523/JNEUROSCI.3414-13.2014>
57. Yasui S, Young LR (1975) **Perceived visual motion as effective stimulus to pursuit eye movement system** *Science* **190**:906–8 <https://doi.org/10.1126/science.1188373>
58. Zucker RS, Regehr WG (2002) **Short-term synaptic plasticity** *Annu Rev Physiol* **64**:355–405 <https://doi.org/10.1146/annurev.physiol.64.092501.114547>

Author information

Toshitake Asabuki

Department of Bioengineering, Imperial College London, London, UK, RIKEN Center for Brain Science, RIKEN ECL Research Unit, Wako, Japan, RIKEN Cluster for Pioneering Research, Wako, Japan

ORCID iD: [0000-0003-0951-5791](https://orcid.org/0000-0003-0951-5791)

For correspondence: toshitake.asabuki@riken.jp

Claudia Clopath

Department of Bioengineering, Imperial College London, London, UK
ORCID iD: [0000-0003-4507-8648](https://orcid.org/0000-0003-4507-8648)

For correspondence: c.clopath@imperial.ac.uk

Editors

Reviewing Editor

Julijana Gjorgjieva

Technical University of Munich, Freising, Germany

Senior Editor

Michael Frank

Brown University, Providence, United States of America

Reviewer #2 (Public review):

Summary:

This work proposes a synaptic plasticity rule which explains the generation of learned stochastic dynamics during spontaneous activity. The proposed plasticity rule assumes that excitatory synapses seek to minimize the difference between the internal predicted activity and stimulus-evoked activity, and inhibitory synapses tries to maintain the E-I balance by matching the excitatory activity. By implementing this plasticity rule in a spiking recurrent neural network, the authors show that the state-transition statistics of spontaneous excitatory activity agrees with that of the learned stimulus patterns, which is reflected in the learned excitatory synaptic weights. The authors further demonstrate that inhibitory connections contribute to well-defined state-transitions matching the transition patterns evoked by the stimulus. Finally, they show that this mechanism can be expanded to more complex state-transition structures including songbird neural data.

Strengths:

This study makes an important contribution to computational neuroscience, by proposing a possible synaptic plasticity mechanism underlying spontaneous generations of learned stochastic state-switching dynamics that are experimentally observed in the visual cortex and hippocampus. This work is also very clearly presented and well-written, and the authors conducted comprehensive simulations testing multiple hypotheses. Overall, I believe this is a well-conducted study providing interesting and novel aspects on the capacity of recurrent spiking neural networks with local synaptic plasticity.

Weaknesses:

This study is very well-thought out and theoretically valuable to the neuroscience community, and I think the main weaknesses are in regard to how much biological realism is taken into account. For example, the proposed model assumes that only synapses targeting excitatory neurons are plastic, and uses an equal number of excitatory and inhibitory neurons. The model also assumes Markovian state dynamics while biological systems can depend more on history. This limitation, however, is acknowledged in the Discussion. Finally, to simulate spontaneous activity, the authors use a constant input of 0.3 throughout the study. Different amplitudes of constant input may correspond to different internal states,

so it will be more convincing if the authors test the model with varying amplitudes of constant inputs.

Comments on revisions:

The authors have addressed all of the previously raised concerns satisfactorily, by running extra simulations with a biologically plausible composition of excitatory and inhibitory neurons, plasticity assumed for all synapses, and varied amounts of constant inputs representing internal states or background activities. While in some of these cases the stochastic dynamics during spontaneous activity change or do not replicate those of the learned stimulus patterns as well as before, these extended studies provide thorough evaluations of the strengths and limitations of the proposed plasticity rule as the underlying mechanism of stochastic dynamics during spontaneous activity. Overall, the revision has strengthened the paper significantly.

