Systems/Circuits

Intrinsic Neural Excitability Biases Allocation and Overlap of Memory Engrams

Geoffroy Delamare,[1] Douglas Feitosa Tomé,[1,2] and Claudia Clopath[1]
[1]Department of Bioengineering, Imperial College London, London SW7 2AZ, UK and [2]Institute of Science and Technology Austria, Klosterneuburg 3400, Austria

Memories are thought to be stored in neural ensembles known as engrams that are specifically reactivated during memory recall. Recent studies have found that memory engrams of two events that happened close in time tend to overlap in the hippocampus and the amygdala, and these overlaps have been shown to support memory linking. It has been hypothesized that engram overlaps arise from the mechanisms that regulate memory allocation itself, involving neural excitability, but the exact process remains unclear. Indeed, most theoretical studies focus on synaptic plasticity and little is known about the role of intrinsic plasticity, which could be mediated by neural excitability and serve as a complementary mechanism for forming memory engrams. Here, we developed a rate-based recurrent neural network that includes both synaptic plasticity and neural excitability. We obtained structural and functional overlap of memory engrams for contexts that are presented close in time, consistent with experimental and computational studies. We then investigated the role of excitability in memory allocation at the network level and unveiled competitive mechanisms driven by inhibition. This work suggests mechanisms underlying the role of intrinsic excitability in memory allocation and linking, and yields predictions regarding the formation and the overlap of memory engrams.

Key words: **engram; excitability; memory; memory allocation; memory linking; plasticity**

Significance Statement

In the brain, memories are not formed in isolation from each other. For example, two memories of events that happened close in time tend to be linked, so that recalling one memory leads to recall of the second one. Although memories are thought to be formed by strengthening synapses among neurons, understanding memory linking requires us to consider intrinsic properties of the neurons themselves. In this study, we modeled a neural network aiming at explaining how memories are formed and linked in the brain. This model is able to reproduce experimental results and allows us to make predictions about how to link or dissociate memories.

Introduction

Neural circuits have the ability to form and retain memories that last from hours to years. In particular, pioneering anatomical studies (Scoville and Milner, 1957) have suggested that such circuits are located in the hippocampus, although they had long remained unobserved. Over the past decades, technological advances such as neural imaging and optogenetics allowed for the discovery of engram cells in multiple brain regions as the neural substrate for memory storage and retrieval (Josselyn and Tonegawa, 2020). They are defined as a subpopulation of neurons that is initially activated during presentation of a stimulus, followed by transient physical and/or chemical changes that lead to its specific reactivation during memory recall (Josselyn and Tonegawa, 2020). Engram cells have been observed in the hippocampus (Liu et al., 2012), in the amygdala (Morrison et al., 2016; Rashid et al., 2016), and the neocortex (Tonegawa et al., 2015; Kitamura et al., 2017). These studies have shed lights on the ability of neural populations to store and retrieve memories but the exact mechanisms responsible for the formation of memory engrams are not yet fully clear.

The mechanistic understanding of the formation and long-term stability of memory engrams has long been dominated by Hebbian learning (Hebb, 1949). Indeed, most computational

Received May 5, 2023; revised March 25, 2024; accepted March 25, 2024.

Author Contributions: G.D., D.F.T., and C.C. designed research; G.D., D.F.T., and C.C. performed research; G.D. analyzed data; G.D. wrote the first draft of the paper; G.D., D.F.T., and C.C. edited the paper; G.D., D.F.T., and C.C. wrote the paper.

We thank Sadra Sadeh and Inês Completo Guerreiro for helpful comments on the manuscript, Yosif Zaki and Denise J. Cai for useful feedback and members of the Clopath lab for discussion and support. This work was supported by Biotechnology and Biological Sciences Research Council (BB/N013956/1 awarded to C.C.), Wellcome Trust (200790/Z/16/Z awarded to C.C.), the Simons Foundation (564408 awarded to C.C.), and Engineering and Physical Sciences Research Council (EP/R035806/1 awarded to C.C.).

Competing Interest Statement:The authors declare no competing financial interests.

Correspondence should be addressed to Claudia Clopath at c.clopath@imperial.ac.uk.

https://doi.org/10.1523/JNEUROSCI.0846-23.2024

models have focused on synaptic mechanisms, such as long-term potentiation or depression (Bliss and Collingridge, 1993; Josselyn and Tonegawa, 2020) that have been able to provide insight into the formation and stability of neural assemblies (Litwin-Kumar and Doiron, 2014; Zenke et al., 2015). As a result, the contribution of other important mechanisms, like intrinsic excitability (Titley et al., 2017), has remained underexplored. Indeed, previous experimental works have shown that neurons with high excitability are preferentially allocated to memory engrams (Han et al., 2007; Silva et al., 2009; Zhou et al., 2009). Interestingly, learning is known to transiently increase neural excitability, reducing the afterhyperpolarization of neurons over several hours (Thompson et al., 1996; Oh et al., 2003). This transient increase is likely due to the learning-induced expression of the transcription factor cAMP Response Element-Binding Protein (CREB) (Han et al., 2007; Silva et al., 2009; Rashid et al., 2016), which is known to play a role in regulating neural excitability (Dong et al., 2006). As a result, time-varying excitability may account for overlapping neural ensembles encoding memories of events that are temporally linked (Sehgal et al., 2018), namely events spaced by a short temporal delay, as observed in the lateral amygdala (Rashid et al., 2016), the hippocampal dorsal CA1 (Cai et al., 2016; Shen et al., 2022), and the retrosplenial cortex (Sehgal et al., 2021).

Attractor networks (Amit, 1989) have been previously used to describe the recurrent network properties of overlapping memory engrams (Gastaldi et al., 2021) but without taking excitability into account. Previous theoretical works have described how the dynamics of plasticity-related proteins and excitability can lead to co-allocation of memories at the dendritic level (Kastellakis et al., 2016; Sehgal et al., 2021; Chowdhury et al., 2022). Here, we extend these computational works to focus on the role of excitability in the recurrent neural network dynamics and its link with behavioral output. Combining synaptic plasticity and activity-dependent intrinsic excitability, our model is able to explain—at the mechanistic level—experimental findings regarding overlapping neural ensembles and memory linking. Moreover, we uncover the potential mechanisms allowing neurons to compete for allocation to memory engrams, as observed experimentally. Our results suggest that the temporal linking of memory engrams arises from co-activation of different neural ensembles, mediated by the interaction of time-varying excitability and synaptic plasticity. Finally, our model makes testable predictions about how the balance among inhibition, feed-forward inputs, and excitability is crucial for determining the extent of overlap among engrams of temporally close events.

Methods

Rate model. Our rate-based model consists of a single recurrent neural network of N neurons (with firing rate r_i, $1 \le i \le N$) which receives inputs from an external region of N^{in} neurons (with firing rate r_i^{in}, $1 \le i \le N^{\text{in}}$). The weights between the input region and the network are given by the matrix W^{FF} (Fig. 1*b*). Recurrent connections are given by the weight matrix W. Inhibition is introduced as $I = I_0 + I_1 \sum_{j=1}^{N} r_j$, where I_0 sets a baseline inhibition level and I_1 scales an inhibition term proportional to the sum of the firing rates of the N neurons. Finally, excitability is added as a time-varying threshold $\epsilon_i(t)$ of the input–output function. The rate dynamics of a neuron i is therefore given by

$$\tau_r \frac{dr_i}{dt} + r_i = \text{ReLU}\left(\sum_{j=1}^{N} W_{ij} r_j + \sum_{j=1}^{N^{\text{in}}} W_{ij}^{FF} r_j^{\text{in}} - I + \epsilon_i(t)\right), \quad (1)$$

where τ_r is the decay time of the rates and ReLU is the rectified linear activation function. In Figure 5, the dashed lines correspond to the case where inhibition is reduced, i.e., I_1 is set to a lower value I_1^- (see Table of parameters).

In Figure 7*d*, noise is applied as random variable sampled at each integration step from a uniform distribution centered around 0 and of amplitude 4 (a.u.).

