Organization and Priming of Long-term Memory Representations with Two-phase Plasticity

Jannik Luboeinski1,2 · Christian Tetzlaff1,2

Received: 27 April 2021 / Accepted: 18 April 2022 / Published online: 25 June 2022
© The Author(s) 2022

Abstract
Background / Introduction In recurrent neural networks in the brain, memories are represented by so-called Hebbian cell assemblies. Such assemblies are groups of neurons with particularly strong synaptic connections formed by synaptic plasticity and consolidated by synaptic tagging and capture (STC). To link these synaptic mechanisms to long-term memory on the level of cognition and behavior, their functional implications on the level of neural networks have to be understood.

Methods We employ a biologically detailed recurrent network of spiking neurons featuring synaptic plasticity and STC to model the learning and consolidation of long-term memory representations. Using this, we investigate the effects of different organizational paradigms, and of priming stimulation, on the functionality of multiple memory representations. We quantify these effects by the spontaneous activation of memory representations driven by background noise.

Results We find that the learning order of the memory representations significantly biases the likelihood of activation towards more recently learned representations, and that hub-like overlap structure counters this effect. We identify long-term depression as the mechanism underlying these findings. Finally, we demonstrate that STC has functional consequences for the interaction of long-term memory representations: 1. intermediate consolidation in between learning the individual representations strongly alters the previously described effects, and 2. STC enables the priming of a long-term memory representation on a timescale of minutes to hours.

Conclusion Our findings show how synaptic and neuronal mechanisms can provide an explanatory basis for known cognitive effects.

Keywords Long-term memory · Recency · Interference · Priming · Synaptic tagging and capture · Recurrent spiking neural networks

Introduction

Synaptic tagging and capture (STC) provides a well-established hypothesis of the processes underlying synaptic consolidation, which has been investigated by a multitude of experimental and theoretical studies [1–9]. However, the dynamics resulting from the interplay between STC and recurrent network structures, which are characteristic for brain structures like the hippocampus and the neocortex [10, 11], and their relation to long-term memory still remain unclear.

Neuronal activity causes the increase or decrease in the transmission strength of synapses via long-term potentiation (LTP) and long-term depression (LTD). The early phase of these phenomena typically lasts for up to a few hours [12, 13]. To be maintained, the synaptic changes have to undergo a consolidation process, which is described by the STC hypothesis [1, 4]. When a group of neurons in a recurrent neural network experiences strong (learning) stimulation, this causes the formation of a Hebbian cell assembly, which thence represents a memory of the stimulation. Hebbian cell assemblies are groups of neurons exhibiting particularly strong synaptic connections, giving rise to facilitated activation of the assembly neurons [14–17]. In our previous study
[9], we developed a computational model and showed that a memory representation can be encoded by the early-phase of long-term plasticity, while STC subsequently transfers the synaptic changes to the late phase, yielding a long-term memory representation. Here, we develop this model further to investigate how the encoding and consolidation of multiple memory representations guides their interactions. This enables us to explore cognitive implications of our two-phase plasticity model.

In the brain, spontaneous activity is ubiquitous and has been shown to inherently arise from deterministic dynamics in neural networks [18, 19]. In theoretical models, spontaneous activity is usually elicited by background noise. This approach is biologically well-grounded through matching power spectra of in vivo neuronal activity on the one hand and of the stochastic processes used to model background noise on the other hand [20, 21]. Spontaneous activity causes the reactivation of existent cell assemblies via neuronal avalanches [22, 23] and thereby becomes relevant for cognition. In the current study, we use the concept of cell assembly reactivation by spontaneous activity to model the reactivation dynamics of multiple long-term memories, learned and consolidated with different protocols, as considered in neuropsychological experiments. Namely, we compare our modeling results to two different classes of experiments from cognitive psychology: free recall experiments, and priming experiments.

Free recall is a memory test that requires subjects to reproduce information from memory, for example a sequence of words, without cuing. While several models have been developed to describe free recall (e.g., [24–30]), its biological underpinnings remain elusive. With respect to the serial position of the input items, free recall experiments typically exhibit three prominent characteristics [31–35]: 1. primacy, the better recall of the items learned first, 2. contiguity, the increased probability of recalling neighboring items, and 3. recency, the better recall of the most recently learned items. Primacy is commonly related to extended rehearsal, in which the recency effect could play an indirect role [27, 36, 37]. Recency was first thought to be a solely short-term memory-dependent effect, but has then also been found to persist in long-term memory [27, 33, 38–40]. The recency effect is usually much stronger than the primacy effect [32, 33, 40], and the two effects are often jointly referred to as serial-position effect. The serial-position effect has been found to be expressed differently in experiments with different delay times before recall [33, 39, 41, 42]. Through these diverse findings, free recall experiments have essentially instigated the distinction between short-term and long-term memory storage in the human brain (also see [34]). To theoretically explain serial-position effects in long-term memory, abstract memory models have been developed [27, 37, 39]. Furthermore, two recent studies [28, 43] have shown that, rather than working-memory associations as in previous models [24, 26], longer-lasting associations between items seem to account for the effects of free recall. Prior familiarity with the learned items seems, however, not to play an important role [44]. With our model, we link neuronal and synaptic dynamics including STC to emergent network principles. Thereby, we complement the literature by providing a mechanistic explanation of the recency effect in long-term memory representations that are learned from scratch.

Priming (not to be confused with primacy) describes a large set of psychological phenomena across a multitude of timescales, which can occur either consciously or unconsciously [45–47]. All priming phenomena have in common that the recall of a certain memory is enhanced following its reactivation by a “priming stimulus.” Many studies on priming have targeted processes in the intersection between psychology and economics, for instance, biases in deciding to buy a certain product. As an example, if we see an advertisement of some brand of orange juice that we have had before, and soon afterwards go to the supermarket to do the weekly shopping, the likelihood of buying exactly that brand seen in the advertisement will be increased. Here, the advertisement acts as a priming stimulus [45, 48, 49]. For priming on a timescale of seconds, Mongillo et al. [50] have provided a model featuring the interplay of short-term plasticity and long-term plasticity. Furthermore, the authors discussed the possibility of transferring this concept to the processes on the next, longer timescales: early- and late-phase long-term plasticity. We take on this idea and show that the STC mechanisms can account for a priming effect on a timescale of hours.

**Methods**

**Model**

To simulate the dynamics of memory representations, we used a network model that comprises spiking neurons and synapses with detailed plasticity features. In this subsection, we first present our mathematical description of neurons and synapses, which is depicted in Fig. 1a, as well as the structure of our network at the population level. After that, we explain the different overlap paradigms and stimulation protocols. The parameters that we used are given in Tables 1 and 2. The model is based on our previous study [9], which should be consulted for a more detailed description.

**Neuron Model**

The dynamics of the membrane potential $V(t)$ of the Leaky Integrate-and-Fire neuron $i$ is described by [51]:

$$i$$ Springer
with reversal potential $V_{\text{rev}}$, membrane time constant $\tau_m$, membrane resistance $R_m$, synaptic weights $w_{ji}$, spike times $t_{ij}$, axonal delay time $t_{\text{ax.delay}}$, synaptic time constant $\tau_{\text{syn}}$, external background current $I_{\text{bg}}(t)$, and external stimulus current $I_{\text{stim}}(t)$. If $V_i$ crosses the threshold $V_{\text{th}}$, a spike is generated. The spike time $t_{\text{spike}}$ is then stored, and the membrane potential is reset to $V_{\text{reset}}$, where it remains for the refractory period $t_{\text{ref}}$. Under basal conditions, the membrane potential dynamics is mainly driven by an external background current that accounts for synaptic inputs from outside the network, described by an Ornstein–Uhlenbeck process:

$$\tau_m \frac{dV_i(t)}{dt} = -V_i(t) + V_{\text{rev}} + \sum_j \sum_{i,j} w_{ji} \exp \left( -\frac{(t - t_{ij} - t_{\text{ax.delay}})}{\tau_{\text{syn}}} \right) + R_m \left( I_{\text{bg}}(t) + I_{\text{stim}}(t) \right)$$

(1)

with mean current $I_0$ and white-noise standard deviation $\sigma_{\text{wn}}$. In this equation, $I(t)$ is Gaussian white noise with mean zero and standard deviation $1 / \sqrt{\Delta t}$ that approaches infinity for $\Delta t \to 0$. The Ornstein–Uhlenbeck process has the same colored-noise power spectrum as the fluctuating input to cortical neurons coming from a large presynaptic population and is therefore well-suited to model the background noise in our model [20].

In addition to the background noise, we use another Ornstein–Uhlenbeck process to model the stimulus current $I_{\text{stim}}(t)$ for learning, recall, and priming (cf. the subsection "Learning, Recall, and Priming Stimulation" below):

$$\tau_{\text{syn}} \frac{dI_{\text{stim}}(t)}{dt} = -I_{\text{stim}}(t) + I_0 + \sigma_{\text{wn}} \cdot \Gamma(t)$$

(3)

with mean current $I_0$ and white-noise standard deviation $\sigma_{\text{wn}}$. In this equation, $I(t)$ is Gaussian white noise with mean zero and standard deviation $1 / \sqrt{\Delta t}$ that approaches infinity for $\Delta t \to 0$. The Ornstein–Uhlenbeck process has the same colored-noise power spectrum as the fluctuating input to cortical neurons coming from a large presynaptic population and is therefore well-suited to model the background noise in our model [20].