<https://doi.org/10.7554/eLife.95243.2.sa2>

Reviewer #3 (Public review):

Summary:

Asabuki and Clopath study stochastic sequence learning in recurrent networks of Poisson spiking neurons that obey Dale's law. Inspired by previous modeling studies, they introduce two distinct learning rules, to adapt excitatory-to-excitatory and inhibitory-to-excitatory synaptic connections. Through a series of computer experiments, the authors demonstrate that their networks can learn to generate stochastic sequential patterns, where states correspond to non-overlapping sets of neurons (cell assemblies) and the state-transition conditional probabilities are first-order Markov, i.e., the transition to a given next state only depends on the current state. Finally, the authors use their model to reproduce certain experimental songbird data involving highly-predictable and highly-uncertain transitions between song syllables. While the findings are only moderately surprising, this is a well-written and welcome detailed study that may be of interest to experts of plasticity and learning in recurrent neural networks that respect Dale's law.

Strengths:

This is an easy-to-follow, well-written paper, whose results are likely easy to reproduce. The experiments are clear and well-explained. In particular, the study of the interplay between excitation and inhibition (and their different plasticity rules) is a highlight of the study. The study of songbird experimental data is another good feature of this paper; finches are classical model animals for understanding sequence learning in the brain. I also liked the study of rapid task-switching, it's a good-to-know type of result that is not very common in sequence learning papers.

Weaknesses:

One weakness I see in this paper is the derivation of the learning rules, which is semi-heuristic. The paper studies Poisson spiking neurons, for which learning rules can be derived from a statistical objective, typically maximum likelihood, as previously done in the cited literature. The authors provide a brief section connecting the learning rules to gradient descent on objective functions, but the link is only heuristic or at least not entirely presented. The reason is that the neural network state is not fully determined by (or "clamped to") the target during learning (for instance, inhibitory neurons do not even have a target assigned). So, the (total) gradient should take into account the recurrent contributions from other neurons, and equation 13 does not appear to be complete/correct to me. Moreover, the target firing rate is a mixture of external currents with currents arising from other neurons in the

recurrent network. The authors ideally should start from an actual distribution matching objective (e.g., KL divergence, and not such a squared error), so that their main claims immediately follow from the mathematical derivations. Along the same line, it would be excellent to get some additional insights on the interaction of the two distinct plasticity rules, one of the highlights of the study. This could be naturally achieved by relating their distinct rules to a common principled objective.

The other major weakness (albeit one that is clearly discussed by the authors) is that the study assumes that every excitatory neuron is directly given its target state when learning. In machine learning language, there are no 'hidden' excitatory neurons. While this assumption greatly simplifies the derivation of efficient and biologically-plausible learning rules that can be mapped to synaptic plasticity, it also limits considerably the distributions that can be learned by the network, more precisely to those that satisfy the Markov property.

<https://doi.org/10.7554/eLife.95243.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public Review):

In the presented manuscript, the authors investigate how neural networks can learn to replay presented sequences of activity. Their focus lies on the stochastic replay according to learned transition probabilities. They show that based on error-based excitatory and balance-based inhibitory plasticity networks can selforganize towards this goal. Finally, they demonstrate that these learning rules can recover experimental observations from song-bird song learning experiments.

Overall, the study appears well-executed and coherent, and the presentation is very clear and helpful. However, it remains somewhat vague regarding the novelty. The authors could elaborate on the experimental and theoretical impact of the study, and also discuss how their results relate to those of Kappel et al, and others (e.g., Kappel et al (doi.org/10.1371/journal.pcbi.1003511)).

We agree with the reviewer that our previous manuscript lacked comparison with previously published similar works. While Kappel et al. demonstrated that STDP in winner-take-all circuits can approximate online learning of hidden Markov models (HMMs), a key distinction from our model is that their neural representations acquire deterministic sequential activations, rather than exhibiting stochastic transitions governing Markovian dynamics. Specifically, in their model, the neural representation of state B would be different in the sequences ABC and CBA, resulting in distinct deterministic representations like ABC and C'B'A', where 'A' and 'A'' are represented by different neural states (e.g., activations of different cell assemblies). In contrast, our network learns to generate stochastically transitioning cell assemblies which replay Markovian trajectories of spontaneous activity obeying the learned transition probabilities between neural representations of states. For example, starting from reactivation from assembly 'A', there may be an 80% probability to transition to assembly 'B' and 20% to 'C'. Although Kappel et al.'s model successfully solves HMMs, their neural representations do not themselves stochastically transition between states according to the learned model. Similar to the Kappel et al.'s model, while the models proposed in Barber (2002) and Barber and Agakov (2002) learn the Markovian statistics, these models learned a static spatiotemporal input patterns only and how assemblies of neurons show stochastic transition in spontaneous activity has been still unclear. In contrast with these models, our