Weights dynamics. The feed-forward weights W^{FF} are static and define three receptive fields (RFs) that model three different contexts (Fig. 1*b*). Neurons 1–15 respond preferentially to the first context, neurons 16–30 to the second context, and neurons 31–45 to the third context. All-to-all recurrent connections W are plastic and the weights W_{ij} between each presynaptic neuron i and postsynaptic neuron j follow a Hebbian rule given by

$$\tau_W \frac{dW_{ij}}{dt} = (1 + \text{US}(t)) * \tanh\left(r_i * (r_j - r_j^0(t))\right), \quad (2)$$

where τ_W is the learning time constant and US(t) is the unconditioned stimulus (US) which is equal to US$^+$ when US is applied (synchronously with stimulation of the context) and 0 otherwise. r_j^0 is the temporal mean over a time window δ of the firing rate of neuron j, given by

$$r_j^0(t) = \frac{1}{\delta} \int_{t-\delta}^{t} r_j(t')\mathrm{d}t'. \quad (3)$$

An upper cap $W_{\max}$ and a lower cap $W_{\min}$ are applied to the recurrent weights W to prevent them from being negative or too high.

Intrinsic neural excitability. Intrinsic neural excitability follows dynamics that have been previously hypothesized to be due to the increase in the CREB transcription factor following learning (Silva et al., 2009). Each neuron's initial excitability ϵ_i^0 is sampled from a half-normal distribution of mean 0 and standard deviation 0.5. If the firing rate of a given neuron i reaches a set active threshold θ, its excitability ϵ_i moves from its initial value ϵ_i^0 to a higher value E, 2.9 s after neurons are tagged. Excitability then decays to ϵ_i^0 with a time scale τ_ϵ:

$$\tau_\epsilon \frac{d\epsilon_i}{dt} + \epsilon_i = \epsilon_i^0. \quad (4)$$

Note that we did not consider any increase in excitability following recall in all figures but Figure 7*c*. In the latter, neurons are tagged during recall of the first memory and excitability of the active neurons is increased between 300 ms and 500 ms after recall onset, before decaying to baseline.

Stimulation protocol. During training, either the first, second, or last third of the N^{in} neurons from the input region are activated, namely their firing rates are set to a fixed value r_{CS}. This activation is repeated N_{stim} times for a duration ΔT, with an interstimulus delay ΔS. During the recall period, one of the three contexts is presented for a duration ΔT. During training in all figures (except Fig. 3), shock is applied synchronously with the context, namely the value US in the learning rule (Eq. 2) is set to a non-zero value US$^+$. In Figure 3, shock is applied during the last presentation of the blue context (when specified).

During the simulations where excitability is manipulated (Figs. 4 and 5), the first N_{ϵ^+} neurons received an enhanced excitability $\epsilon^{\text{increase}}$, that is added to their initial excitability ϵ_i^0 during presentation of the first context. Then, when inhibition is applied (i^+), the first N_{i^+} neurons receive an external negative current i.

Behavioral read-out. We introduced a read-out variable of the memory strength in order to compare it with the freezing levels measured in experiments. To that end, we modeled this memory strength using an ideal observer defined as

$$F = \sum_{i \in \Omega} \int_{t_r}^{t_r + \Delta T'} r_i(t)\mathrm{d}t, \quad (5)$$

where t_r is the onset of the recall stimulation and $\Delta T' = 100$ ms is the integration time, corresponding to the temporal window during which the neuron is active. Ω denotes the ensemble of neurons belonging to the engram, namely the set of neurons crossing the active threshold θ during the recall time window.

The engram overlap (Figs. 6 and 7) is computed as the number of neurons responding to both recall of the first context and another context (either the same context, a novel context, or a context that was presented with a 6 h or 24 h delay) divided by the number of neurons responding to the first context.

Exclusion criteria. During some simulations, the firing rates of some neurons increased and reached non-realistic values. We defined a threshold of 100 Hz and decided to exclude any trials where the firing rate of any neuron reached 100 Hz at any time point. Around 10% of the trials were typically excluded.

Integration. Integration was done using Euler's method on Python. A time step of 0.5 ms was used during and 3 s after training sessions, and during and 300 ms after recall sessions. Between training and recall, a time step of 20 s was used. Firing rates of neurons below 10^{-5} were set to 0.

Table of parameters. An initial set of parameters was used in most of the figures except in Figure 3. This initial configuration was chosen to match previous experimental results in the lateral amygdala (Rashid et al., 2016). In Figure 3, a second set of parameters was used to match engrams overlap measures observed in the hippocampal dCA1 (Cai et al., 2016).

Results

Formation of a single memory engram in a recurrent network with excitability

In order to study the effect of excitability in memory allocation and linking, we built a rate-based model with feed-forward and recurrent connections, equipped with excitability and Hebbian plasticity. Excitability of each neuron i is modeled as a time-varying threshold ϵ_i of the input–output function (Methods, Eq. 1). This excitability is initially sampled from a random distribution (Methods) and changes to a higher value when the neuron's firing rate reaches a threshold θ before decaying to its initial value (Methods, Eq. 4). Feed-forward inputs are defined as a single layer divided in three subpopulations corresponding to different contexts. Feed-forward weights are set as a diagonal block structure to define three receptive fields (RFs, Methods, Fig. 1*a* and *b*), such that presenting a context increases the input current to a subpopulation of neurons in the main region. Recurrent connections are assumed to be all-to-all and plastic, according to a Hebbian learning rule (Methods, Eq. 2), and initialized at 0. We then stimulated the network by presenting different contexts (training phase, Methods).

We first observed that, after presenting the first context, the firing rates of neurons responding preferentially to this context (Fig. 1*d*, left) are above the "active" threshold θ, which we defined as the threshold above which neurons are classified as active. This was not the case for the other neurons in the network (Fig. 1*d*, left). Analyzing the recurrent weights matrix revealed that learning led to the formation of an assembly of neurons strongly connected to each other (Fig. 1*c*). The weights between neurons outside the assembly, however, have not significantly changed from their initial value equal to zero. We then sought to test the ability of the network to perform pattern completion. To this end, we stimulated the network with a partial cue and measured memory retrieval. We observed that stimulating four out of the seven neurons composing the assembly, namely those that were tagged as active during training (Methods), is enough to activate all the neurons in the assembly (Fig. 1*e*, right). This result shows that stimulating a subset of neurons of the assembly is sufficient to activate other neurons in the same assembly through strong intra-assembly connections.

Parameter	Description	Figures 1, 2, and 4–7	Figure 3
N	Number of neurons in the main region	60	
N^{in}	Number of neurons in the input layer	30	
τ_ϵ	Excitability decay time constant	24 h | 12 h	
E	Excitability increase	3.5	4
τ_r	Rate decay time constant	15 ms	
θ	Active threshold	6 Hz	4 Hz
I_0	Inhibition baseline	6	
I_1	Inhibition multiplicative factor	0.9	
I_1^-	Reduced inhibition multiplicative factor	0.88	
ΔT	Stimulation duration	40 ms	
N_{stim}	Number of stimulation per training sessions	20	15
ΔS	Interstimulus delay	150 ms	
W^{FF}_{RF}	Feed-forward weights corresponding to the RF neurons	0.3	
$W^{FF}_{non\text{-}RF}$	Feed-forward weights corresponding to the non-RF neurons	0.2	
τ_w	Hebbian rule time constant	750 ms	500 ms
W_{max}	Upper cap for plastic recurrent weights	1	
W_{min}	Lower cap for plastic recurrent weights	0	
US^+	Unconditioned stimulus strength	1	0.5
r_{CS}	Firing rate of the neurons from the input layer when activated	4 Hz	
δ	Averaging time of firing rates	15 s	
N_{ϵ^+}	Number of neurons that receive enhanced excitability	8	–
$\epsilon^{increased}$	Increase in excitability	5	–
N_{i^+}	Number of neurons that are inhibited	6	–
i	Inhibition strength	0.5	–

Intrinsic neural excitability induces overlap among memory engrams of temporally close events

Next, we investigated the effect of presenting a second context to the network either 6 h or 24 h after the first one (Methods, Fig. 2*a*), inspired by previous experiments (Rashid et al., 2016). We designed our model in such a way that, after learning, excitability of neurons taking part in the newly formed assembly is increased, before slowly decaying to their baseline level (Methods, Eq. 4). Specifically, there is a transient increase in excitability after stimulating the network by the first context (Fig. 2*b*, red triangle) and the second context (Fig. 2*b*, blue triangle). We then measured memory recall for both contexts successively (Fig. 2*a*). Presenting a second context after a 24 h delay led to the formation of a second neural assembly in the recurrent weights matrix, distinct from the first one (Fig. 2*c*, top, and *e*). This second assembly is composed of neurons that are responsive to only the second context. As in the previous section, neurons that are part of both assemblies are reactivated independently during memory recall (Fig. 2*d*, top). Interestingly, we found that if the second context is presented after a 6 h delay, some off-diagonal weights are also reinforced for neurons responding preferentially to the second context (Fig. 2*c*, bottom). This suggests that the two memories are encoded by overlapping neural representations in the case where the contexts are presented 6 h apart but not 24 h apart. Indeed, when recalling the second memory, we observed a co-activation of neurons that take part in the first assembly in the case of a 6 h delay (Fig. 2*d*, bottom). We can