In addition to the background noise, we use another Ornstein–Uhlenbeck process to model the stimulus current $I_{\text{stim}}(t)$ for learning, recall, and priming (cf. the subsection "Learning, Recall, and Priming Stimulation" below):

$$\tau_{\text{syn}} \frac{dI_{\text{stim}}(t)}{dt} = -I_{\text{stim}}(t) + I_0 + \sigma_{\text{wn}} \cdot \Gamma(t)$$

(3)

Mean and standard deviation of this process are defined by putative input spikes from $N_{\text{stim}}$ neurons, occurring at the frequency $f_{\text{stim}}$ and conveyed by synapses of weight $h_0$ (for the unit time of one second) to the neurons of our network that are to be stimulated.
Model of Synapses and STC

All spikes \( k \) that occur in neuron \( j \) are transmitted to \( i \), if there is a synaptic connection from neuron \( j \) to neuron \( i \). The postsynaptic current caused by a presynaptic spike then depends on the total weight or strength of the synapse, which is given by:

\[
    w_{ji} = \begin{cases} 
        h_{ji} + h_0 \cdot z_{ji} & \text{for } E \to E, \\
        w_{ei} & \text{for } E \to I, \\
        w_{ie} & \text{for } I \to E, \\
        w_{ii} & \text{for } I \to I. 
    \end{cases} 
\]  

(4)

Based on our previous study [9], we selected the values of the inhibition strength \( w_{ie} \) and \( w_{ii} \) (Table 1) such that we obtained excitation–inhibition balance with the same background current (cf. Eq. 2) as in [9].

While in the model all synaptic connections involving inhibitory neurons are constant, the weight of \( E \to E \) connections comprises two variable contributions to account for STC: the early-phase weight \( h_{ji} \), and the late-phase weight \( z_{ji} \). The dynamics of the early-phase weight is given by:

\[
    \frac{dh_{ji}(t)}{dt} = \tau_h \cdot \left( h_{ji} \cdot \Theta(c_{ji}(t) - \theta_h) - \frac{h_{ji}(t)}{\tau_{pl}} \right) + \xi(t), 
\]

where \( \Theta(\cdot) \) is the Heaviside function, \( \tau_h \) is a time constant, \( c_{ji}(t) \) is the calcium concentration at the postsynaptic site, and \( h_{ji}(t) \) describes additional special conditions for the induction of plasticity. The first term on the right-hand side describes the decay of the early-phase weight to the median initial value \( h_0 \), the second term describes early-phase LTP with rate \( \gamma_p \), given that calcium is above the threshold \( \theta_p \), the third term describes early-phase LTD with rate \( \gamma_d \), given that calcium is above the threshold \( \theta_d \), and the fourth term describes calcium-dependent fluctuations with standard deviation \( \sigma_{pl} \).

### Table 1

| Symbol | Value | Description | Refs. |
|--------|-------|-------------|-------|
| \( \Delta \tau \) | 0.2 ms | Duration of one time step for numerical computation | This study |
| \( \tau_m \) | 10 ms | Membrane time constant | Gerstner and Kistler [51], Dayan and Abbott [74] |
| \( \tau_{syn} \) | 5 ms | Synaptic time constant, also for external input current | Gerstner and Kistler [51], Dayan and Abbott [74], Roth and van Rossum [111] |
| \( t_{ax.delay} \) | 3 ms | Axonal spike delay | Gerstner and Kistler [51], Lin and Faber [112] |
| \( t_{ref} \) | 2 ms | Duration of the refractory period | Dayan and Abbott [74], Kobayashi et al. [113] |
| \( R_m \) | 10 MΩ | Membrane resistance | Dayan and Abbott [74] |
| \( V_{rev} \) | −65 mV | Reversal (equilibrium) potential | Dayan and Abbott [74] |
| \( V_{reset} \) | −70 mV | Reset potential | Dayan and Abbott [74] |
| \( V_{th} \) | −55 mV | Threshold potential to be crossed for spiking | Dayan and Abbott [74] |
| \( I_0 \) | 0.15 nA | Mean of the external background current | Luboeinski and Tetzlaff [9] |
| \( \sigma_{wn} \) | 0.05 nA \( t^{1/2} \) | Standard deviation for Gaussian noise in the background current | Luboeinski and Tetzlaff [9] |
| \( N_e \) | 2500 | Number of neurons in the excitatory population | This study |
| \( N_i \) | 625 | Number of neurons in the inhibitory population | This study |
| \( \rho \) | 0.1 | Probability of a connection existing between two neurons | Le Duigou et al. [11] |
| \( h_0 \) | 4.20075 mV | Median initial excitatory–excitatory coupling strength | Luboeinski and Tetzlaff [9], Li et al. [54] |
| \( w_{ei} \) | 2 \( h_0 \) | Excitatory–inhibitory coupling strength | Luboeinski and Tetzlaff [9] |
| \( w_{ie} \) | 3.5 \( h_0 \) | Excitatory–inhibitory coupling strength | This study |
| \( w_{ii} \) | 3.5 \( h_0 \) | Inhibitory–inhibitory coupling strength | This study |
| \( f_{stim} \) | 100 Hz | Frequency of stimulation | Li et al. [54], Sajikumar et al. [114] |
| \( N_{stim} \) | 25 | Number of input neurons for stimulation | Luboeinski and Tetzlaff [9] |
| \( n_{thresh} \) | 10 | Detection threshold for an avalanche | This study |
Table 2  Values were used as given in this table, unless stated otherwise

| Parameters for synaptic plasticity | Value   | Description                                                                 | Refs.                                                                 |
|-----------------------------------|---------|-----------------------------------------------------------------------------|----------------------------------------------------------------------|
| $t_{c,delay}$                     | 0.0188 s| Delay of postsynaptic calcium influx after presynaptic spike                 | Graupner and Brunel [52]                                            |
| $c_{pre}$                         | 0.6     | Presynaptic calcium contribution, in vivo adjusted                           | Graupner and Brunel [52], Higgins et al. [53], Li et al. [54]         |
| $c_{post}$                        | 0.1655  | Postsynaptic calcium contribution, in vivo adjusted                          | Graupner and Brunel [52], Higgins et al. [53], Li et al. [54]         |
| $\tau_c$                          | 0.0488 s| Calcium time constant                                                       | Graupner and Brunel [52], Li et al. [54]                             |
| $\tau_p$                          | 688.4 s | Early-phase time constant                                                   | Graupner and Brunel [52], Li et al. [54]                             |
| $\tau_z$                          | 60 min  | Protein time constant                                                       | Clopath et al. [5], Li et al. [54]                                   |
| $v_{th}$                          | 40 Hz   | Firing rate threshold for LTP induction                                      | Sajikumar et al. [3], Bear [115]                                    |
| $\gamma_p$                        | 1645.6 Hz| Potentiation rate                                                           | Graupner and Brunel [52], Li et al. [54]                             |
| $\gamma_d$                        | 313.1   | Depression rate                                                             | Graupner and Brunel [52], Li et al. [54]                             |
| $\theta_p$                        | 3       | Calcium threshold for potentiation                                          | Li et al. [54]                                                       |
| $\theta_d$                        | 1.2     | Calcium threshold for depression                                           | Li et al. [54]                                                       |
| $\sigma_{pl}$                     | 2.90436 mV | Standard deviation for plasticity fluctuations                           | Graupner and Brunel [52], Li et al. [54]                             |
| $\alpha$                          | 1       | Protein synthesis rate                                                      | Clopath et al. [5], Li et al. [54]                                   |
| $\theta_{pro}$                    | 2.10037 mV | Protein synthesis threshold                                  | Li et al. [54]                                                       |
| $\theta_{tag}$                    | 0.840149 mV | Tagging threshold                                          | Li et al. [54]                                                       |

The calcium concentration $c_{ji}(t)$ at the postsynaptic site is driven by the pre- and postsynaptic spikes $n$ and $m$, respectively:

$$\frac{dc_{ji}(t)}{dt} = -\frac{c_{ji}(t)}{\tau_c} + c_{pre} \sum_n \delta(t - t_j^n - t_{c,delay}) + c_{post} \sum_m \delta(t - t_m^n),$$

(6)

where $\delta(\cdot)$ is the Dirac delta distribution, $\tau_c$ is a time constant, $c_{pre}$ is the contribution of presynaptic spikes, $c_{post}$ is the contribution of postsynaptic spikes, $t_j^n$ and $t_m^n$ are spike times, and $t_{c,delay}$ is the delay of calcium triggered by presynaptic spikes.

