PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Homeostatic control of synaptic rewiring in recurrent networks induces the formation of stable memory engrams

Júlia V. Gallinaro[1,2☯], **Nebojša Gašparović**[1☯], **Stefan Rotter**[1*]

1 Bernstein Center Freiburg & Faculty of Biology, University of Freiburg, Freiburg im Breisgau, Germany,
2 Bioengineering Department, Imperial College London, London, United Kingdom

☯ These authors contributed equally to this work.
* stefan.rotter@bio.uni-freiburg.de

Check for updates

OPEN ACCESS

Citation: Gallinaro JV, Gašparović N, Rotter S (2022) Homeostatic control of synaptic rewiring in recurrent networks induces the formation of stable memory engrams. PLoS Comput Biol 18(2): e1009836. https://doi.org/10.1371/journal.pcbi.1009836

Editor: Mark C. W. van Rossum, University of Nottingham, UNITED KINGDOM

Received: May 16, 2021

Accepted: January 14, 2022

Published: February 10, 2022

Copyright: © 2022 Gallinaro et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All essential scripts for simulation and analysis are provided on the repository https://github.com/juliavg/engram_structural_plasticity.

Funding: Supported by Erasmus Mundus / EuroSPIN (EU; JVG, NG), Deutsche Forschungsgemeinschaft (DFG; grant EXC 1086; NG, SR) and Carl Zeiss Foundation (SR). The HPC facilities are funded by the State of Baden-Württemberg through bwHPC and DFG grant INST 39/963-1 FUGG (SR). The funders had no role in

Abstract

Brain networks store new memories using functional and structural synaptic plasticity. Memory formation is generally attributed to Hebbian plasticity, while homeostatic plasticity is thought to have an ancillary role in stabilizing network dynamics. Here we report that homeostatic plasticity alone can also lead to the formation of stable memories. We analyze this phenomenon using a new theory of network remodeling, combined with numerical simulations of recurrent spiking neural networks that exhibit structural plasticity based on firing rate homeostasis. These networks are able to store repeatedly presented patterns and recall them upon the presentation of incomplete cues. Storage is fast, governed by the homeostatic drift. In contrast, forgetting is slow, driven by a diffusion process. Joint stimulation of neurons induces the growth of associative connections between them, leading to the formation of memory engrams. These memories are stored in a distributed fashion throughout connectivity matrix, and individual synaptic connections have only a small influence. Although memory-specific connections are increased in number, the total number of inputs and outputs of neurons undergo only small changes during stimulation. We find that homeostatic structural plasticity induces a specific type of "silent memories", different from conventional attractor states.

Author summary

Memories are thought to be stored in groups of strongly connected neurons, or engrams. Much effort has been put into understanding engrams, but currently there is no definitive consensus about how they are formed. Hebbian plasticity was proposed to underlie engram formation. Hebbian plasticity occurs when synaptic weights are strengthened between pairs of neurons with correlated activity. Hebbian plasticity, therefore, is thought to rely on a mechanism that can detect correlation between neurons. However, the increased synaptic weight implies increased correlation, which creates a positive feedback loop that can lead to runaway growth. Avoiding such an unfavorable condition would require regulatory mechanisms that are much faster than the ones which have been

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

observed in the brain. Here we show that a structural plasticity rule that is based on firing rate homeostasis can also lead to the formation of engrams. In this case, engram formation does not rely on a mechanism that traces correlation. Instead, stronger connectivity between correlated neurons emerges as a network effect, based on self-organized rewiring. Our work proposes a different possibility how engrams could be created. This new perspective could be instrumental for better understanding the process of memory formation.

Introduction

Memories are thought to be stored in the brain using cell assemblies that emerge through coordinated synaptic plasticity [1]. Cell assemblies with strong enough recurrent connections lead to bistable firing rates, which allows a network to encode memories as dynamic attractor states [2, 3]. If strong excitatory recurrent connections are counteracted by inhibitory plasticity, "silent" memories are formed [4–6]. In principle, assemblies can be generated by strengthening already existing synapses [2, 3], but potentially also by increasing connectivity among neurons. It has been shown that attractor networks can emerge through the creation of such neuronal clusters [7].