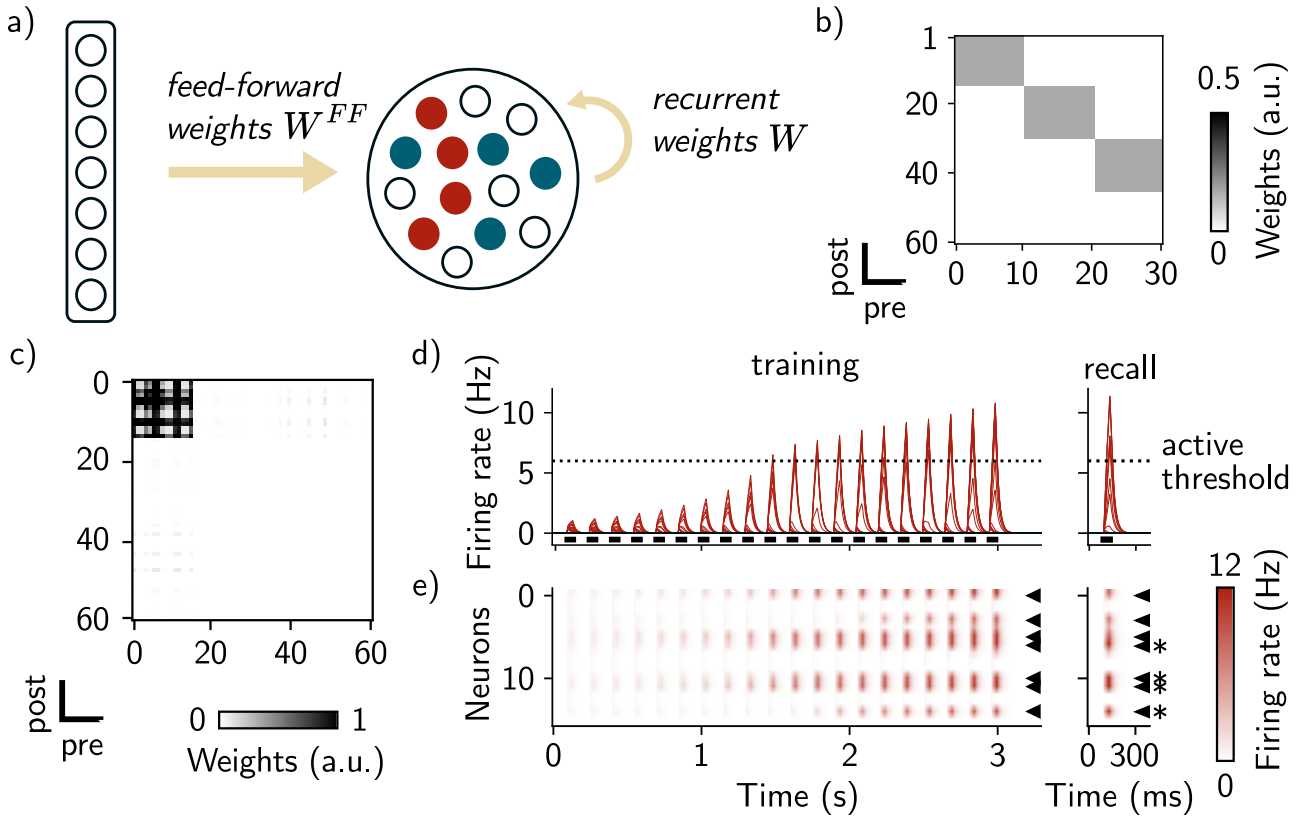

Figure 1. Encoding a single memory engram in a recurrent network equipped with intrinsic neural excitability. ***a***, Diagram of the network architecture. Neurons in the input layer (left) project to the network with feed-forward connections W^{FF} (right). The feed-forward weights are defined such that neurons in the main layer have RFs (only two are shown here, red and blue neurons). ***b***, Fixed, feed-forward synaptic weights matrix W^{FF}. RFs are defined by strong weights in a block-diagonal structure. When presenting the first context for example, the first 10 neurons of the input region are activated which in turn stimulate preferentially the first 15 neurons in the main region. ***c***, Recurrent weights matrix after training. The block structure shows a neural assembly with stronger connections between neurons responding to the red context (0–15). ***d***, Firing rates of all the neurons across time during training (left) and recall (right). Neurons responding preferentially to the context are shown in red while the other neurons are shown in black and do not respond to the stimulus. Black bars show presentations of the stimulus to the network. The dashed line is the "active" threshold, i.e., the threshold above which neurons are classified as active. ***e***, Firing rate of each neuron responding preferentially to the first context during training (left) and recall (right). Black arrows indicate seven neurons that were tagged as "active" during training and that were reactivated during recall. During the recall phase, four of these seven neurons (black stars) were stimulated.

therefore quantify the overlap between the two assemblies, namely the number of neurons that were active during recall of both contexts, and found that it is higher in the case where the events were separated by 6 h relative to 24 h (Figs. 6 and 7). This overlap seems to be independent of the engram size (Fig. 7*b*). We also observed similar results when excitability was increased following recall or when the activation threshold during training was heterogeneous (Methods, Fig. 7*c*). The presence of ongoing noise throughout the simulation did not affect the results (Methods, Fig. 7*d*). In contrast, no overlap was observed when excitability was kept fixed (Fig. 7*c*).

Linking memories at the behavioral level in a fear conditioning simulation

We then asked whether this structural overlap among memory engrams could lead to memory linking. To that end, we modeled a fear conditioning experiment (Cai et al., 2016). We measured the strength of the memory recall as an ideal observer, namely a read-out value that is proportional to the sum of the neurons' firing rate, integrated over the duration of the recall stimulus (Methods, Eq. 5). The US was introduced as a multiplicative term in the Hebbian learning rule (Methods, Eq. 2) such that the recurrent weights are preferentially increased when the US is applied.

We presented three distinct contexts to the network, separated by 7 d and 5 h (Fig. 3*a*). In order to test memory linking, we presented the last context a second time, now paired with the US (Fig. 3*a*, in blue) and measured the memory strength of each of the three contexts. We observed that the memory strength was high when presenting either the blue context, which was paired with the US, or the yellow context, which was not paired with the US but was initially separated by 5 h relative to the shocked context (Fig. 3*b*). Conversely, presenting the red context, delayed by 7 d, elicited a level of memory recall comparable to the control case when no US was applied (Fig. 3*b*). Similarly to previous results (Fig. 2), the memory ensemble associated to the yellow contexts overlaps with the one associated to the blue context, but not the red one (Fig. 3*c*–*e*). Our model was then able to show that two memories encoded close in time tend to be linked in such a way that recalling either memory can lead to a similar behavioral output, as shown experimentally (Cai et al., 2016; Shen et al., 2022).