In addition to the calcium dependence of early-phase plasticity, which is based on previous studies [9, 52–54], we imposed special conditions for the occurrence of plasticity to enable the stable formation, consolidation, and activation of large cell assemblies. In their original study [52], Graupner and Brunel had considered pre- and postsynaptic firing rates up to 50 Hz. However, a drawback in their calcium-based model is the chance that high spike rates on only one side of the synapse suffice to trigger LTP, which is physiologically rather not plausible. To account for the evidence that the firing of only one neuron (i.e., on one side of the synapse) is not sufficient to evoke LTP [12, 14, 55], we introduced an additional threshold $v_{th}$ which allows the induction of LTP only if both pre- and postsynaptic firing rate cross this threshold. Furthermore, we blocked early-phase plasticity outside the learning and priming phases to enable fast computation and stable activation dynamics. This is explained by a lack of novelty outside the learning and priming phases, which goes along with lowered neuromodulator concentrations [56, 57] and thereby prevents learning, as experimental studies have demonstrated [58, 59]. In summary, the two “special plasticity conditions” that we added to the model (cf. Fig. 1a) can be put down formally in the following way:

$$b_i(t) = \Theta(v_i(t) - \theta_p) \cdot \Theta(v_i(t) - v_{th}) \cdot \Theta(v_i(t) - v_{tha}) + \Theta((v_i(t) - \theta_p) \cdot \Theta(v_i(t) - v_{th}) \cdot \Theta(v_i(t) - v_{tha}) \cdot NM,$$

(7)

where the first term in the square brackets describes the LTP induction depending on the calcium concentration $c_{ji}$ and the firing rate threshold $v_{th}$, and the second term describes the LTD induction which only depends on calcium. The firing rate function $v_i(t)$ describes the firing rate of neuron $i$ at time $t$, computed using a sliding window of 0.5 s. The factor $NM$ is the neuromodulatory impact, which we assume unity during learning and priming phases, and zero otherwise.
Driven by the calcium-based early-phase dynamics, the late-phase synaptic weight is given by:

\[
\tau_{\text{z}} \frac{dz_{ji}(t)}{dt} = p_i(t) \cdot (1 - z_{ji}(t)) \cdot \Theta(h_{ji}(t) - h_0) - \theta_{\text{tag}}
\]

\[-p_i(t) \cdot (z_{ji} + 0.5) \cdot \Theta(h_0 - h_{ji}(t)) - \theta_{\text{tag}}).\]  

(8)

with the protein amount \(p_i(t)\), the late-phase time constant \(\tau_{\text{z}}\), and the tagging threshold \(\theta_{\text{tag}}\). The first term on the right-hand side describes late-phase LTP, while the second one describes late-phase LTD. Both depend on the amount of proteins being available. If the early-phase weight change \(|h_{ji}(t) - h_0|\) exceeds the tagging threshold, the synapse is tagged. This can be the case either for positive or for negative weight changes. The presence of the tag leads to the capture of proteins (if \(p_i(t) > 0\)) and thereby gives rise to changes in the late-phase weight.

The synthesis of new proteins requires sufficient changes in early-phase weights, while there is also an inherent decay of the protein amount \([5]\):

\[
\tau_{\text{p}} \frac{dp_i(t)}{dt} = -p_i(t) + \alpha \left[ \left( \sum_j |h_{ji}(t) - h_0| \right) - \Theta_{\text{pro}} \right].
\]  

(9)

For the major part of this study, early-phase weights were initialized with the value \(h_0\), while late-phase weights were initialized with the value 0. However, for the results shown in Suppl. Fig. S2, the weights were initialized according to a log-normal distribution, which is typical for synapses in the brain \([60, 61]\). In that case, the early-phase weights were initialized following the distribution:

\[
h_{\text{init}} = h_0 \cdot \exp \left( \sigma_h \cdot \mathcal{N}(0, 1) \right),
\]  

(10)

with the standard normal variable \(\mathcal{N}(0, 1)\) and the standard deviation \(\sigma_h = h_0/3\). The initial late-phase weights used for the results in Suppl. Fig. S2 were drawn according to the following distribution:

\[
z_{\text{init}} = 0.5 \cdot \exp \left( \sigma_z \cdot \mathcal{N}(0, 1) \right) - 0.5,
\]  

(11)

with the standard deviation \(\sigma_z = 1/3\).

Population Structure

Using the neuron model and the synapse model that have been explained above, we set up a neural network consisting of 2500 excitatory and 625 inhibitory neurons (depicted in Fig. 1b). The ratio of 4:1 between excitatory and inhibitory neurons is typical for cortical and hippocampal networks \([62]\). Some of the excitatory neurons receive specific inputs to form cell assemblies (see the subsection “Learning, Recall, and Priming Stimulation”). The inhibitory population provides feedback inhibition, while not directly being subject to plastic changes. The probability for the existence of a synapse between two neurons is 0.1, across the whole network. This is a plausible value for diverse cortical areas as well as for hippocampal region CA3 \([11, 63]\). For every trial, a new network structure is drawn.

Learning, Recall, and Priming Stimulation

Before we apply the first learning stimulus to our network, we let the initial activity settle for 10.0 seconds. Then, we apply our learning protocol, which delivers three stimulus pulses, of 0.1 seconds duration each, to the neurons dedicated to the first cell assembly (A, the first 600 neurons in the network). Breaks of 0.4 seconds separate the pulses. The stimulation is delivered by an Ornstein–Uhlenbeck process (see Eq. 3 introduced above). After the last pulse has ended, we let the network run under basal conditions for 3.0 seconds (Fig. 1d). For the case of intermediate consolidation, which we consider in one part of our study, we extend this period to 28800.0 seconds, such that the synaptic weights may undergo consolidation through synaptic tagging and capture. Every time that stimulation ends, we wait for 12.5 seconds and then disable early-phase plasticity (except for the decay of early-phase changes; also cf. Eqs. 5 and 7). This is reasonable considering the impact of novelty and neuromodulation on encoding \([58, 59]\).

Then, we apply another learning stimulus of the same kind as the first one to the neurons dedicated to the second cell assembly \(B\). In the non-overlapping case, this set of neurons is completely disjunct from the first set. In the cases with overlaps, 10% of the neurons in the assembly (in the main paper; also cf. Suppl. Table S1) are shared with the other assembly. The size of the overlaps is of the same order of magnitude as experimental findings, which range from around 1% to more than 40% \([64–67]\). After the learning stimulus we let the network again run under basal conditions for 3.0 seconds (28800.0 seconds in the intermediate consolidation case).

Finally, we apply a third learning stimulus of the same kind, this time to the neurons dedicated to the third cell assembly \(C\). In the non-overlapping case, this set of neurons is completely disjunct from the first and second set. For the overlap cases, still, 10% of the neurons between each two assemblies are shared. We denote the exclusive intersections between each two assemblies by \(I_{AB}, I_{AC}, I_{BC}\) (see Fig. 1c), which may make up for half the overlap or for the whole overlap between two assemblies. If existent, the intersection between all three assemblies \(I_{ABC}\) makes up for the other half of each overlap between two assemblies (cf. Suppl. Table S1). Specifically, the different subpopulations are set to occupy the following neuron indices in the network (example for overlapping case “OVERLAP10”):

\(\odot\) Springer
• \( l_{ABC} : 540–569, \)
• \( l_{AC} : 0–29, \)
• \( A : 30–539, \)
• \( l_{AB} : 570–599, \)
• \( B : 600–1109, \)
• \( l_{BC} : 1110–1139, \)
• \( C : 1140–1649. \)

By \( \tilde{A}, \tilde{B}, \) and \( \tilde{C}, \) we denote the subsets of neurons that exclusively belong to one of the assemblies. Note that in the non-overlapping case, \( \tilde{A} = A, \tilde{B} = B, \) and \( \tilde{C} = C. \)

At two points in our study, we use interleaved learning of the three memory representations. For this, instead of the three steps \( A-B-C, \) we employ the following scheme of 15 learning steps: \( A-B-C-A-C-B-B-A-C-A-B-C-C-B-A. \)

After learning, we let the network first run under basal conditions for 3.0 seconds and then consolidate for 28800.0 seconds (eight hours, see Fig. 1e). This causes all early-phase weight changes to either decay or be transferred to the late phase, thereby yielding long-term memory representations. After that, we start measuring the spontaneous activity (see subsection below).

For the testing of memory function, shown in Fig. 2, we apply a recall stimulus to randomly drawn 20% of the neurons of an assembly, delivered by one pulse 0.1 seconds long, with otherwise the same parameters as for the learning stimulus. During this, we leave early-plasticity switched off.

For the priming of a memory representation, which we investigate in the last part of our study, we apply a stimulus to all neurons of one selected assembly (alternatively, to randomly drawn 50% of the assembly neurons, the results of which are only shown in the Supplementary Information). The stimulus is delivered via one pulse 0.1 seconds long, with otherwise the same parameters as for the learning stimulus (cf. Fig. 6a). The priming stimulus is followed by a period of consolidation (and early-phase decay) of varied duration, after which we start measuring the spontaneous activity.

### Determining Avalanches in the Spontaneous Activity

After learning, consolidation, and possibly priming of the cell assemblies, we measure the network activity across three minutes to evaluate the spontaneous activation of the assemblies. To this end, we compute the likelihood that an avalanche with a certain minimum size occurs in a specific assembly, averaged across the whole period of three minutes.

In addition, we measure the plain mean firing rate of each assembly and the control population across the whole period of three minutes.

To determine if an avalanche occurs, we count the number of spikes that are produced by the neurons of each cell assembly in a time frame of 10 ms. Then, we compare these numbers to the threshold value \( n_{\text{thresh}} \) and consider an assembly active if the threshold is reached. Thus, multiple assemblies may be active at the same time. Finally, we compute the likelihood of an avalanche in assembly \( A, \) for instance, by dividing the number of periods \( N_{\text{act}}(A) \) in which the assembly was active by the total number of time frames (given by the total duration of the simulation \( t_{\text{max}}/10 \text{ ms} \)):

\[
I(A) := \frac{N_{\text{act}}(A)}{t_{\text{max}}/10 \text{ ms}}. \quad (12)
\]

The likelihoods for avalanches in \( B \) and \( C \) are computed analogously. The threshold value \( n_{\text{thresh}} \) is chosen to be about the border value of the 99\% quantile for the number of spikes of an assembly in one time frame (cf. Suppl. Fig. S11).