model captures the probabilistic neural state trajectories, allowing spontaneous replay of experienced sequences with stochastic dynamics matching the learned environmental statistics.

We have included new sentences for explain these in ll. 509-533 in the revised manuscript.

Overall, the work could benefit if there was either (A) a formal analysis or derivation of the plasticity rules involved and a formal justification of the usefulness of the resulting (learned) neural dynamics;

We have included a derivation of our plasticity rules in ll. 630-670 in the revised manuscript. Consistent with our claim that excitatory plasticity updates the excitatory synapse to predict output firing rates, we have shown that the corresponding cost function measures the discrepancy between the recurrent prediction and the output firing rate. Similarly, for inhibitory plasticity, we defined the cost function that evaluates the difference between the excitatory and inhibitory potential within each neuron. We showed that the resulting inhibitory plasticity rule updates the inhibitory synapses to maintain the excitation-inhibition balance.

and/or (B) a clear connection of the employed plasticity rules to biological plasticity and clear testable experimental predictions. Thus, overall, this is a good work with some room for improvement.

Our proposed plasticity mechanism could be implemented through somatodendritic interactions. Analogous to previous computational works (Urbanczik and Senn., 2014; Asabuki and Fukai., 2020; Asabuki et al., 2022), our model suggests that somatic responses may encode the stimulus-evoked neural activity states, while dendrites encode predictions based on recurrent dynamics that aim to minimize the discrepancy between somatic and dendritic activity. To directly test this hypothesis, future experimental studies could simultaneously record from both somatic and dendritic compartments to investigate how they encode evoked responses and predictive signals during learning (Francioni et al., 2022).

We have included new sentences for explain these in ll. 476-484 in the revised manuscript.

Reviewer #2 (Public Review):

Summary:

This work proposes a synaptic plasticity rule that explains the generation of learned stochastic dynamics during spontaneous activity. The proposed plasticity rule assumes that excitatory synapses seek to minimize the difference between the internal predicted activity and stimulus-evoked activity, and inhibitory synapses try to maintain the E-I balance by matching the excitatory activity. By implementing this plasticity rule in a spiking recurrent neural network, the authors show that the state-transition statistics of spontaneous excitatory activity agree with that of the learned stimulus patterns, which are reflected in the learned excitatory synaptic weights. The authors further demonstrate that inhibitory connections contribute to well-defined state transitions matching the transition patterns evoked by the stimulus. Finally, they show that this mechanism can be expanded to more complex state-transition structures including songbird neural data.

Strengths:

This study makes an important contribution to computational neuroscience, by proposing a possible synaptic plasticity mechanism underlying spontaneous generations of learned stochastic state-switching dynamics that are experimentally observed in the visual cortex and hippocampus. This work is also very clearly presented and well-written,

and the authors conducted comprehensive simulations testing multiple hypotheses. Overall, I believe this is a well-conducted study providing interesting and novel aspects of the capacity of recurrent spiking neural networks with local synaptic plasticity.

Weaknesses:

This study is very well-thought-out and theoretically valuable to the neuroscience community, and I think the main weaknesses are in regard to how much biological realism is taken into account. For example, the proposed model assumes that only synapses targeting excitatory neurons are plastic, and uses an equal number of excitatory and inhibitory neurons.