Manipulating initial excitability biases neural allocation of memories

Given that excitability is a key mechanism for linking memory engrams, we then asked to what extent excitability could also play a role in biasing memory allocation. To that end, and inspired by previous experiments (Rashid et al., 2016), we

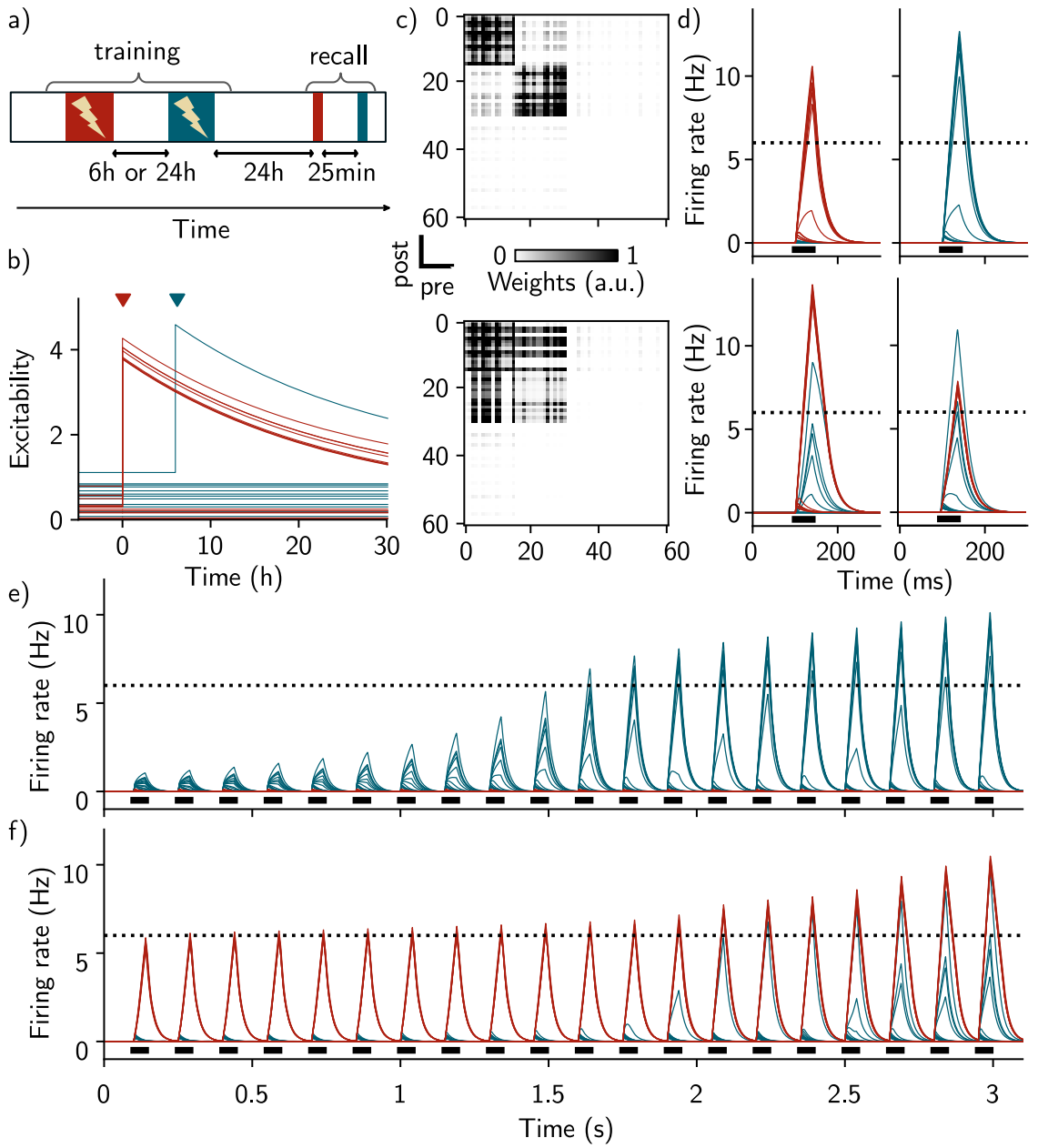

Figure 2. Training-induced increase in excitability induces overlap among memory engrams of temporally close events. ***a***, Simulation protocol for studying the effect of forming two memories, spaced by different temporal delays. During the training phase, two contexts are presented 6 h or 24 h apart. After 24 h, both contexts are recalled, separated by 25 min. ***b***, Time course of excitability for each neuron of the network. Whenever the firing rate of a neuron crosses the active threshold, its excitability moves to a higher value (red and blue triangles corresponding to training on the first and second contexts, respectively), before decreasing to their initial value on a time scale of 24 h (Methods). Red and blue traces correspond to neurons responding preferentially to the first and second contexts, respectively. ***c***, Recurrent weights matrix immediately after training, in the case where the contexts are presented 24 h apart (top) and 6 h apart (bottom). ***d***, Firing rates of individual neurons during recall of the first context (left) and the second context (right), in the case where the events are separated by 24 h (top) and 6 h (bottom). Neurons 1–15 respond preferentially to the first context (red) and neurons 16–30 respond preferentially to the second context (blue). The dashed line corresponds to the active threshold and the black bars to the stimulation. ***e*** Time course of the firing rate of all the neurons during presentation of the second context, when presented 24 h after the first context (same protocol as Fig. 2). ***f***, Same as ***e***, when the second context is presented 6 h after the first one.

increased the initial excitability of a subset of the neurons that respond preferentially to the first context (ϵ^+, Methods, Fig. 4*a*) during training of the first context. We then inhibited this subpopulation during memory recall (i^+, Methods, Fig. 4*c*) and measured the strength of memory recall. We found that inhibiting the subpopulation whose excitability was enhanced reduced the memory strength during recall relative to the control case without manipulation of excitability (Fig. 4*d* and *f*, top). This suggests that neurons with enhanced excitability are preferentially allocated to memory engrams (Fig. 4*g*, left).

Next, we presented two different contexts to the network, separated by 6 h or 24 h, and tested how increasing excitability to a subset of neurons during formation of the first memory could bias the overlap between the two memory engrams. We inhibited the subpopulation that received enhanced excitability during recall of the second context and we measured the memory strength of the second context (Fig. 4*b* and *f*, bottom). In the