### Computational Implementation and Software Used

We have used C++ in the ISO 2011 standard to implement our simulations. To compile and link the code, we employed g++ in version 7.5.0 with boost in version 1.65.1.

For further data analysis, we used Python 3.7.3 with NumPy 1.20.1 and pandas 1.0.3. For the creation of plots, we used gnuplot 5.0.3, as well as Matplotlib 3.1.0.

We generated random numbers using the generator ‘minstd_rand0’ from the C++ standard library, while the system time served as the seed. We implemented a loop in our code which ensured that for each distribution a unique seed was used. The simulations that had to be performed for this study were computationally extremely demanding (each running for several hours). Thus, it was very fortunate that we had the opportunity to run most of our jobs on the computing cluster of the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG).

All data underlying the results presented in this study are contained in this manuscript and the Supplementary Information. The data can be reproduced using the simulation code and the analysis scripts that we have released under the Apache-2.0 license [68]. The code continues to be developed on https://github.com/jlubo/memory-consolidation-stc.

### Results

In this study, we have employed a spiking recurrent neural network model based on a previous model of memory consolidation via synaptic tagging and capture (STC) [9]. The STC principles provide a biologically plausible underpinning of the long-term maintenance of synaptic plasticity and, thereby, of long-term memory representations [4, 8, 9, 55]. In a proof-of-principle approach, we investigate multiple paradigms of the formation, consolidation, and interaction of three cell assemblies in the recurrent network. In the first
part of our study, we measure the likelihood of activation of these three assemblies and show that two factors mainly guide their activation: 1. the learning order, and 2. the type of overlaps between the assemblies. By mechanistically describing the learning order effect, our model provides a possible neural correlate for the recency effect in free recall experiments. In the second part of our study, we focus on the functional consequences of STC with respect to such interactions. To this end, we first show that intermediate consolidation phases in the stimulation protocol strongly alter the learning order effect. Thus, we make predictions about memory dynamics on different time scales which can

Fig. 2 Three non-overlapping memory representations are successfully learned and consolidated. Learning order: A-B-C. a Raster plot showing the spiking in the whole network before and after learning and consolidation. At specified times, additional input is applied to 20% of the neurons of one assembly. Each assembly comprises 600 neurons. Spikes of excitatory and inhibitory neurons are shown by blue and red dots, respectively. Avalanches occurring in an assembly are indicated by bars in the respective color. b Firing rates corresponding to the raster plots (computed via bins of 10 ms). Shaded areas indicate the time frame of additional stimulation. The dashed line indicates the detection threshold for avalanches. In the two leftmost panels, gray bars next to the peaks emphasize the increase after consolidation in the mean value of the firing rate of A during the induced activation (+26.3%) and in the duration of the activation (+11.1%). The measure Q based on the stimulation for learning and recall also indicates pattern completion (cf. [9]). c Weight distribution of the synapses within the assemblies, revealing the spread of LTP and LTD. Weight values were discretized into 100 bins, and the relative frequency was normalized across all bins and all groups of synapses. d Abstract weight matrix showing the mean weight within and between all subpopulations of the excitatory population. The asymmetric nature of the matrix reveals the directionality of the coupling strengths. Data in c and d were averaged over 10 trials and taken after learning and consolidation relative to the initial value before learning ($h_0$).
be tested in neuropsychological experiments. Furthermore, depending on consolidation, we encounter a differential effect of hub-like overlap structures, indicating that STC-based mechanisms might be linked to the diverse behavior observed for similarity of memories, specifically, to retroactive interference and enhanced recall [69–73]. As a final step, we show that the interaction between early- and late-phase long-term plasticity constitutes a synaptic mechanism for priming phenomena lasting from minutes to hours.

The neural network that we used consisted of 2500 excitatory and 625 inhibitory leaky integrate-and-fire neurons. To form a cell assembly, we applied strong stimulation to a subset of 600 excitatory neurons in the network and thereby elicited synaptic plasticity, modeled by a spike-timing-dependent calcium model and a firing rate condition. This early-phase plasticity could be followed by consolidation via STC, which transfers substantial weight changes (i.e., weight changes that have sufficed to set a tag) to the long-lasting late phase. The synaptic and neuronal level of the model is depicted in Fig. 1a, linked to a schematic of the network level in Fig. 1b (for more details, see “Methods” and [9]).

After phases of learning and consolidation, we investigated the activation of the assemblies by considering neuronal avalanches within each assembly [22]. To consider an assembly active, we utilized the 99% quantile of the number of spikes within a time window as threshold value. Hence, we considered the 1% of the time windows with most spikes (for more details, see “Methods” and Suppl. Fig. S11). Note that in contrast to the colloquial sense of the word “avalanche,” neural avalanches are defined to refer, depending on the time window, not only to longer but also to relatively short periods of activation.

Indicated by findings on precisely timed events [74–76], the likelihood of occurrence of neuronal avalanches seems more relevant to behavioral function than the mere average firing rate within each assembly. Nevertheless, by comparing the results of the likelihood of avalanche occurrence to the average firing rates (Suppl. Figs. S6 and S7), we found that similar conclusions can be drawn for both measures.

**Learning and Consolidation of Multiple Long-term Memory Representations**

Before investigating the spontaneous activation of long-term memory representations, we will demonstrate how the network learns and consolidates them. We start from an unbiased network in which all weights have the same value. Then, learning proceeds according to the protocol shown in Fig. 1d, followed by consolidation and the investigation of the spontaneous activity shown in Fig. 1e. Substantially extended breaks between learning stimuli, such that “intermediate consolidation” occurs in those phases, are considered in a later part of this study. In the main text, we consider four organizational paradigms, shown in Fig. 1c. The differences between these paradigms are visualized by small pictures. The different paradigms will be examined in the next subsection, while we restrict ourselves to the non-overlapping case in this subsection.

As a first step, after learning and consolidating three non-overlapping memory representations, we tested whether all three assemblies served as functional memory representations. To this end, we examined the ability to recall the learned and consolidated assemblies by applying stimulation to 20% of the neurons of each assembly, subsequently (Fig. 2a,b). We found the response caused by this stimulation to be stronger and to last longer than the response to the same stimulation in a naive network before learning (the example for A is shown in Fig. 2a,b in the two leftmost panels), which confirms the functionality of the memory representations. Moreover, we observed avalanches within the assemblies during spontaneous activity after learning and consolidation (Fig. 2a,b). Note that the synchronicity in the activity is also increased after learning and consolidation (Fig. 2a). Since the average excitatory synaptic weight in our network after learning and consolidation is increased (103.8%, integrated over all weights in Fig. 2c), the synchronicity can be explained by findings on network dynamics made by Brunel [77]. These findings show that the strengthening of excitatory connections in a network causes a shift toward a regime of synchronized regular activity. Therefore, we conclude that our network operates in the critical regime close to the phase transition, which is typical for the occurrence of avalanches [22, 23].

In addition to the activity, we scrutinized the weight structure in the network after learning and consolidation. As expected, we found that the weights within each assembly are much higher than the weights in the remainder of the network (Fig. 2c). Remarkably, however, the weight distributions of the three assemblies differ. This is already a hint to the bias in assembly activation introduced by the learning order, which we will investigate in the following subsection. Complementary to the weight distribution, the abstract weight matrix in Fig. 2d reveals the directionality of the coupling strengths within and between subpopulations of the network after learning and consolidation. The matrix reflects the highly potentiated weights within each cell assembly, but also the depressed weights between the assemblies. Presumably, this depression arises from activation of the other assemblies upon learning a new assembly. The weights involving control neurons, on the other hand, exhibit only slight plasticity. Remarkably, synapses incoming from assembly neurons to control neurons are on average slightly depressed, while synapses outgoing from control neurons remain unchanged. Comparing with the weights after learning but before consolidation (Suppl. Fig. S1) reveals that at that time there is also depression
in the synapses outgoing from control neurons, which is yet too weak to become consolidated. The depression of the weights presynaptic to the control neurons probably also emerges during learning of the assemblies, while the weights postsynaptic to the control neurons only receive indirect activation during learning. The higher calcium increase through presynaptic spikes in our model might increase this effect further.

Our results on learning, consolidation, and the interaction between memory representations (investigated below) also hold considering randomly initialized weights that follow a log-normal distribution (Suppl. Fig. S2). Such a distribution has been found for synaptic strengths in the brain (cf. [60, 61]) and can therefore mimic the influence of previously acquired memory representations in the network.

For the overlapping paradigm, the weight distribution and the abstract weight matrix are given in Suppl. Figs. S3 and S4. Those results reveal further asymmetries for the weights of the intersections between assemblies, exemplifying the rich dynamics covered by our model.

**Learning Order and Overlaps Guide Spontaneous Activity**

After the learning and consolidation of the three assemblies, we systematically investigated the spontaneous reactivation of the assemblies driven by background noise. We collected data from three minutes, which is a time span similar to the recall period in free recall experiments [32, 34, 41]. A sample of spontaneous activity and the detected avalanches is shown in Fig. 3a, b.