We agree with the reviewer. The network shown in the previous manuscript consists of an equal number of excitatory and inhibitory neurons, which seems to lack biological plausibility. Therefore, we first tested whether a biologically plausible scenario would affect learning performance by setting the ratio of excitatory to inhibitory neurons to 80% and 20% (Supplementary Figure 7a; left). Even in such a scenario, the network still showed structured spontaneous activity (Supplementary Figure 7a; center), with transition statistics of replayed events matching the true transition probabilities (Supplementary Figure 7a; right). We then asked whether the model with our plasticity rule applied to all synapses would reproduce the corresponding stochastic transitions. We found that the network can learn transition statistics but only under certain conditions. The network showed only weak replay and failed to reproduce the appropriate transition (Supplementary Fig. 7b) if the inhibitory neurons were no longer driven by the synaptic currents reflecting the stimulus, due to a tight balance of excitatory and inhibitory currents on the inhibitory neurons. We then tested whether the network with all synapses plastic can learn transition statistics if the external inputs project to the inhibitory neurons as well. We found that, when each stimulus pattern activates a non-overlapping subset of neurons, the network does not exhibit the correct stochastic transition of assembly reactivation (Supplementary Fig. 7c). Interestingly, when each neuron's activity is triggered by multiple stimuli and has mixed selectivity, the reactivation reproduced the appropriate stochastic transitions (Supplementary Fig. 7d).

We have included these new results as new Supplementary Figure 7 and they are explained in ll.215-230 in the revised manuscript.

The model also assumes Markovian state dynamics while biological systems can depend more on history. This limitation, however, is acknowledged in the Discussion.

We have included the following sentence to provide a possible solution to this limitation: "Therefore, to learn higher-order stochastic transitions, recurrent neural networks like ours may need to integrate higher-order inputs with longer time scales." in ll.557-559 in the revised manuscript.

Finally, to simulate spontaneous activity, the authors use a constant input of 0.3 throughout the study. Different amplitudes of constant input may correspond to different internal states, so it will be more convincing if the authors test the model with varying amplitudes of constant inputs.

We thank the reviewer for pointing this out. In the revised manuscript, we have tested constant input with three different strengths. If the strength is moderate, the network showed accurate encoding of transition statistics in the spontaneous activity as we have seen in Fig.2. We have additionally shown that the weaker background input causes spontaneous activity with lower replay rate, which in turn leads to high variance of encoded transition, while stronger inputs make assembly replay transitions more uniform. We have included these

new results as new Supplementary Figure 6 and they are explained in ll.211214 in the revised manuscript.

Reviewer #3 (Public Review):

Summary:

Asabuki and Clopath study stochastic sequence learning in recurrent networks of Poisson spiking neurons that obey Dale's law. Inspired by previous modeling studies, they introduce two distinct learning rules, to adapt excitatory-to-excitatory and inhibitory-to-excitatory synaptic connections. Through a series of computer experiments, the authors demonstrate that their networks can learn to generate stochastic sequential patterns, where states correspond to non-overlapping sets of neurons (cell assemblies) and the state-transition conditional probabilities are first-order Markov, i.e., the transition to a given next state only depends on the current state. Finally, the authors use their model to reproduce certain experimental songbird data involving highly-predictable and highly-uncertain transitions between song syllables.

Strengths:

This is an easy-to-follow, well-written paper, whose results are likely easy to reproduce. The experiments are clear and well-explained. The study of songbird experimental data is a good feature of this paper; finches are classical model animals for understanding sequence learning in the brain. I also liked the study of rapid task-switching, it's a good-to-know type of result that is not very common in sequence learning papers.

Weaknesses:

While the general subject of this paper is very interesting, I missed a clear main result. The paper focuses on a simple family of sequence learning problems that are well-understood, namely first-order Markov sequences and fully visible (nohidden-neuron) networks, studied extensively in prior work, including with spiking neurons. Thus, because the main results can be roughly summarized as examples of success, it is not entirely clear what the main point of the authors is.