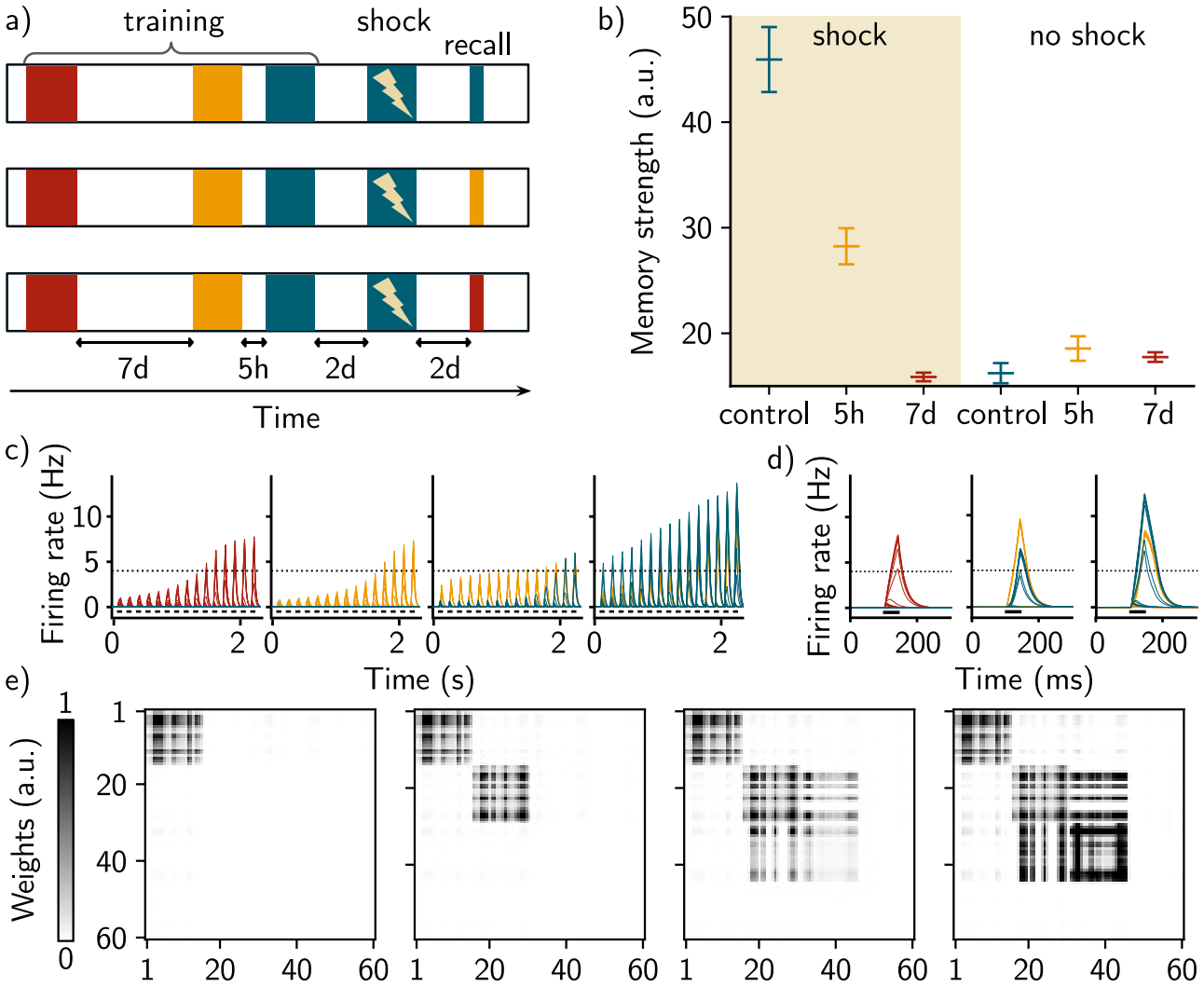

Figure 3. Memory linking in a fear conditioning simulation. ***a***, Simulation protocol of fear conditioning. Three contexts are used and are delayed by either 7 d (red and yellow) or 5 h (yellow and blue). Shock is then applied 2 d later in the blue context by pairing the context with an unconditioned stimulus (Methods). Memories are then recalled independently for each of the three contexts. ***b***, Memory strength read-out upon recalling the three memories, when shock is applied (left) or not (right). $n = 10$ simulations and data are shown as mean ± s.e.m. ***c***, Firing rates of all the neurons across time during training in the fear conditioning protocol depicted in ***a***. Each color represents a population of neurons that receives high input current when one of the contexts is presented. The first, second, and third panels correspond, respectively, to the presentation of the first, second, and third contexts, with a delay of 7 d (between the first and the second one) and 5 h (between the second and the third one). Shock is applied 2 d after the first presentation of the third context (rightmost panel). The dashed line corresponds to the active threshold and the black bars to the stimulation. ***d***, Firing rate of the neurons during recall of the first, second, and third memory, respectively, 2 d after the shock. ***e***, Corresponding recurrent weights matrices after presentation of each of the three contexts (first three panels) and after the second presentation of the third context paired with the shock (rightmost panel).

case where the events were separated by 6 h, inhibiting the subpopulation resulted in a reduction of the memory strength as compared to the case where the events were delayed by 24 h (Fig. 4*e* and *f*, bottom). Given that the subpopulation ϵ^+ is composed of neurons responding preferentially to the first context, this suggests that this subpopulation preferentially took part in the overlap between the two memory engrams (Fig. 4*g*, right).

Inhibition-induced competition among neurons crucially regulate memory allocation for temporally close events

Finally, we sought to evaluate how much neurons compete for memory allocation. To that end, we repeated the same protocol as in the previous section, but inhibiting the subpopulation (i^+) during presentation of the second context, instead of during recall (Fig. 5*a*). We observed that the formation of the second memory was impaired when the contexts were presented 6 h apart compared to the 24 h delay (Fig. 5*b*, solid lines). Note that this is the case whether or not the subpopulation (ϵ^+) is inhibited (i^+) during recall (Fig. 5*b*, solid lines).

We then hypothesized that this memory impairment was driven by inhibition. We repeated the simulation as above while reducing the amount of inhibition in the network (Methods) during presentation of the second context, as inspired by previous experiments (Rashid et al., 2016). We observed that the memory strength was less impaired by inhibition of the subpopulation (i^+) when the network inhibition was reduced (Fig. 5*b*, dashed lines). Indeed, the memory strength of the second context increased more for a delay of 6 h compared to 24 h, relative to the case with baseline inhibition (Fig. 5*b*, dashed lines).

The balance among inhibition, feed-forward inputs, and excitability is crucial for forming overlaps

Overall, we found that excitability can induce overlap between memory engrams. This overlap is dependent on the temporal

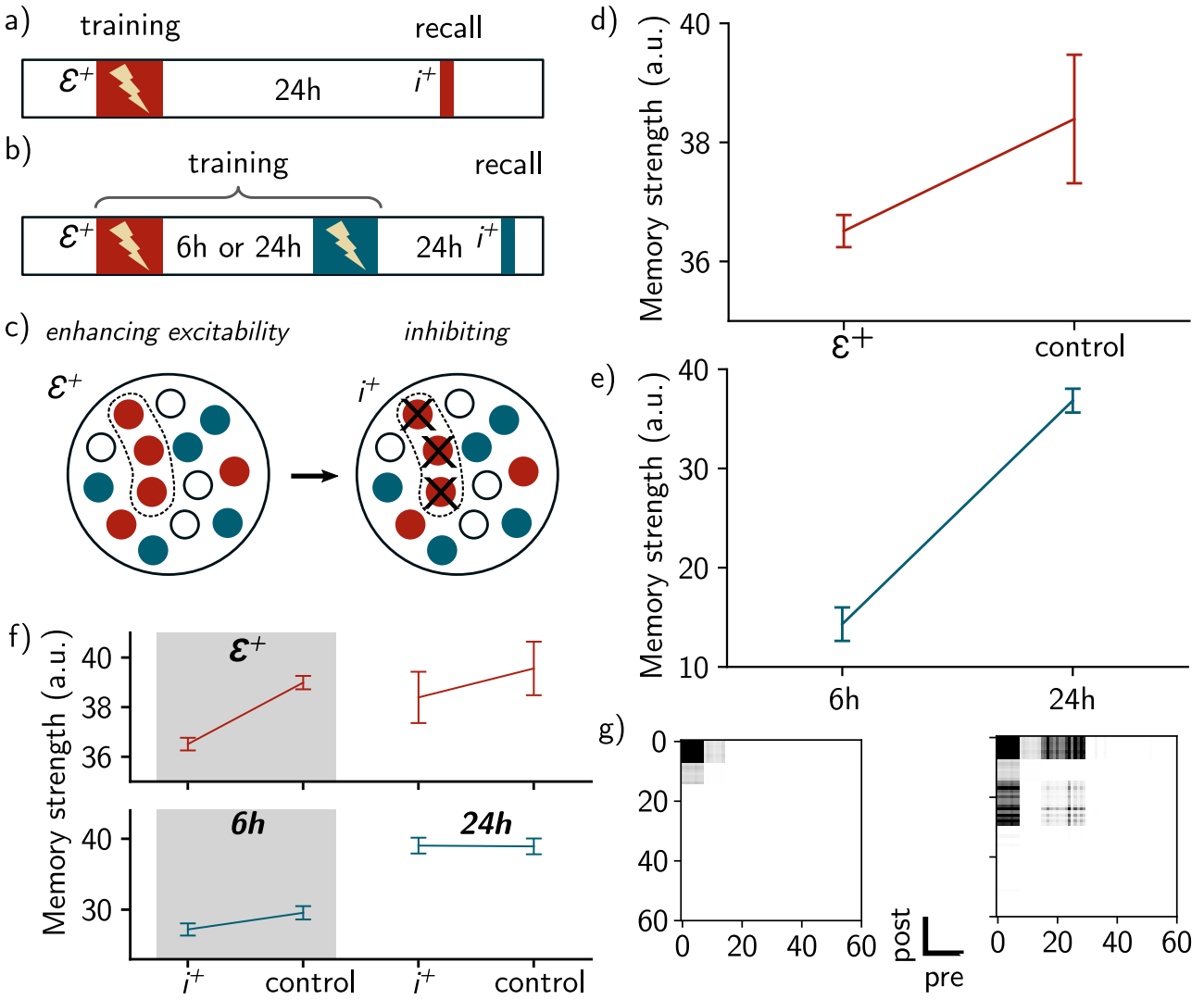

Figure 4. Neurons with enhanced excitability are preferentially allocated to memory engrams and overlapping ensembles. ***a***, Protocol for biasing memory allocation to a subpopulation of neurons. Excitability of a subset of neurons is enhanced during training (ϵ^+, top). During recall, this subpopulation is blocked (i^+, Methods). In the control case, excitability is not manipulated. ***b***, Protocol for biasing the overlap to a subset of neurons. Again, excitability is enhanced (ϵ^+) for a subset of neurons during presentation of the first context (red). Then, a second context is presented (blue) after 6 h or 24 h, and the memory strength of the second context is measured while blocking the subpopulation that received enhanced excitability (i^+). ***c***, Spatial representation of the protocol: a subset of N_{ϵ^+} neurons receives an enhanced excitability ϵ^+, that is added to their initial excitability. During inhibition, N_{i^+} neurons from this subset are inhibited, receiving a negative current i^+ (Methods). ***d***, Memory strength of the context while blocking the subset of neurons, in the case where excitability is enhanced (ϵ^+) or not (control). ***e***, Memory strength when recalling the second memory in ***b***, when the two contexts are separated by either 6 h or 24 h. For all simulations, $n = 50$ simulations and data are shown as mean ± s.e.m. ***f***, Same as ***d*** and ***e***, when the subset is either inhibited (i^+) or not (control) during recall. For each conditions, $n = 50$ simulations and data are shown as mean ± s.e.m. ***g***, Recurrent weight matrix after presentation of the first context (left) and the second context (right) for a 6 h delay.

delay between the two contexts in a manner consistent with experimental findings in the lateral amygdala (Rashid et al., 2016) and in the hippocampal dorsal CA1 (Cai et al., 2016). Our result predicts that the engram overlap arises from reactivation of the first ensemble when forming the second memory (Fig. 2*f*). Indeed, co-activation of neurons encoding the first memory (red traces) along with neurons responding preferentially to the second context (blue traces) lead to strengthening of the weights between these two ensembles, due to Hebbian plasticity (Fig. 2*c* and *f*). We varied the temporal delay between the two contexts and found that the amount of overlap decreases when this delay increases (Fig. 6*b*).