Overlap of memory representations on the network level is related to memory similarity on the cognitive level [28, 29, 64, 66, 78, 79] and is therefore of huge relevance to the understanding of memory dynamics. Here, we consider four different paradigms of overlaps. These are shown in Fig. 1c (also see “Methods”), where small images picture possible real-life examples of non-overlapping and overlapping memory representations: representations of entities from three different categories, like chicken, house, and bicycle, do not overlap, while representations from the same category, like three types of birds, overlap. Mixed cases occur if two entities are

![Fig. 3 Spontaneous activity of long-term memory representations, driven by background noise.](image)

**Fig. 3** Spontaneous activity of long-term memory representations, driven by background noise. **a** Raster plot showing the spontaneous spiking activity in the whole network induced by background noise. Each assembly comprises 600 excitatory neurons (spikes indicated by blue dots). Spikes of inhibitory neurons are shown by red dots. Avalanches occurring in the assemblies are indicated by bars in the assembly colors. Assemblies do not overlap and are learned in the order: A-B-C. **b** Firing rates corresponding to the raster plot (computed via bins of 10 ms). The dashed line indicates the threshold for the detection of an avalanche. **c** Overview of the likelihood of avalanches in different organizational paradigms, learned with standard protocol (order: A-B-C). Averaged over mean likelihood from 10 networks. Error bars show the 95% confidence interval. Braces and plus/minus signs indicate differences to the “Overlap 10%” paradigm. **d** Overview of the likelihood of avalanches as in c, but learned with extended, interleaved protocol (order: A-B-C-A-C-B-A-C-A-B-C-C-B-A). Data in c and d were averaged over 10 trials.
in the same category while there is no common category for all three, for instance, bicycle and horse are both means of transportation, and chicken and horse are both farm animals. In the Supplementary Information to our study, we present results on more paradigms, including such with overlaps of varied percentage (Suppl. Figs. S6 and S7). For an overview over these paradigms, see Suppl. Table S1. The most important result from these additional paradigms is that in the majority of cases, larger overlaps further increase the effects that we describe in the following.

Considering the mean likelihood of activation in each of the four paradigms from Fig. 1c with assemblies that were learned in the order A-B-C, we encounter the first main result of our study: the learning order significantly biases the spontaneous activity, such that the activation of more recently learned assemblies is promoted (shown in Fig. 3c). While this is already the case without overlap, overlap between assemblies increases the effect even further.

Our second finding is indicated in Fig. 3c by braces with a plus or minus sign. The braces demonstrate differences between the paradigms with less overlaps and the “Overlap 10%” paradigm. The lack of an overlap between two assemblies (in the abundance of the other overlaps) creates a hub-like structure, where the assembly that overlaps with the two other assemblies is the hub. We find that such hub-like structure reduces the learning order effect for the two assemblies with less overlaps. This means, the activation of the assembly that is disadvantaged by the learning order effect (A in the “no AC” paradigm, B in the “no BC” paradigm) is increased, while the activation of the assembly favored by the learning order effect (C) is decreased. As we will show below, this effect as well as the learning order effect itself are essentially caused by LTD.

For gathering further knowledge about the learning order effect and to examine other effects without the bias of the learning order, we tried to eradicate the learning order effect (also cf. the following subsection on priming). To this end, we employed a learning protocol for the interleaved learning of the three assemblies, using an extended sequence with the order A-B-C-A-C-B-B-A-C-A-B-C-C-B-A (random with the constraint that every assembly occurs equally often). Compared to Fig. 1d, the interleaved protocol simply contains more steps, but otherwise stays the same. In the non-overlapping paradigm, the interleaved learning successfully eradicates any significant impact of the learning order (Fig. 3d). In the other paradigms, however, a certain impact remains. Note that the most recently learned assembly in the extended sequence is A, which is also the assembly reactivated most often.

In our findings shown above, we had seen lower weights in the previously learned assemblies as compared to the newly learned ones (cf. Fig. 2c). Assuming that there might be a decrease in weights (i.e., LTD) in the pre-existing assemblies upon learning a new assembly, we considered the temporal evolution of the synaptic weights with our standard protocol (Fig. 4a). The temporal traces show that the weights within assembly A decrease during learning stimulation for B and C. Comparing these results to the outcome of a standard protocol in which we blocked LTD (Fig. 4c; cf. “Methods”), we see that the decrease is also blocked. A similar difference is evident in the late phase (Fig. 4b, d). We further found that the weight distribution in the case where LTD is blocked is almost equal for all three assemblies (Fig. 5a and Suppl. Fig. S5c), which is also in contrast to the standard case (cf. inset in Fig. 5a).

In our model, LTD is caused by medium synaptic calcium levels (cf. Eq. 5), which are evoked by moderate neuronal activity (cf. Eq. 6). Thus, the weakening of pre-existing assemblies seems to originate from their moderate activation evoked by the learning stimulus that acts on the network to form a new assembly. The interleaved protocol employs several times more learning stimuli than the standard protocol to counteract the weakening of individual assemblies (Fig. 4e, f and Suppl. Fig. S5b), which might be the reason why it also enables the attenuation of the learning order effect (cf. Fig. 3d).

Based on these findings, we reasoned that LTD could be the cause of the learning order effect. Therefore, as a next step, we investigated the spontaneous activation in the case of blocked LTD. We found what had already been hinted by the weight distribution (cf. Fig. 5a): the learning order effect indeed vanishes if there is no LTD (Fig. 5c). In addition, the previously observed effect of hub-like structures vanishes. Now, the activation of the hub assembly (B in the “no AC” paradigm, A in the “no BC” paradigm) clearly stands out, as it is expected for an LTP-only model. Furthermore, the overall likelihood of activation increases. In contrast to that, an LTD-only condition would presumably prevent the learning of new memory representations, while stimulation could still cause the weakening of existing memory representations.

### Impact of Synaptic Tagging and Capture on Organizational Effects

Next, we investigated the functional consequences of STC with respect to the effects arising from the different organizational paradigms which we characterized in the previous subsection. To this end, we considered the activation of assemblies that had undergone intermediate consolidation after being learned, meaning that the breaks shown in Fig. 1d lasted for 8 hours instead of 3 seconds. Although LTD was switched on for these investigations, much of the learning order effect vanishes, similarly to the no-LTD case (Fig. 5b, d and Suppl. Fig. S5d). The effect of hub-like structures is similar to the no-LTD case as well. It seems that here, LTD is not significantly expressed in the synapses.
within the assemblies, even though LTD occurs in other synapses (cf. Fig. 5b and inset in Fig. 5a). Apparently, if only one assembly is learned before the next one follows after hours, much of the depression which evokes the learning order effect is too weak to become consolidated. This indicates that late-phase weight change is more than just a low-pass filtered version of early-phase weight change (also cf. Suppl. Fig. S1 and Suppl. Fig. S5). Instead, the criticality of cooperation between synapses to trigger protein synthesis, and of the synaptic tag, renders the transition from early- to late phase weight a nonlinear process.

Fig. 4 Temporal evolution of synaptic weights following different stimulation protocols. The plots show the weights within assembly A for the non-overlapping paradigm. a, b Standard learning/consolidation protocol (order: A-B-C); c, d like the standard protocol but with LTD blocked; e, f extended, interleaved protocol (order: A-B-C-A-C-B-B-A-C-A-B-C-C-B-A). Panels b, d, f show data from the same simulations as a, c, e but on a longer timescale, revealing consolidation. Data in all panels were averaged over 10 trials; error bands represent the 95% confidence interval, while mostly too small to be visible. Values are given relative to the initial value before learning ($h_0$); late-phase weight has been shifted by 1 (cf. Methods) to be visualized accordingly.
and substantial late-phase LTD, leading to a strong learning order effect and its reduction for assemblies with less overlap, the “intermediate consolidation” protocol rather causes operation in the regime of late-phase LTP only, leading to a weaker learning order effect and promoted activation of hub assemblies. This indicates that the interaction of plasticity processes on different timescales, namely early- and late-phase synaptic plasticity described by STC, enables differential memory effects at the cognitive level.

**Priming on Long Timescales**

After investigating the functional impact of different organizational paradigms on recency, interference, and enhanced activation through similarity, we provide in this subsection a proof of principle to show that our two-phase plasticity model is also able to describe priming on long timescales. To this end, we use a protocol that includes a brief additional stimulus delivered to all neurons of one selected assembly after learning and consolidation (Fig. 6a). This stimulation causes the respective assembly to be transiently primed for later reactivation, which we investigate by considering the likelihood of avalanche occurrence through background noise as it has been done in the previous subsections. Our results demonstrate that there exists a mechanism for priming enabled by the interplay of early- and late-phase long-term plasticity (Fig. 6b–g). The priming decays on a timescale of minutes to hours, corresponding to the dynamics of early-phase plasticity.

For investigating priming in the absence of other effects, we sought to attenuate the learning order effect and therefore used the extended interleaved protocol introduced previously. For the non-overlapping case, this approach works fine as it almost entirely eliminates the influence of the learning order (Fig. 6b). However, in the overlapping case (Fig. 6c), the learning order effect still plays a minor role, since in the basal state, A is more active than B and C. Finally, we have to...
mention that in rare cases, priming stimulation after learning with the interleaved protocol seems to disrupt the recallable memory structure and render reactivation impossible (this occurred in one of ten trials).