We apologize the reviewer that our main claim was not clear. While various computational studies have suggested possible plasticity mechanisms for embedding evoked activity patterns or their probability structures into spontaneous activity (Litwin-Kumar et al., Nat. Commun. 2014, Asabuki and Fukai., Biorxiv 2023), how transition statistics of the environment are learned in spontaneous activity is still elusive and poorly understood. Furthermore, while several network models have been proposed to learn Markovian dynamics via synaptic plasticity (Brea, et al. (2013); Pfister et al. (2004); Kappel et al. (2014)), they have been limited in a sense that the learned network does not show stochastic transition in a neural state space. For instance, while Kappel et al. demonstrated that STDP in winner-take-all circuits can approximate online learning of hidden Markov models (HMMs), a key distinction from our model is that their neural representations acquire deterministic sequential activations, rather than exhibiting stochastic transitions governing Markovian dynamics. Specifically, in their model, the neural representation of state B would be different in the sequences ABC and CBA, resulting in distinct deterministic representations like ABC and C'B'A', where 'A' and 'A' are represented by different neural states (e.g., activations of different cell assemblies). In contrast, our network learns to generate stochastically transitioning cell assemblies that replay Markovian trajectories of spontaneous activity obeying the learned transition probabilities between neural representations of states. For example, starting from reactivation from assembly 'A', there may be an 80% probability to transition to assembly 'B' and 20% to 'C'. Although Kappel et al.'s model successfully solves HMMs, their neural representations do not themselves stochastically transition between

states according to the learned model. Similar to the Kappel et al.'s model, while the models proposed in Barber (2002) and Barber and Agakov (2002) learn the Markovian statistics, these models learned a static spatiotemporal input patterns only and how assemblies of neurons show stochastic transition in spontaneous activity has been still unclear. In contrast with these models, our model captures the probabilistic neural state trajectories, allowing spontaneous replay of experienced sequences with stochastic dynamics matching the learned environmental statistics.

We have explained this point in ll.509-533 in the revised manuscript.

Going into more detail, the first major weakness I see in this paper is the heuristic choice of learning rules. The paper studies Poisson spiking neurons (I return to this point below), for which learning rules can be derived from a statistical objective, typically maximum likelihood. For fully-visible networks, these rules take a simple form, similar in many ways to the E-to-E rule introduced by the authors. This more principled route provides quite a lot of additional understanding on what is to be expected from the learning process.

We thank the reviewer for pointing this out. To better demonstrate the function of our plasticity rules, we have included the derivation of the rules of synaptic plasticity in ll. 630-670 in the revised manuscript. Consistent with our claim that excitatory plasticity updates the excitatory synapse to predict output firing rates, we have shown that the corresponding cost function measures the discrepancy between the recurrent prediction and the output firing rate. Similarly, for inhibitory plasticity, we defined the cost function that evaluates the difference between the excitatory and inhibitory potential within each neuron. We showed that the resulting inhibitory plasticity rule updates the inhibitory synapses to maintain the excitation-inhibition balance.

For instance, should maximum likelihood learning succeed, it is not surprising that the statistics of the training sequence distribution are reproduced. Moreover, given that the networks are fully visible, I think that the maximum likelihood objective is a convex function of the weights, which then gives hope that the learning rule does succeed. And so on. This sort of learning rule has been studied in a series of papers by David Barber and colleagues [refs. 1, 2 below], who applied them to essentially the same problem of reproducing sequence statistics in recurrent fully-visible nets. It seems to me that one key difference is that the authors consider separate E and I populations, and find the need to introduce a balancing I-to-E learning rule.

The reviewer's understanding that inhibitory plasticity to maintain EI balance is one of a critical difference from previous works is correct. However, we believe that the most striking point of our study is that we have shown numerically that predictive plasticity rules enable recurrent networks to learn and replay the assembly activations whose transition statistics match those of the evoked activity. Please see our reply above.

Because the rules here are heuristic, a number of questions come to mind. Why these rules and not others - especially, as the authors do not discuss in detail how they could be implemented through biophysical mechanisms? When does learning succeed or fail? What is the main point being conveyed, and what is the contribution on top of the work of e.g. Barber, Brea, et al. (2013), or Pfister et al. (2004)?