The creation of clusters through changes in connectivity between cells would require synaptic rewiring, or structural plasticity. Structural plasticity has been frequently reported in different areas of the brain, and sprouting and pruning of synaptic contacts was found to be often activity-dependent [8–10]. Sustained turnover of synapses, however, poses a severe challenge to the idea of memories being stored in synaptic connections [11]. Interestingly, recent theoretical work has shown that stable assemblies can be maintained despite ongoing synaptic rewiring [12, 13].

The formation of neuronal assemblies, or clusters, is traditionally attributed to Hebbian plasticity, driven by the correlation between pre- and postsynaptic neuronal activity on a certain time scale. For a typical Hebbian rule, a positive correlation in activity causes an increase in synaptic weight, which in turn increases the correlation between neuronal firing. This positive feedback cycle can result in unbounded growth, runaway activity and an essential dynamic instability of the network, if additional regulatory mechanisms are lacking. In fact, neuronal networks of the brain appear to employ homeostatic control mechanisms that regulate neuronal activity [14], and actively stabilize the firing rate of individual neurons at specific target levels [15, 16]. However, even though homeostatic mechanisms have been reported in experiments to operate on a range of different time scales, they seem to be too slow to trap the instabilities caused by Hebbian learning rules [17]. All things considered, it remains to be elucidated, what are the exact roles of Hebbian and homeostatic plasticity, and how these different processes interact to form cell assemblies in a robust and stable way [18].

Concerning the interplay between Hebbian and homeostatic plasticity, we have recently demonstrated by simulations that homeostatic structural plasticity alone can lead to the formation of assemblies of strongly interconnected neurons [19]. Moreover, we found that varying the strength of the stimulation and the fraction of stimulated neurons in combination with repetitive protocols can lead to even stronger assemblies [20]. In both papers, we used a structural plasticity model based on firing rate homeostasis, which had been used before to study synaptic rewiring linked with neurogenesis [21, 22], and the role of structural plasticity after focal stroke [23, 24] and after retinal lesion [25]. This model has also been used to study the emergence of criticality in developing networks [26] and other topological aspects of plastic

networks [27]. Similarly to models with inhibitory plasticity [4–6], the memories formed in networks of this type represent a form of silent memory that is not in any obvious way reflected by neuronal activity.

The long-standing discussion about memory engrams in the brain has been revived recently. Researchers were able to identify and manipulate engrams [28], and to allocate memories to specific neurons during classical conditioning tasks [29]. These authors have also emphasized that an engram is not yet a memory, but merely the physical substrate of a potential memory in the brain [28]. Similar to the idea of a memory trace, it should provide the necessary conditions for a retrievable memory to emerge. Normally, the process of engram formation is thought to involve the strengthening of already existing synaptic connections. Here, we propose that new engrams could also be formed by a special form of synaptic clustering with increased synaptic connectivity among participating neurons, but the total number of incoming and outgoing synapses remaining approximately constant.

To demonstrate the feasibility of the idea, we performed numerical simulations of a classical conditioning task in a recurrent network with structural plasticity based on firing rate homeostasis. We were able to show that the cell assemblies formed share all characteristics of a memory engram. We further explored the properties of the formed engrams and developed a mean-field theory to explain the mechanisms of memory formation with homeostatic structural plasticity. We showed that these networks are able to effectively store repeatedly presented patterns. The formed engrams implement a special type of silent memory, which normally exists in a quiescent state and can be successfully retrieved using incomplete cues.

Results

Formation of memory engrams by homeostatic structural plasticity

We simulated a classical conditioning paradigm using a recurrent network. The network was composed of excitatory and inhibitory leaky integrate-and-fire neurons, and the excitatory-to-excitatory connections were subject to structural plasticity regulated by firing rate homeostasis. All neurons in the network received a constant background input in the form of independent Poisson spike trains. Three different non-overlapping subsets of neurons were sampled randomly from the network. The various stimuli considered here were conceived as increased external input to one of the specific ensembles, or combinations thereof. As the stimuli were arranged exactly as in behavioral experiments, we also adopted their terminology "unconditioned stimulus" (US) and "conditioned stimulus" (C1 and C2). The unconditioned response (UR) was conceived as the activity of a single readout neuron, which received input from the ensemble of excitatory neurons associated with the US (Fig 1A, top).