The mean total weight shown in Fig. 6d, e exhibits roughly the same qualitative time course as the activation (Fig. 6b, c). Interestingly, in the non-overlapping paradigm, the priming stimulation causes the primed memory representation to behave in an almost attractor-like manner with a likelihood of activation of more than 0.4 (Fig. 6b; also see “Discussion”). It should further be noted that in Fig. 6d, the mean weight in assembly B decreases to a lower value than
it had before priming, which may be explained by strong depression of a few synapses in $B$ entering the late phase, while most synapses are subject to early-phase potentiation. Considering the mean early-phase weight (Fig. 6f, g), we find that potentiated weights of assembly $B$ and depressed weights of assemblies $A$ and $C$ correlate with the priming effect of $B$. Around 420 min after priming, the early-phase changes have mostly vanished, while the activation of the assemblies (Fig. 6b, d) has also returned to its baseline level before priming.

Data for more paradigms are shown in Suppl. Fig. S8. The case where only 50% of the assembly neurons receive priming stimulation is shown in Suppl. Fig. S9 (more paradigms in Suppl. Fig. S10). The response to this stimulation leads to the interesting indication that priming can also be mainly driven by LTD. However, this result is not statistically significant, and further investigation on this topic was beyond the scope of this study.

Discussion

In this study, we have shown by proof of principle that synaptic mechanisms of calcium-based plasticity and STC enable to robustly learn, consolidate, organize, and prime long-term memory representations. Our results demonstrate that these features give rise to specific characteristics as they are found in neuropsychological experiments. The robustness of our model against parameter variations is shown by the relative similarity of the results for the standard protocol and the intermediate consolidation protocol (Figs. 3c and 5d), as well as for early- and late-phase weights (Suppl. Fig. S1c and Fig. 3c), and for late-phase weights resulting from randomly initialized weights (Suppl. Fig. S2c).

Measuring the likelihood of spontaneous activation of three cell assemblies in a spiking recurrent neural network after learning and consolidation, we found two factors that guide the activation: 1. the learning order, and 2. the overlap between the assemblies. As an experimental study has shown [66], overlap can account for memory association while not being necessary for the recall of the original separate memories, which matches our findings for the non-overlap and overlap paradigms. There is, moreover, experimental evidence that memories can be potentiated and depotentiated independently of each other [80], which also seems possible in our model. Data from free recall experiments typically exhibit primacy (high probability of activating the first learned items), recency (high probability of activating the last learned items), and contiguity (higher probability of transition between neighboring items) [32, 34]. By our model, we may have provided a mechanistic explanation for the recency effect. However, the direct reproduction of data from free recall experiments that also include the typically weaker primacy and contiguity effects (e.g., [35, 40, 81, 82]) would require a larger number of cell assemblies, which was beyond the scope of this study. Since our detailed network model has a high demand for computational resources, the present study was limited to proof-of-principle investigations with medium-sized networks of 3125 neurons and three assemblies. Nevertheless, in the future, more computational power will enable the simulation of larger networks, which may serve to directly match data from free recall experiments and to provide a mechanistic explanation of primacy and contiguity on a timescale of hours, complementary to previous models [29, 30, 43]. This will, however, require careful tuning of certain plasticity parameters in a biologically plausible way. Related to this, it is important to note that storing more and more memory representations will cause interference and can eventually lead to catastrophic forgetting [83, 84]. It has, however, been shown that multi-phase plasticity can counter catastrophic forgetting [85–87]. Studying this with our model could be the goal of a future study serving to explore possibilities of technical application. Another interesting approach to relate to models targeting memory capacity (e.g., [87, 88]) can be to adjust our network such that the memory representations constitute attractors. This would enable the study of transitions between attractor states, which have been proposed to account for the switching between neuropsychological concepts [29, 89, 90], building on the vast amount of previous studies on attractor neural networks. First investigations of attractor dynamics with a slightly modified version of our model have been presented in [91].

In addition to recency in free recall, our model may provide a mechanistic basis for retroactive interference [69–71] and enhanced recall of similar memories [72, 73]. Our findings show that it depends on the learning/consolidation protocol whether a hub-like overlap structure enhances or attenuates the activation of particular memory representations (see Figs. 3c and 5d). Furthermore, while overlap causes memory-deteriorating interference in the “standard” protocol (Fig. 3c and Suppl. Fig. S6), it causes memory enhancement in the case of longer intervals between learning stimuli with otherwise equal parameters (“intermediate consolidation” protocol; Fig. 5d and Suppl. Fig. S7). This suggests that in addition to general memory improvement [9], STC can play a functional role in enhancing the retrievability of overlapping memory representations.

From our theoretical investigations, we made the prediction that learning order effects may be essentially caused by LTD (Fig. 5a, c). This can be tested experimentally by a protocol that blocks LTD, for example, via the mGlu receptor antagonist MCPG or via BDNF [92, 93]. Such experiments could, for instance, target the recall of contextual or fear memories (cf. [67, 94]). At the same time, it is important to think about other mechanisms that could disrupt cell assem-
blies to yield learning order effects. Homeostatic plasticity [95, 96] and synaptic rewiring [97] constitute possible candidates for this, although acting on longer timescales than the plasticity processes that we have considered here.

Our theoretical findings are supported by experiments which have shown that overlap of assemblies acquired on similar timescales as in our “intermediate consolidation” protocol can enhance recall [67, 94]. However, these studies also suggest that memories acquired even further apart in time tend to be separated, which may be investigated in future studies that should consider effects of processes acting on longer timescales. Furthermore, experimental investigation of our “intermediate consolidation” protocol with human subjects in a free recall task may yield a large gain in knowledge. The difference between this protocol and typical free recall protocols is that after learning a memory representation, we wait until it is consolidated (i.e., for about 8 hours) and then learn the next representation, whereas in free recall experiments, the whole list of items is usually learned in short time before consolidation can take place (similar to our “standard” protocol). The experimental investigation of the “intermediate consolidation” protocol would thus enable to test our predictions and yield further interesting insights into the interaction of short- and long-term memory. An issue could be that the subjects would have to stay awake for the whole duration of the protocol, which is about 24 hours, to avoid a biasing effect from sleep consolidation. However, similar experiments have already been done, testing the impact of sleep deprivation on free recall performance [40, 98].

Our study mainly focused on classical Hebbian cell assemblies based on potentiated excitatory synapses. While plasticity of inhibitory synapses, so-called inhibitory plasticity, can enhance excitation–inhibition balance in spiking neural networks [99] and enables functional discrimination between memory representations [100], balanced network states and functional memory recall can be achieved without requiring this type of plasticity (cf. [9, 101]). While it would be very interesting to investigate the influence of inhibitory plasticity on our model, adding such further dynamics will render the simulation and analysis even more complicated.

For our simulations, we assumed early-phase plasticity to be disabled during certain phases. As we pointed out in the “Methods” section, this can be justified by lack of novelty or attention which comes along with lowered neuromodulator concentrations [56, 57], thereby preventing learning (cf. [58, 59]). If, however, early-phase plasticity were enabled all the time, spontaneous activity could exert the following influence on the structure of a cell assembly: either the assembly would stay unaffected, as we have shown for smaller cell assemblies in our previous study [9], or, LTD would cause degradation by downregulating the synaptic weights, or, LTP would cause reinforcement by increasing the synaptic weights. While the latter is not possible due to the firing rate threshold in our model, we assume that spontaneous occurrence of LTD, if any, would not qualitatively change our results. Furthermore, the analysis of our model in [91] shows that input of up to ∼3 Hz will not elicit early-phase plasticity.

In the final part of our study, we modeled the priming of one of the three previously learned and consolidated memory representations on a timescale of minutes to hours. We demonstrated that this effect can be explained by the interplay between early- and late-phase long-term plasticity as follows: After consolidation, the structure of the memory representations is permanently stored by the late-phase weights, while additional transient information, like priming information, is stored by early-phase weights, which decay across minutes to hours. However, our results also indicate that priming information may “spill over” to the late phase, causing permanent changes in the structure of a memory representation. In our study, the priming effect is strongest if priming is only 10 minutes ago, and it declines until it has vanished almost completely after 7 hours (Fig. 6b, c). This time dependence, as well as the impact of priming in general, may be affected by the strength and duration of the priming stimulus which “loads the representation into short-term memory.” Moreover, the effect of the temporal delay may be influenced by a wide range of further experimental conditions, making it difficult to draw a general conclusion, which has been pointed out by others earlier [102]. Since our results shall constitute a proof of principle, we leave it to further studies to investigate the influence of different kinds of stimulation.

We showed that our model can account for direct positive priming, but our model can likely be extended with relatively low effort to account for other types of priming. We have already encountered predominance of LTD in the case where only 50% of the assembly neurons receive priming stimulation (Suppl. Figs. S9 and S10). While this paradigm still causes a slightly positive priming effect, lowering this fraction even further or simply applying background activity in the presence of early-phase plasticity should revert the outcome and cause a negative priming effect. Such further results would be of great importance, especially because a mechanistic biological theory for negative priming still remains to be found (cf. [46, 103, 104]).