We also predict that increasing the level of inhibition I_0 leads to a decrease in the overlap between the two ensembles. Indeed, if the two events are separated by 6 h, increasing I_0 leads to a decrease in the overlap as compared to the control case (Fig. 6*c*, I_0^+). We found that decreasing the excitability decay timescale τ_ϵ also leads to a decrease in the overlap (Fig. 6*c*, τ_ϵ^-). Indeed, the excitability increase following learning needs to be above a threshold otherwise the first ensemble cannot be reactivated even if the second context is presented after 6 h (Fig. 7*j*). Similarly, the overlap among memory engrams decreases with the active threshold set during training (Fig. 7*a*).

Finally, we predict that the network can only form overlapping memory engrams if the feed-forward weights that do not form RFs ($W^{FF}_{\text{non-RF}}$) are within a defined range (Fig. 7*i*). If these weights are too low, neurons that are not preferentially activated by the second context cannot be reactivated when presenting this context after 6 h. On the other hand, if they are too high, the ensembles overlap independently of the temporal delay between

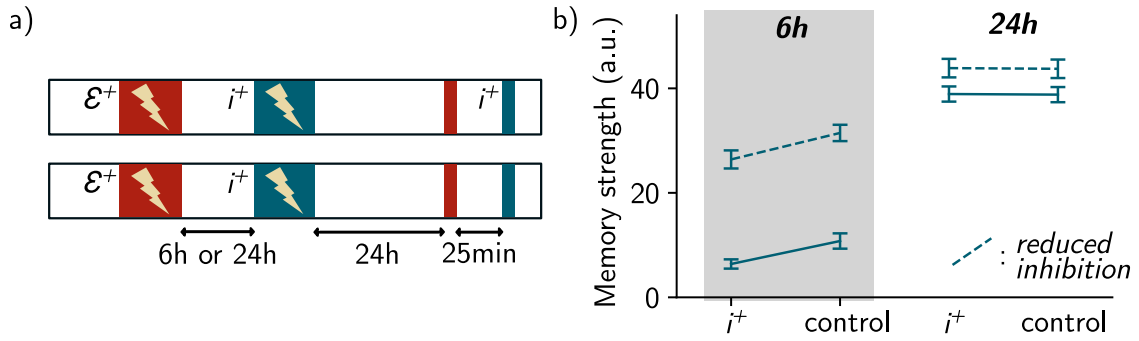

Figure 5. Neurons compete for allocation to memory engrams. ***a***, Same protocol as Figure 4*b* but with the inhibition of the subset (i^+) applied during presentation of the second context. The subset is then either inhibited (i^+) or not (control) during recall. ***b***, Fear measurement when recalling the second memory when the events are separated by either 6 h or 24 h. The dashed line corresponds to the case where global inhibition I_1 is reduced (Methods). For all simulations, $n = 50$ simulations and data are shown as mean ± s.e.m. (nine trials were excluded, Methods).

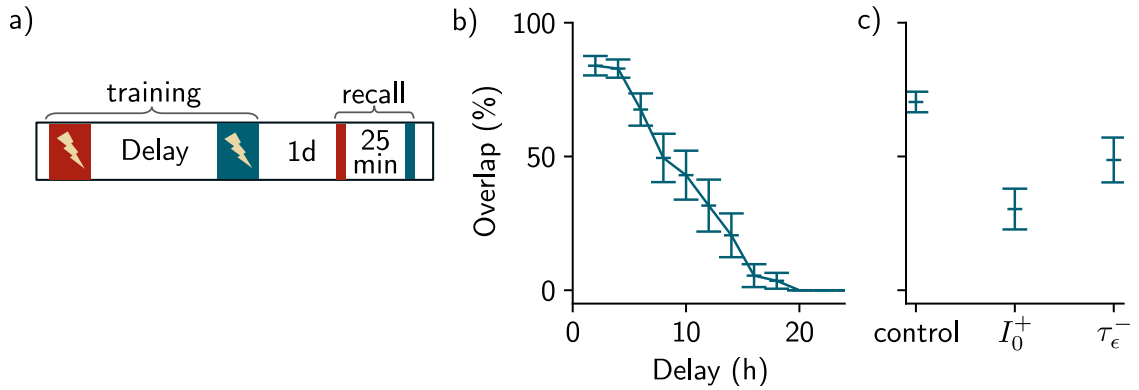

Figure 6. Predictions regarding the overlap among memory engrams. ***a***, Protocol for forming overlap among memory engrams. During training, two contexts were presented, separated by a given temporal delay. The recall protocol allows for measuring the amount of overlap between engrams associated to the first context (red) and the second context (blue). ***b***, Overlap among engrams against the temporal delay between the contexts. For each temporal delay, $n = 20$ simulations (three were excluded, Methods). Results are shown as mean ± s.e.m. ***c***, Overlap obtained for a 6 h delay in the control case, the case where inhibition was increased I_0^+ and the case where the excitability decay time was decreased τ_ϵ^-. For the control case, $n = 50$ simulations (five were excluded, Methods) and for I_0^+ and τ_ϵ^-, $n = 20$ simulations. Data are shown as mean ± s.e.m.

the two contexts. In that case, we even observed an overlap with the novel context (Fig. 7*i*, yellow line).

Discussion

Learning a single memory

We first showed that our network is able to form memories when stimulated by a feed-forward input. We attributed this formation to synaptic plasticity, independently from the dynamics of neural excitability. Indeed, the time scale of excitability is much slower than the time scale of Hebbian plasticity in our model, suggesting that the initial activation of the neurons by feed-forward inputs leads to strengthening of the synaptic weights through Hebbian learning. The structure of the neural assembly holding the formed memory is similar to previous theoretical works that have used attractor networks (Amit, 1989; Gastaldi et al., 2021). These neural assemblies are formed during the learning phase and are reactivated during memory recall. Importantly, memories can be recalled even when stimulated by a partial cue, suggesting that the neural activity is driven by recurrent connections. In particular, the structural change that leads to the formation of neural assemblies leads to the formation of a memory, as suggested in previous definitions of engram cells (Josselyn and Tonegawa, 2020).

The ability of the network to perform pattern completion is complemented by its ability to perform pattern separation. However, we did not evaluate pattern separation in our model. This would be interesting in particular because a recent study has shown that increased neural excitability in dentate gyrus improves pattern completion and separation (Pignatelli et al., 2019). The role of the recall-induced increase in excitability could explain this improvement. This could be directly tested in our model, for example by presenting a conflicting cue to the network and measuring pattern separation after recall of the memory.

In our model, the range of the firing rates of neurons during stimulations is consistent with electrophysiological recordings during fear conditioning in the amygdala (Lee et al., 2021) and in CA1 (Cohen et al., 2017). Here, we did not consider spontaneous firing rate apart from when introducing background noise in Figure 7*d*. Finally, our model does not take into account the evolution of memory engrams across time. Indeed, when the network is not stimulated, the recurrent weights are kept static. In this framework, the delay between the training phase and the recall phase has no impact on the recurrent weights.