Fig 1B illustrates the protocol of the conditioning experiment simulated here. During a baseline period, the engrams representing US, C1 and C2 were stimulated once, one after the other. The stimulation consisted of increasing the rate of the background input by 40% for a period of 2*mathrms*. In this phase, the activity of the US ensemble was high only upon direct stimulation (Fig 1B, middle). The baseline period was followed by an encoding period, in which the C1 engram was always stimulated together with the US engram, while the C2 engram was always stimulated in isolation. Simultaneous stimulation of neurons in a recurrent network with homeostatic structural plasticity can lead to the formation of reinforced ensembles [19], which are strengthened by repetitive stimulation [20]. This is also what happened here: After the encoding period, each of the three neuronal ensembles had increased within-ensemble connectivity, as compared to baseline. Memory traces, or engrams, were formed (Fig 1D). Moreover, the US and C1 engrams also had higher bidirectional across-ensemble connectivity, representing an association between their corresponding memories.

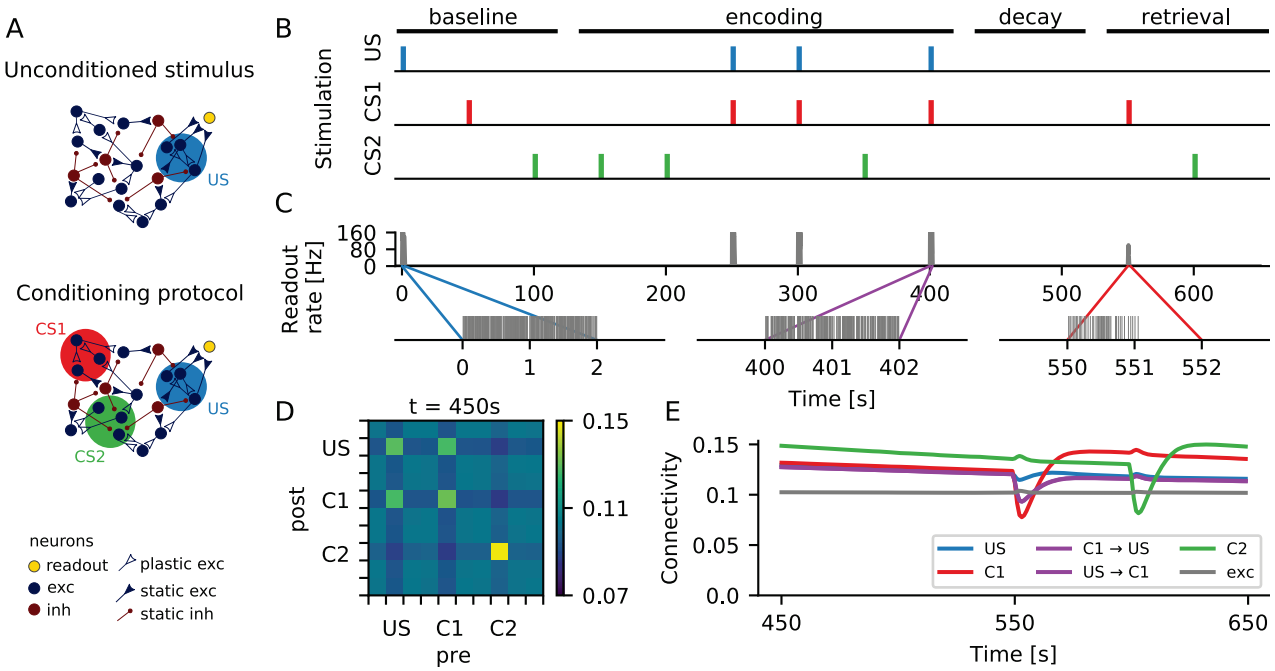


Fig 1. Formation of memory engrams in a neuronal network with homeostatic structural plasticity. (A) In a classical conditioning scenario, an unconditioned stimulus (US) was represented by a group of neurons that were connected to a readout neuron (yellow) via static synapses. The readout neuron spiked whenever neuronal activity representing the US was above a certain threshold (top). During a conditioning protocol, two other groups of neurons (CS1 and CS2) were chosen to represent two different conditioned stimuli (bottom). (B) During stimulation, the external input to a specific group of neurons was increased. The color marks indicate when each specific group was being stimulated. During the "encoding" phase, CS1 was always stimulated together with US, and CS2 was always stimulated alone. The time axis matches that of panel C (top). (C) Firing rate (top) and spike train (bottom) of the readout neuron. Before the paired stimulation ("baseline"), the readout neuron responded strongly only upon direct stimulation of the neuronal ensemble corresponding to the US. After the paired stimulation ("retrieval"), however, a presentation of C1 alone triggered a strong response of the readout neuron. This was not the case for a presentation of C2 alone. (D) Coarse grained connectivity matrix. Neurons are divided into 10 groups of 100 neurons each, shown is the average connectivity within each group. After encoding, the connectivity matrix indicates that engrams were formed, and we found enhanced connectivity within all three ensembles as a consequence of repeated stimulation. Bidirectional inter-connectivity across different engrams, however, is only observed for the pair C1-US that experienced paired stimulation. (E) Average connectivity as a function of time. The connectivity dynamics shows that engram identity was strengthened with each stimulus presentation, and that engrams decayed during unspecific external stimulation.