Although various studies have found priming to be relatively independent of the hippocampus [105, 106], other studies have demonstrated a functional relation [107]. Neocortical regions adjacent to the hippocampus have, furthermore, been found to be critical for priming [106]. For this reason, neural networks with hippocampal and neocortical characteristics as the ones presented here are important to understand general processes that could underlie priming. A tuned version of our model might, moreover, account for semantic...
priming, which is the case if the stimulation of an assembly that overlaps with a second assembly causes primed activation of the second assembly (e.g., in the way that the word “chicken” would prime the word “duck”) [47, 102, 108, 109]. Finally, the fact that the interplay of early- and late-phase plasticity enables priming is another hint that the these two phases might constitute correlates for the concepts of retrieval strength and storage strength [9, 110].

Summary

We have shown that early- and late-phase LTP on the one hand and LTD on the other hand, implemented by calcium-based synaptic plasticity and STC, can critically influence the functionality of memory representations. Our results provide neural correlates to at least partially explain different cognitive effects. In the future, our model could serve as a mechanistic basis for the more detailed reproduction of findings obtained from neuropsychological experiments.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12559-022-10021-7.

Acknowledgements We thank the members of the Department of Computational Neuroscience, especially Sebastian Schmitt, for many helpful comments on this study. The research was funded by the German Research Foundation (CRC1286, project C1, project #419866478) and by the H2020 – FETPROACT project Plan4Act (#732266).

Author Contributions Conceptualization, methodology, and investigation are contributed by CT and JL; software and simulations are contributed by JL; data curation, formal analysis, and visualization are contributed by JL; funding acquisition and supervision are contributed by CT; writing and editing are contributed by CT and JL.

Funding Open Access funding was enabled and organized by Projekt DEAL.

Declarations

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of Interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. References

1. Frey U, Morris RGM. Synaptic tagging and long-term potentiation. Nature. 1997;385:533–6.
2. Sajikumar S, Frey JU. Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. Neurobiol Learn Mem. 2004;82(1):12–25.
3. Sajikumar S, Navakkode S, Frey JU. Identification of compartment- and process-specific molecules required for “synaptic tagging” during long-term potentiation and long-term depression in hippocampal CA1. J Neurosci. 2007;27(19):5068–80.
4. Redondo RL, Morris RGM. Making memories last: the synaptic tagging and capture hypothesis. Nat Rev Neurosci. 2011;12:17–30.
5. Clopath C, Ziegler L, Vasilaki E, Büssing L, Gerstner W. Tag-triggered consolidation: a model of early and late long-term potentiation and depression. PLoS Comput Biol. 2008;4(12):e1000248.
6. Barrett AB, Billings GO, Morris RGM, van Rossum MCW. State based model of long-term potentiation and synaptic tagging and capture. PLoS Comput Biol. 2009;5:e1000259.
7. Smolen P, Baxter DA, Byrne JH. Molecular constraints on synaptic tagging and maintenance of long-term potentiation: a predictive model. PLoS Comput Biol. 2012;8(8):e1002620.
8. Ziegler L, Zenke F, Kastner DB, Gerstner W. Synaptic consolidation: from synapses to behavioral modeling. J Neurosci. 2015;35(3):1319–1334.
9. Luboinski J, Tetzlaff C. Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks. Commun Biol. 2021;4(275).
10. Amaral DG, Ishizuka N, Claiborne B. Neurons, numbers and the hippocampal network. In: Understanding the Brain Through the Hippocampus. vol. 83 of Progress in Brain Research. Elsevier; 1990. p. 1–11.
11. Le Duigou C, Simonnet J, Teleńczuk M, Fricker D, Miles RM. Recurrent synapses and circuits in the CA3 region of the hippocampus: an associative network. Front Cell Neurosci. 2014;7:262.
12. Bliss TV, Collingridge GL. A synaptic model of memory: long-term potentiation in the hippocampus. Nature. 1993;361(6407):31.
13. Abraham WC. How long will long term potentiation last? Philos Trans R Soc B. 2003;358:735–44.
14. Hebb DO. The Organization of Behavior. 1st ed. New York/NY, USA: Wiley; 1949.
15. Martin SJ, Grimwood PD, Morris RG. Synaptic plasticity and memory: an evaluation of the hypothesis. Ann Rev Neurosci. 2000;23(1):649–711.
16. Buzsáki G. Neural syntax: cell assemblies, synapsembles, and readers. Neuron. 2010;68(3):362–85.
17. Eichenbaum H. Barlow versus Hebb: When is it time to abandon the notion of feature detectors and adopt the cell assembly as the unit of cognition? Neurosci Lett. 2017.
18. Hartmann C, Lazar A, Nessler B, Triesch J. Where’s the noise? Key features of spontaneous activity and neural variability arise through learning in a deterministic network. PLoS Comput Biol. 2015;11(12):e1004640.
19. Effenberger F, Jost J, Levina A. Self-organization in balanced state networks by STDP and homeostatic plasticity. PLoS Comput Biol. 2015;11(9):e1004420.
20. Destexhe A, Rudolph M, Paré D. The high-conductance state of neocortical neurons in vivo. Nat Rev Neurosci. 2003;4(9).
21. Moreno-Bote R, Parga N. Response of integrate-and-fire neurons to noisy inputs filtered by synapses with arbitrary time-scales: firing rate and correlations. Neural Comput. 2010;22(6).
22. Plenz D, Thiagarajan TC. The organizing principles of neuronal avalanches: cell assemblies in the cortex? Trends Neurosci. 2007;30(3):101–10.

23. Tetzlaff C, Okujeni S, Egert U, Wörgötter F, Butz M. Self-organized criticality in developing neuronal networks. PLoS Comput Biol. 2010;6(12):e1001013.

24. Raaijmakers JG, Shiffrin RM. Search of associative memory. Psychol Rev. 1981;88(2):93.

25. Bradski G, Carpenter GA, Grossberg S. STORE working memory networks for storage and recall of arbitrary temporal sequences. Biol Cybern. 1994;71(6):469–80.

26. Howard MW, Kahana MJ. A distributed representation of temporal context. J Math Psychol. 2002;46(3):269–99.

27. Davelaar EJ, Goshen-Gottstein Y, Ashkenazi A, Haarmann HJ, Usher M. The demise of short-term memory revisited: empirical and computational investigations of recency effects. Psychol Rev. 2005;112(1):3.

28. Romani S, Pinkovskiy I, Rubin A, Tsodyks M. Scaling laws of associative memory retrieval. Neural Comput. 2013;25(10):2523–44.

29. Recanatesi S, Katkov M, Romani S, Tsodyks M. Neural network model of memory retrieval. Front Comput Neurosci. 2015;9(149).

30. Lansner A, Marklund P, Sikström S, Nilsson LG. Reactivation in working memory: an attractor network model of free recall. PLoS One. 2013;8(8):e73776.

31. Ebbinghaus H. Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie. Leipzig, Germany: Duncker & Humbolt; 1885.

32. Murdock BB Jr. The serial position effect of free recall. J Exp Psychol. 1962;64(5):482.

33. Bjork RA, Whitten WB. Recency-sensitive retrieval processes in long-term free recall. Cogn Psychol. 1974;6(2):173–89.

34. Howard MW, Kahana MJ. Contextual variability and serial position effects in free recall. J Exp Psychol Learn Mem Cogn. 1999;25(4):923.

35. Sederberg PB, Miller JF, Howard MW, Kahana MJ. The temporal contiguity effect predicts episodic memory performance. Mem Cogn. 2010;38(6):689–99.

36. Brodie DA, Murdock BB Jr. Effect of presentation time on nominal and functional serial-position curves of free recall. J Verbal Learn Verbal Behav. 1977;16(2):185–200.

37. Brown GD, Neath I, Chater N. A temporal ratio model of memory. Psychol Rev. 2007;114(3):539.

38. Greene RL. Sources of recency effects in free recall. Psychol Bull. 1986;99(2):221.

39. Nairne JS. The loss of positional certainty in long-term memory. Psychol Sci. 1992;3(3):199–202.

40. de Almeida Valverde Zanini G, Tufik S, Andersen ML, da Silva RCM, Bueno OFA, Rodrigues CC, et al. Free recall of word lists under total sleep deprivation and after recovery sleep. Sleep. 2014;37(12):2083–95.

41. Glanzer M, Cunitz AR. Two storage mechanisms in free recall. J Verbal Learn Verbal Behav. 1966;5(4):351–60.

42. Pacheco D, Verschure PF. Long-term spatial clustering in free recall. Memory. 2017;25(6):798–806.

43. Katkov M, Romani S, Tsodyks M. Effects of long-term representations on free recall of unrelated words. Learn Mem. 2015;22(2):101–8.

44. Coutanche MN, Koch GE, Paulus JP. Influences on memory for naturalistic visual episodes: sleep, familiarity, and traits differentially affect forms of recall. Learn Mem. 2020;27(7):284–91.

45. Janiszewski C, Wyer RS Jr. Content and process priming: a review. J Consum Psychol. 2014;24(1):96–118.

46. Elgendi M, Kumar P, Barbic S, Howard N, Abbott D, Cichocki A. Subliminal priming—state of the art and future perspectives. Behav Sci. 2018;8(6):54.

47. Bermeitinger C. Priming. In: Psychology and Mental Health: Concepts, Methodologies, Tools, and Applications. Hershey, PA/USA: IGI Global; 2015. p. 42–88.

48. Nedungadi P. Recall and consumer consideration sets: influencing choice without altering brand evaluations. J Consum Res. 1990;17(3):263–76.