Our proposed plasticity mechanism could be implemented through somatodendritic interactions. Analogous to previous computational works (Senn, Asabuki), our model suggests that somatic responses may encode the stimulus-evoked neural activity states, while dendrites encode predictions based on recurrent dynamics that aim to minimize the discrepancy between somatic and dendritic activity. To directly test this hypothesis, future experimental

studies could simultaneously record from both somatic and dendritic compartments to investigate how they encode evoked responses and predictive signals during learning.

To address the point of the reviewer, we conducted additional simulations to test where the model fails. We found that the model with our plasticity rule applied to all synapses only showed faint replays and failed to replay the appropriate transition (Supplementary Fig. 7b). This result is reasonable because the inhibitory neurons were no longer driven by the synaptic currents reflecting the stimulus, due to a tight balance of excitatory and inhibitory currents on the inhibitory neurons. Our model predicts that mixed selectivity in the inhibitory population is crucial to learn an appropriate transition statistics (Supplementary Fig. 7d). Future work should clarify the role of synaptic plasticity on inhibitory neurons, especially plasticity at I to I synapses. We have explained this result as new supplementary Figure 7 in the revised manuscript.

The use of a Poisson spiking neuron model is the second major weakness of the study. A chief challenge in much of the cited work is to generate stochastic transitions from recurrent networks of deterministic neurons. The task the authors set out to do is much easier with stochastic neurons; it is reasonable that the network succeeds in reproducing Markovian sequences, given an appropriate learning rule. I believe that the main point comes from mapping abstract Markov states to assemblies of neurons. If I am right, I missed more analyses on this point, for instance on the impact that varying cell assembly size would have on the findings reported by the authors.

The reviewer's understanding is correct. Our main point comes from mapping Markov statistics to replays of cell assemblies. In the revised manuscript, we performed additional simulations to ask whether varying the size of the cell assemblies would affect learning. We ran simulations with two different configurations in the task shown in Figure 2. The first configuration used three assemblies with a size ratio of 1:1.5:2. After training, these assemblies exhibited transition statistics that closely matched those of the evoked activity (Supplementary Fig. 4a,b). In contrast, the second configuration, which used a size ratio of 1:2:3, showed worse performance compared to the 1:1.5:2 case (Supplementary Fig. 4c,d). These results suggest that the model can learn appropriate transition statistics as long as the size ratio of the assemblies is not drastically varied.

Finally, it was not entirely clear to me what the main fundamental point in the HVC data section was. Can the findings be roughly explained as follows: if we map syllables to cell assemblies, for high-uncertainty syllable-to-syllable transitions, it becomes harder to predict future neural activity? In other words, is the main point that the HVC encodes syllables by cell assemblies?

The reviewer's understanding is correct. We wanted to show that if the HVC learns transition statistics as a replay of cell assemblies, a high-uncertainty syllable-to-syllable transition would make predicting future reactivations more difficult, since trial-averaged activities (i.e., poststimulus activities; PSAs) marginalized all possible transitions in the transition diagram.

(1) Learning in Spiking Neural Assemblies, David Barber, 2002. URL: <https://proceedings.neurips.cc/paper/2002/file/619205da514e83f869515c782a328d3c-Paper.pdf>

(2) Correlated sequence learning in a network of spiking neurons using maximum likelihood, David Barber, Felix Agakov, 2002. URL: <http://web4.cs.ucl.ac.uk/staff/D.Barber/publications/barber-agakovTR0149.pdf>

Recommendations for the authors:

Reviewer #1 (Recommendations For The Authors):

In more detail:

A) Theoretical analysis

The plasticity rules in the study are introduced with a vague reference to previous theoretical studies of others. Doing this, one does not provide any formal insight as to why these plasticity rules should enable one to learn to solve the intended task, and whether they are optimal in some respect. This becomes noticeable, especially in the discussion of the importance of inhibitory balance, which does not go into any detail, but rather only states that its required, both in the results and discussion sections. Another unclarity appears when error-based learning is discussed and compared to Hebbian plasticity, which, as you state, "alone is insufficient to learn transition probabilities". It is not evident how this claim is warranted, nor why error-based plasticity in comparison should be able to perform this (other than referring to the simulation results). Please either clarify formally (or at least intuitively) how plasticity rules result in the mentioned behavior, or alternatively acknowledge explicitly the (current) lack of intuition.