Overlap between memory engrams of temporally close events

Here, we built a model that is able to reproduce the overlap among memory engrams of events that are temporally linked. These overlaps have been observed in the amygdala (Rashid et al., 2016), the hippocampus (Cai et al., 2016), and the retrosplenial cortex (Sehgal et al., 2021) while it has been reported

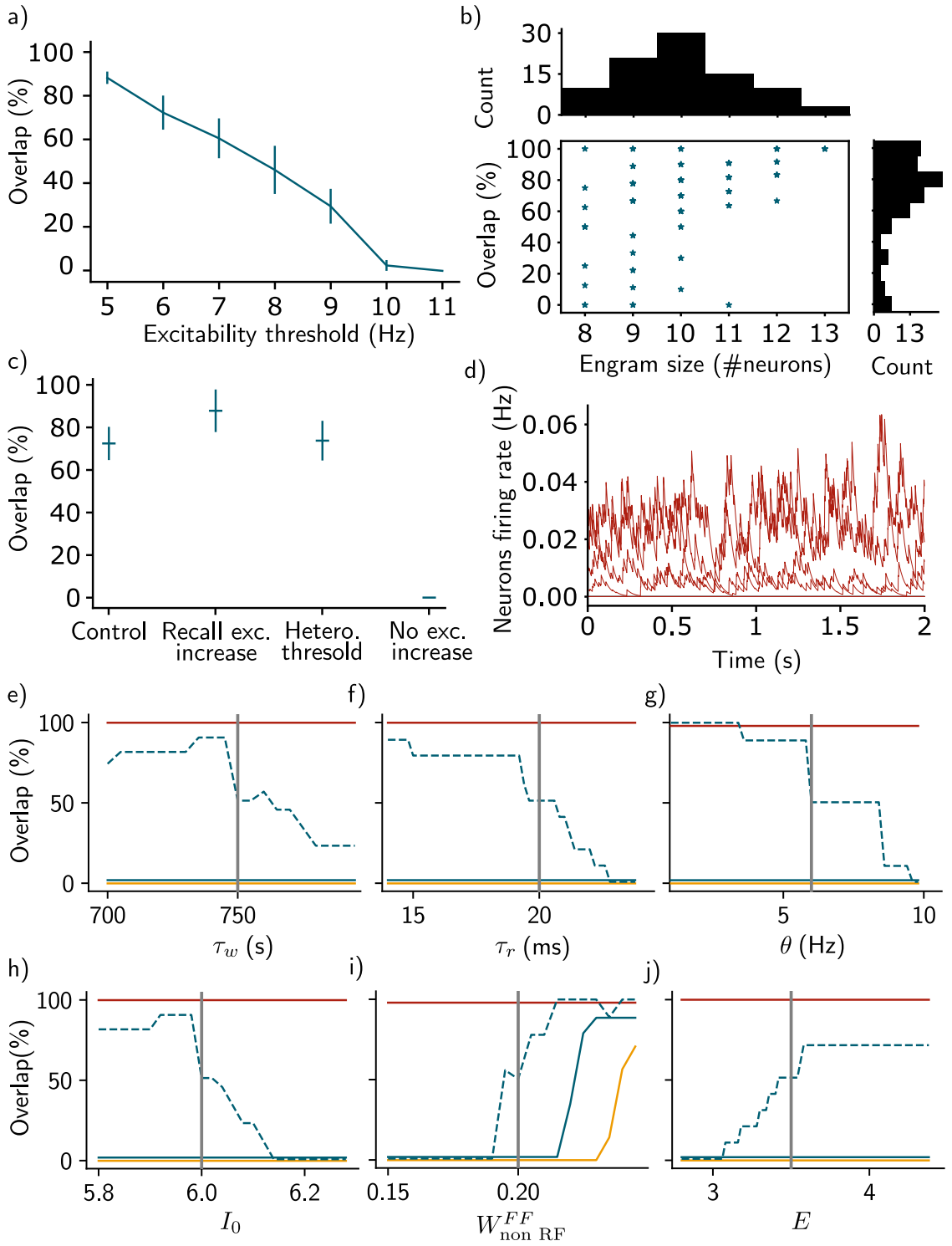

Figure 7. Further model analysis. ***a***, Overlap among two engrams when the contexts are separated by 6 h (same protocol as Fig. 2, Methods) against the activation threshold during training. Here, the activation threshold during recall remains fixed (6 Hz). $n = 10$ simulations (one was excluded, Methods) and data are shown as mean ± s.e.m. ***b***, Overlap among two engrams when the contexts are separated by 6 h against the size of the first engram. Each dot shows one simulation. The top and right histograms show the distribution of engram sizes and overlaps, respectively. $n = 100$ simulations (11 were excluded, Methods). ***c***, Overlap among two engrams when the contexts are separated by 6 h in four scenarios: (1) control case (as in Fig. 2), (2) excitability is increased following recall of the memories, (3) activation threshold during training is sample from a half-normal distribution of mean 6 Hz and standard deviation 1 Hz, and (4) excitability is left at baseline during the entire simulation. $n = 10$ simulations (one was excluded, Methods) and data are shown as mean ± s.e.m. ***d***, Neurons firing rate 1 h after learning the first memory in the presence of ongoing noise throughout the whole simulation, when two contexts are presented 6 h apart. Traces in red correspond to neurons responding preferentially to the first context, other neurons' firing rate are at baseline. ***e***, Overlap measured during the same protocol as Figure 2 as a function of the main parameters of the model (Methods, Table of parameters). The dashed blue line and the solid blue line correspond to the 6 h delay and the 24 h delay, respectively. The red line corresponds to the case where the same context is presented and the yellow to the case where a novel context is presented. $n = 50$ simulations (five were excluded, Methods) and data are shown as mean ± s.e.m. Each gray line corresponds to the parameter that has been selected for the simulations in the main figures.

that these three regions are involved in a memory consolidation process known as systems consolidation (Kitamura et al., 2017; Tonegawa et al., 2018). However, it remains unclear what information is transferred from one region to another and investigating a potential transfer of overlap between brain regions would help understand how temporal memory linking evolves over the course of systems consolidation.

Structural overlap of memory engrams induces memory linking

Here, we reproduced a fear conditioning experiment by modeling the US as a three-factor learning rule. We reasoned that an aversive event (e.g., foot shock) would increase some neuromodulators that in turn modulate plasticity. On the other hand, our measure of memory strength (averaged firing rates of the neurons belonging to the engram) is comparable to the fear readout in experimental works on rodents (freezing response), which is also a proxy for memory strength.

After showing that memories of temporally close events are structurally linked in overlapping neural ensembles, we showed that this overlapping structure can induce memory linking. In line with recent experimental studies (Cai et al., 2016; Yokose et al., 2017), we observed in our model that this overlap supports memory linking as the fear associated with one context can be transferred to another context (Fig. 3).

We also note that this memory linking is a result of the short temporal delay between contexts. Indeed, in contrast to previous studies (de Sousa et al., 2021; Gastaldi et al., 2021), the overlap between two memory engrams is independent from the conceptual relation between the contexts in question, which we did not consider here. However, it is possible that this overlap supports the formation of mnemonic structures, as they have been observed in the hippocampus for instance (Deuker et al., 2016; Barron et al., 2020). Further work could be done to investigate the importance of overlapping memory engrams for more complex cognitive processes such as inferential reasoning (Zeithamova et al., 2012; Barron et al., 2020).

Manipulating excitability biases neural allocation of memories

In our model, memory allocation is determined by two main factors. On the one hand, engrams are preferentially allocated to neurons that receive increased feed-forward inputs, as defined in the feed-forward weights (Fig. 1). On the other hand, we showed that memory allocation is also biased toward neurons with high excitability (Fig. 4), consistent with previous studies (Zhou et al., 2009; Rashid et al., 2016). However, for the sake of simplicity, we did not explicitly probe the relative importance of the feed-forward weights compared to intrinsic excitability. To that end, it would be necessary to introduce some variability in the structure of the RFs, and subsequently investigate how neuronal memory allocation is impacted by feed-forward inputs versus excitability dynamics.

Here, we considered that excitability is mainly governed by the transcription factor CREB. We modeled its dynamics by increasing excitability instantaneously after learning and then allowing it to decay over a time scale of a few hours, as motivated by several experimental studies (Moyer et al., 1996; Thompson et al., 1996; Oh et al., 2003; Kitagawa et al., 2017). Although the results might be similar, it is important to note that these dynamics are conceptualized and that other mechanisms that are not considered here are also known to regulate neural excitability. For instance, internalization of Kir2.1 channel increases neural excitability during memory recall (Pignatelli et al., 2019) while the expression of the C-C chemokine receptor type 5 is known to decrease excitability (Zhou et al., 2016; Shen et al., 2022). Adult-born neurons are also known to be more excitable than their counterparts (Silva et al., 2009). Overall, the exact relationship between firing rate and change in excitability is not known. For this reason, we abstracted the dynamics of excitability by only considering the increase of excitability following learning and a slow decay over a timescale of several hours. Finally, memory allocation may also be influenced by other mechanisms beyond the scope of the present study such as synaptic tagging and spine clustering (Rogerson et al., 2014; Kastellakis et al., 2016).