49. Bermeitinger C, Goelz R, Johr N, Neumann M, Ecker UK, Doerr R. The hidden persuaders break into the tired brain. J Exp Soc Psychol. 2009;45(2):320–6.

50. Mongillo G, Barak O, Tsodyks M. Synaptic theory of working memory. Science. 2008;319(5869):1543–6.

51. Gerstner W, Kistler WM. Spiking Neuron Models. 1st ed. Cambridge, UK: Cambridge University Press; 2002.

52. Graupner M, Brunel N. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. Proc Natl Acad Sci USA. 2012;109:3991–6.

53. Higgins D, Graupner M, Brunel N. Memory maintenance in synapses with calcium-based plasticity in the presence of background activity. PLoS Comput Biol. 2014;10(10):e1003834.

54. Li Y, Kulvicius T, Tetzlaff C. Induction and consolidation of calcium-based homo- and heterosynaptic potentiation and depression. PLoS One. 2016;11:e0161679.

55. Abraham WC, Jones OD, Glanzman DL. Is plasticity of synapses the mechanism of long-term memory storage? NPJ Sci Learn. 2019;4(1):1–10.

56. Lisman J, Grace AA, Duzel E. A neoHebbian framework for episodic memory: role of dopamine-dependent late LTP. Trends Neurosci. 2011;34(10):536–47.

57. Duszkiewicz AJ, McNamara CG, Takeuchi T, Genzel L. Novelty and dopaminergic modulation of memory Persistence: a tale of two systems. Trends Neurosci. 2019;42(2):102–14.

58. Pezze M, Bazz T. Dopaminergic modulation of hippocampus-dependent learning: blockade of hippocampal D1-class receptors during learning impairs 1-trial place memory at a 30-min retention delay. Neuropharmacology. 2012;63(4):710–8.

59. Lindskog M, Kim M, Wikström MA, Blackwell KT, Kotaleski JH. Transient calcium and dopamine increase PKA activity and DARPP-32 phosphorylation. PLoS Comput Biol. 2006;2(9):e119.

60. Barbour B, Brunel N, Hakim V, Nadal JP. What can we learn from synaptic weight distributions? Trends Neurosci. 2007;30(12):622–9.

61. Buzsáki G, Mizuseki K. The log-dynamic brain: how skewed distributions affect network operations. Nat Rev Neurosci. 2014;15(4):264–78.

62. Braitenberg V, Schiz A. Cortex: statistics and geometry of neuronal connectivity. 2nd ed. Berlin, Germany: Springer; 1998.

63. Sjöström PJ, Turrigiano GG, Nelson SB. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron. 2001;32:1149–64.

64. De Falco E, Ison MJ, Fried I, Quiroga RQ. Long-term coding of information and functional serial-position curves of free recall. J Verbal Learn Verbal Behav. 2007;46(6):539–53.
68. Luboeinski J. Simulation code and analysis scripts for recurrent synaptic neural networks with memory consolidation based on synaptic tagging and capture; 2022. Available from: https://doi.org/10.5281/zenodo.4429195.
69. Anderson MC, Neely JH. Interference and inhibition in memory retrieval. In: Memory. San Diego/CA, USA: Academic Press; 1996. p. 237–313.
70. Wickelgren WA. Acoustic similarity and retroactive interference in short-term memory. J Verbal Learn Verbal Behav. 1965;4(1):53–61.
71. Anderson JR. Interference: the relationship between response latency and response accuracy. J Exp Psychol Hum Learn Mem. 1981;7(5):326.
72. Preston AR, Eichenbaum H. Interplay of hippocampus and prefrontal cortex in memory. Curr Biol. 2013;23(17):R764–73.
73. Ghosh VE, Gilboa A. What is a memory schema? A historical perspective on current neuroscience literature. Neuropsychologia. 2014;53:104–14.
74. Dayan P, Abbott LF. Theoretical Neuroscience. 1st ed. Cambridge, MA, USA: MIT Press; 2001.
75. Jezek K, Henriksen EJ, Treves A, Moser EI, Moser MB. Theta-paced flickering between place-cell maps in the hippocampus. Nature. 2011;478(7368):246–9.
76. Russo E, Durstewitz D. Cell assemblies at multiple time scales with arbitrary lag constellations. eLife. 2017;6:e19428.
77. Brunel N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci. 2000;8(3):183–208.
78. McKenzie S, Frank AJ, Kinsky NR, Porter B, Rivière PD, Lohnas LJ, Polyn SM, Kahana MJ. Expanding the scope of memory engrams. PLoS Comput Biol. 2021;17(12):e1009691.
79. Abdou K, Shehata M, Choko K, Nishizono H, Matsuo M, Muramatsu SI, et al. Synapse-specific representation of the identity of overlapping memory engrams. Science. 2018;360(6394):1227–31.
80. Lohnas LJ, Polyn SM, Kahana MJ. Expanding the scope of memory search: modeling intralist and interlist effects in free recall. Psychol Rev. 2015;122(2):337.
81. Long NM, Danoff MS, Kahana MJ. Recall dynamics reveal the retrieval of emotional context. Psychon Bull Rev. 2015;22(5):1328–33.
82. Amit DJ. Modeling Brain Function: The World of Attractor Neural Networks. Cambridge, UK: Cambridge University Press; 1989.
83. McClelland JL, McNaughton BL, O’Reilly RC. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev. 1995;102(3):419.
84. Gardner-Medwin AR. Doubly modifiable synapses: a model of short and long term auto-associative memory. Proc R Soc B. 1989;238(1291):137–54.
85. Fusi S, Drew PJ, Abbott LF. Cascade models of synaptically stored memories. Neuron. 2005;45(4):599–611.
86. Päpper M, Kempter R, Leibold C. Synaptic tagging, evaluation of memories, and the distal reward problem. Learn Mem. 2011;18:58–70.
87. Boboeva V, Brasslett R, Treves A. The capacity for correlated semantic memories in the cortex. Entropy. 2018;20(11):824.
88. Russo E, Treves A. Cortical free-association dynamics: distinct phases of a latching network. Phys Rev E. 2012;85(5):051920.
89. Recanatesi S, Katkov M, Tsodyks M. Memory states and transitions between them in attractor neural networks. Neural Comput. 2017;29(10):2684–711.
90. Luboeinski J. The Role of Synaptic Tagging and Capture for Memory Dynamics in Spiking Neural Networks [Dissertation]. University of Göttingen; 2021. Available from: https://doi.org/10.5384/gediiss-463.
91. Cho K, Aggleton JP, Brown M, Bashir Z. An experimental test of the role of postsynaptic calcium levels in determining synaptic strength using perirhinal cortex of rat. J Physiol. 2001;532(2):459–66.
92. Ikegaya Y, Ishizaka Y, Matsuki N. BDNF attenuates hippocampal LTD via activation of phospholipase C: implications for a vertical shift in the frequency-response curve of synaptic plasticity. Eur J Neurosci. 2002;16(1):145–8.
93. Rashid AJ, Yan C, Mercaldo V, Hsiang HLL, Park S, Cole CJ, et al. Competition between engrams influences fear memory formation and recall. Science. 2016;353(6297):383–7.
94. Tetzlaff C, Koldzschejski C, Timme M, Tsodyks M, Wörgötter F. Synaptic scaling enables dynamically distinct short- and long-term memory formation. PLoS Comput Biol. 2013;9(10):e1003307.
95. Ren X, Coutanche MN. Sleep reduces the semantic coherence of memory recall: An application of latent semantic analysis to investigate memory reconstruction. Psychon Bull Rev. 2021;28:1336–43.
96. Y. Ikegaya Y, Ishizaka Y, Matsuki N. BDNF attenuates hippocampal LTD via activation of phospholipase C: implications for a vertical shift in the frequency-response curve of synaptic plasticity. Eur J Neurosci. 2002;16(1):145–8.
97. Tetzlaff C, Koldzschejski C, Timme M, Tsodyks M, Wörgötter F. Synaptic scaling enables dynamically distinct short- and long-term memory formation. PLoS Comput Biol. 2013;9(10):e1003307.
98. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
99. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
100. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
101. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
102. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
103. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
104. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
105. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
106. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
107. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
108. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
109. Tenenbaum JM, Hastie R, Tenenbaum JM. A method for selecting hyperparameters in machine learning. J Mach Learn Res. 2011;12:927–82.
110. Bjork RA, Bjork EL. A new theory of disuse and an old theory of stimulus fluctuation. In: From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes. vol. 2. Hillsdale/NJ, USA: L. Erlbaum; 1992. p. 35–67.

111. Roth A, van Rossum MCW. In: Computational Modeling Methods for Neuroscientists. 1st ed. Cambridge/MA, USA: MIT Press; 2009. p. 139–160.

112. Lin JW, Faber DS. Modulation of synaptic delay during synaptic plasticity. Trends Neurosci. 2002;25(9):449–55.

113. Kobayashi R, Tsubo Y, Shinomoto S. Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. Front Comput Neurosci. 2009;3(9).

114. Sajikumar S, Navakkode S, Sacktor TC, Frey JU. Synaptic tagging and cross-tagging: the role of protein kinase Mζ in maintaining long-term potentiation but not long-term depression. J Neurosci. 2005;25(24):5750–6.

115. Bear MF. A synaptic basis for memory storage in the cerebral cortex. Proc Natl Acad Sci USA. 1996;93:13453–9.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.