The lack of formal discussion is a relevant shortcoming compared to previous research that showed very similar results with formally more rigorous and principled approaches. In particular, Kappel et al derived explicitly how neural networks can learn to sample from HMMs using STDP and winner-take-all dynamics. Even though this study has limitations, the relation with respect to that work should be made very clear; potentially the claims of novelty of some results (sampling) should be adjusted accordingly. See also Yanping Huang, Rajesh PN Rao (NIPS 2014), and possibly other publications. While it might be difficult to formally justify the learning rules post-hoc, it would be very helpful to the field if you very clearly related your work to that of others, where learning rules have been formally justified, and elaborate on the intuition of how the employed rules operate and interact (especially for inhibition).

Lastly, while the importance of sampling learned transition probabilities is discussed, the discussion again remains on a vague level, characterized by the lack of references in the relevant paragraphs. Ideally, there should be a proof of concept or a formal understanding of how the learned behaviour enables to solve a problem that is not solved by deterministic networks. Please incorporate also the relation to the literature on neural sampling/planning/RL etc. and substantiate the claims with citations.

We have included sentences in ll. 691-696 in the revised manuscript to explain that for Poisson spiking neurons, the derived learning rule is equivalent to the one that minimizes the Kullback-Leibler divergence between the distributions of output firing and the dendritic prediction, in our case, the recurrent prediction (Asabuki and Fukai; 2020). Thus, the rule suggests that the recurrent prediction learns the statistical model of the evoked activity, which in turn allows the network to reproduce the learned transition statistics.

We have also added a paragraph to discuss the differences between previously published similar models (e.g., Kappel et al.). Please see our response above.

B) Connection to biology

The plasticity rules in the study are introduced with a vague reference to previous theoretical studies of others. Please discuss in more detail if these rules (especially the error-based learning rule) could be implemented biologically and how this could be achieved. Are there connections to biologically observed plasticity? E.g. for error-based plasticity has been discussed in the original publication by Urbanczik and Senn, or more recently by Mikulasch et al (TINS 2023). The biological plausibility of inhibitory balance

has been discussed many times before, e.g. by Vogels and others, and a citation would acknowledge that earlier work. This also leaves the question of how neurons in the songbird experiment could adapt and if the model does capture this well (i.e., do they exhibit E-I balance? etc), which might be discussed as well.

Last, please provide some testable experimental predictions. By proposing an interesting experimental prediction, the model could become considerably more relevant to experimentalists. Also, are there potentially alternative models of stochastic sequence learning (e.g., Kappel et al)? How could they be distinguished? (especially, again, why not Hebbian/STDP learning?)

We have cited the Vogels paper to acknowledge the earlier work. We have also included additional paragraphs to discuss a possible biologically plausible implementation of our model and how our model differs from similar models proposed previously (e.g., Kappel et al.). Please see our response above.

Other comments

As mentioned, a derivation of recurrent plasticity rules is missing, and parameters are chosen ad-hoc. This leaves the question of how much the results rely on the specific choice of parameters, and how robust they are to perturbations. As a robustness check, please clarify how the duration of the Markov states influences performance. It can be expected that this interacts with the timescale of recurrent connections, so having longer or shorter Markov states, as it would be in reality, should make a difference in learning that should be tested and discussed.

We thank the reviewer for pointing this out. To address this point, we performed new simulations and asked to what extent the duration of Markov states affect performance. Interestingly, even when the network was trained with input states of half the duration, the distributions of the durations of assembly reactivations remain almost identical to those in the original case (Supplementary Figure 3a). Furthermore, the transition probabilities in the replay were still consistent with the true transition probabilities (Supplementary Figure 3b). We have also included the derivation of our plasticity rule in ll. 630-670 in the revised manuscript.