We also showed that artificially increasing excitability in an ensemble of neurons could also bias co-allocation of this ensemble to further memories as shown in previous experimental findings (Rashid et al., 2016). This result arose naturally in our model because neurons with enhanced excitability are preferentially allocated to the first memory, and will then overlap with the second engram (Fig. 4*g*).

Role of inhibition in memory allocation and linking

Finally, we showed that neurons can compete for memory allocation and that the outcome of this competition is determined by both the initial excitability of neurons and the amount of inhibition in the network. We first showed that blocking neurons which received enhanced excitability during presentation of the first context impaired learning of a second context presented shortly after, suggesting that these neurons have a competitive advantage over the others for memory allocation. Second, we showed that reducing inhibition restored the ability of the network to learn the second memory, suggesting that competition is driven by inhibition.

In our model, neurons with a higher initial excitability are favored for memory allocation and inhibit the remaining neurons, preventing them from taking part in a memory engram. This process has been previously shown experimentally (Han et al., 2007; Rashid et al., 2016) and this study provides a computational model that sheds light on the underlying competitive mechanism. Finally, we use a homogeneous global inhibition model, but further studies could explore the effect of populations of different inhibitory cell types on engram overlap and memory linking.

Conclusion

In summary, we have built a recurrent neural network model that can reproduce the experimentally observed neuronal overlap between temporally linked memory engrams by combining both synaptic plasticity and neural excitability. Our results suggest that engram overlaps are crucially determined by the balance among inhibition, feed-forward inputs, and excitability.

References

Amit DJ (1989) *Modeling brain function: the world of attractor neural networks*. Cambridge: Cambridge University Press. https://www.cambridge.org/core/books/modeling-brain-function/2EA95FDABF616D187220A6B9596091B7

Barron HC, et al. (2020) Neuronal computation underlying inferential reasoning in humans and mice. Cell 183:228–243.e21.

Bliss TVP, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31–39.

Cai DJ, et al. (2016) A shared neural ensemble links distinct contextual memories encoded close in time. Nature 534:115–118.

Chowdhury A, Luchetti A, Fernandes G, Filho DA, Kastellakis G, Tzilivaki A, Ramirez EM, Tran MY, Poirazi P, Silva AJ (2022) A locus coeruleus-

dorsal CA1 dopaminergic circuit modulates memory linking. Neuron 110:3374–3388.e8.
Cohen JD, Bolstad M, Lee AK (2017) Experience-dependent shaping of hippocampal CA1 intracellular activity in novel and familiar environments. eLife 6:e23040.
de Sousa AF, Chowdhury A, Silva AJ (2021) Dimensions and mechanisms of memory organization. Neuron 109:2649–2662.
Deuker L, Bellmund JL, Navarro Schröder T, Doeller CF (2016) An event map of memory space in the hippocampus. eLife 5:e16534.
Dong Y, Green T, Saal D, Marie H, Neve R, Nestler EJ, Malenka RC (2006) CREB modulates excitability of nucleus accumbens neurons. Nat Neurosci 9:475–477.
Gastaldi C, Schwalger T, Falco ED, Quiroga RQ, Gerstner W (2021) When shared concept cells support associations: theory of overlapping memory engrams. PLoS Comput Biol 17:e1009691.
Han J-H, Kushner SA, Yiu AP, Cole CJ, Matynia A, Brown RA, Neve RL, Guzowski JF, Silva AJ, Josselyn SA (2007) Neuronal competition and selection during memory formation. Science 316:457–460.
Hebb DO (1949) *The organization of behavior: a neuropsychological theory*. Mahwah, NJ: L. Erlbaum Associates.
Josselyn SA, Tonegawa S (2020) Memory engrams: recalling the past and imagining the future. Science 367:eaaw4325.
Kastellakis G, Silva AJ, Poirazi P (2016) Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. Cell Rep 17:1491–1504.
Kitagawa H, Sugo N, Morimatsu M, Arai Y, Yanagida T, Yamamoto N (2017) Activity-dependent dynamics of the transcription factor of cAMP-response element binding protein in cortical neurons revealed by single-molecule imaging. J Neurosci 37:1–10.
Kitamura T, Ogawa SK, Roy DS, Okuyama T, Morrissey MD, Smith LM, Redondo RL, Tonegawa S (2017) Engrams and circuits crucial for systems consolidation of a memory. Science 356:73–78.
Lee J, An B, Choi S (2021) Longitudinal recordings of single units in the basal amygdala during fear conditioning and extinction. Sci Rep 11:11177.
Litwin-Kumar A, Doiron B (2014) Formation and maintenance of neuronal assemblies through synaptic plasticity. Nat Commun 5:5319.
Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, Tonegawa S (2012) Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature 484:381–385.
Morrison DJ, Rashid AJ, Yiu AP, Yan C, Frankland PW, Josselyn SA (2016) Parvalbumin interneurons constrain the size of the lateral amygdala engram. Neurobiol Learn Mem 135:91–99.
Moyer JR, Thompson LT, Disterhoft JF (1996) Trace eyeblink conditioning increases CA1 excitability in a transient and learning-specific manner. J Neurosci 16:5536–5546.
Oh MM, Kuo AG, Wu WW, Sametsky EA, Disterhoft JF (2003) Watermaze learning enhances excitability of CA1 pyramidal neurons. J Neurosci 90: 2171–2179.
Pignatelli M, Ryan TJ, Roy DS, Lovett C, Smith LM, Muralidhar S, Tonegawa S (2019) Engram cell excitability state determines the efficacy of memory retrieval. Neuron 101:274–284.e5.
Rashid AJ, et al. (2016) Competition between engrams influences fear memory formation and recall. Science 353:383–387.
Rogerson T, Cai DJ, Frank A, Sano Y, Shobe J, Lopez-Aranda MF, Silva AJ (2014) Synaptic tagging during memory allocation. Nat Rev Neurosci 15:157–169.
Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. J Neurol Neurosurg Psychiatry 20:11–21.
Sehgal M, et al. (2021) Co-allocation to overlapping dendritic branches in the retrosplenial cortex integrates memories across time. Pages: 2021.10.28.466343. Section: New Results.
Sehgal M, Zhou M, Lavi A, Huang S, Zhou Y, Silva A (2018) Memory allocation mechanisms underlie memory linking across time. Neurobiol Learn Mem 153:21–25.
Shen Y, et al. (2022) CCR5 closes the temporal window for memory linking. Nature 606:146–152.
Silva AJ, Zhou Y, Rogerson T, Shobe J, Balaji J (2009) Molecular and cellular approaches to memory allocation in neural circuits. Science 326:391–395.
Thompson LT, Moyer JR, Disterhoft JF (1996) Transient changes in excitability of rabbit CA3 neurons with a time course appropriate to support memory consolidation. J Neurophysiol 76:1836–1849.
Titley HK, Brunel N, Hansel C (2017) Toward a neurocentric view of learning. Neuron 95:19–32.
Tonegawa S, Liu X, Ramirez S, Redondo R (2015) Memory engram cells have come of age. Neuron 87:918–931.
Tonegawa S, Morrissey MD, Kitamura T (2018) The role of engram cells in the systems consolidation of memory. Nat Rev Neurosci 19:485–498.
Yokose J, et al. (2017) Overlapping memory trace indispensable for linking, but not recalling, individual memories. Science 355:398–403.
Zeithamova D, Schlichting ML, Preston AR (2012) The hippocampus and inferential reasoning: building memories to navigate future decisions. Front Hum Neurosci 6:70.
Zenke F, Agnes EJ, Gerstner W (2015) Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nat Commun 6:6922.
Zhou M, et al. (2016) CCR5 is a suppressor for cortical plasticity and hippocampal learning and memory. eLife 5:e20985.
Zhou Y, Won J, Karlsson MG, Zhou M, Rogerson T, Balaji J, Neve R, Poirazi P, Silva AJ (2009) CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nat Neurosci 12:1438–1443.