Similarly, inhibitory plasticity operates with the same plasticity timescale parameter as excitatory plasticity, but, as the authors discuss, lags behind excitatory plasticity in simulation as in experiment. Is this required or was the parameter chosen such that this behaviour emerges? Please clarify this in the methods section; moreover, it would be good to test if the same results appear with fast inhibitory plasticity.

We have performed a new simulation and showed that even when the learning rate of inhibitory plasticity was larger than that of excitatory plasticity, inhibitory plasticity still occurred on a slower timescale than excitatory plasticity. We have included this result in a new Supplementary Figure 2 in the revised manuscript.

What is the justification (biologically and theoretically) for the memory trace h and its impact on neural spiking? Is it required for the results or can it be left away? Since this seems to be an important and unconventional component of the model, please discuss it in more detail.

In the model, it is assumed that each stimulus presentation drives a specific subset of network neurons with a fixed input strength, which avoids convergence to trivial solutions. Nevertheless, we choose to add this dynamic sigmoid function to facilitate stable replay by

regulating neuron activity to prevent saturation. We have explained this point in ll.605-611 in the revised manuscript.

Reviewer #2 (Recommendations For The Authors):

I noticed a couple of minor typos:

Page 3 "underly"->"underlie"

Page 7 "assemblies decreased settled"->"assemblies decreased and settled"

We have modified the text. We thank the reviewer for their careful review.

I think Figure 1C is rather confusing and not intuitive.

We apologize that the Figure 1C was confusing. In the revised figure, we have emphasized the flow of excitatory and inhibitory error for updating synapses.

Reviewer #3 (Recommendations For The Authors):

One possible path to improve the paper would be to establish a relationship between the proposed learning rules and e.g. the ones derived by Barber.

When reading the paper, I was left with a number of more detailed questions I omitted from the public review:

(1) The authors introduce a dynamic sigmoidal function for excitatory neurons, Eq. 3. This point requires more discussion and analysis. How does this impact the results?

In the model, it is assumed that each stimulus presentation drives a specific subset of network neurons with a fixed input strength, which avoids convergence to trivial solutions. Nevertheless, we choose to add this dynamic sigmoid function to facilitate stable replay by regulating neuron activity to prevent saturation. We have explained this point in ll.605-611 in the revised manuscript.

(2) For Poisson spiking neurons, it would be great to understand what cell assemblies bring (apart from biological realism, i.e., reproducing data where assemblies can be found), compared to self-connected single neurons. For example, how do the results shown in Figure 2 depend on assembly size?

We have changed the cell assembly size ratio and how it affects learning performance in a new Supplementary Figure 4. Please see our reply above.

(3) The authors focus on modeling spontaneous transitions, corresponding to a highly stochastic generative model (with most transition probabilities far from 1). A complementary question is that of learning to produce a set of stereotypical sequences, with probabilities close to 1. I wondered whether the learning rules and architecture of the model (in particular under the I-to-E rule) would also work in such a scenario.

We thank the reviewer for pointing this out. In fact, we had the same question, so we considered a situation in which the setting in Figure 2 includes both cases where the transition matrix is very stochastic (prob=0.5) and near deterministic (prob=0.9).

(4) An analysis of what controls the time so that the network stays in a certain state would be welcome.

We trained the network model in two cases, one with a fast speed of plasticity and one with a slow speed of plasticity. As a result, we found that the duration of assembly becomes longer in the slow learning case than in the fast case. We have included these results as Supplementary Figure 5 in the revised manuscript.

*Regarding the presentation, given that this is a computational modeling paper, I wonder whether *all* the formulas belong in the Methods section. I found myself skipping back and forth to understand what the main text meant, mainly because I missed a few key equations. I understand that this is a style issue that is very much community-dependent, but I think readability would improve drastically if the main model and learning rule equations could be introduced in the main text, as they start being discussed.*

We thank the reviewer for the suggestion. To cater to a wider audience, we try to explain the principle of the paper without using mathematical formulas as much as possible in the main text.

<https://doi.org/10.7554/eLife.95243.2.sa0>