https://doi.org/10.1371/journal.pcbi.1009836.g001

Between encoding and retrieval, memory traces remained in a dormant state. Due to the homeostatic nature of network remodeling, the ongoing activity after encoding was very similar to the activity before encoding, but specific rewiring of input and output connections led to the formation of structural engrams. It turns out that these "silent memories" are quite persistent, as "forgetting" is much slower than "learning" them. In later sections, we will present a detailed analysis of this phenomenon. Any silent memory can be retrieved with a cue. In our case, this is a presentation of the conditioned stimulus. Stimulation of C1 alone, but not of C2 alone, triggered a conditioned response (Fig 1C) that was similar to the unconditioned response. Inevitably, stimulation of C1 and C2 during recall briefly destabilized the corresponding cell assemblies (seen as a drop in connectivity in Fig 1E), as homeostatic plasticity was still ongoing. The corresponding engrams then went through a reconsolidation period, during which the within-assembly connectivity grew even higher than before retrieval (Fig 1E, red and green). As a consequence, stored memories got stronger with each recall. Interestingly, as in our case the retrieval involved stimulation of either C1 or C2 alone, the connectivity between the US and C1 engrams decreased a bit after the recall (Fig 1E, purple).

Memories and associations were formed by changes in synaptic wiring, triggered by neuronal activity during the encoding period. They persisted in a dormant state and could be

reactivated by a retrieval cue that reflected the activity experienced during encoding. This setting exactly characterizes a memory engram [28]. In the following sections, we will further characterize the process of formation of a single engram. We will also explore the nature and stability of the formed engram in more detail.

Engrams represent silent memories, not attractors

Learned engrams have a subtle influence on network activity. For a demonstration, we first grew a network under the influence of homeostatic structural plasticity (see Section Grown networks). All neurons received a baseline stimulation in the form of independent Poisson spike trains. We then randomly selected an ensemble E_1 of excitatory neurons and stimulated it repeatedly. Each stimulation cycle was comprised of a period of 150s increased input to E_1 and another 150 s relaxation period with no extra input. After 8 such stimulation cycles, the within-engram connectivity had increased to $\bar{C}_{E_1E_1} \approx 0.21$. At this point, the ongoing activity of the network exhibited no apparent difference to the activity before engram encoding (Fig 2A). Due to the homeostatic nature of structural plasticity, neurons fired on average at their target rate, even though massive rewiring had led to higher within-engram connectivity. Looking closer, however, revealed a conspicuous change in the second-order properties of neuronal ensemble activity. We quantified this phenomenon using the overlap m^{μ}. This quantity reflects

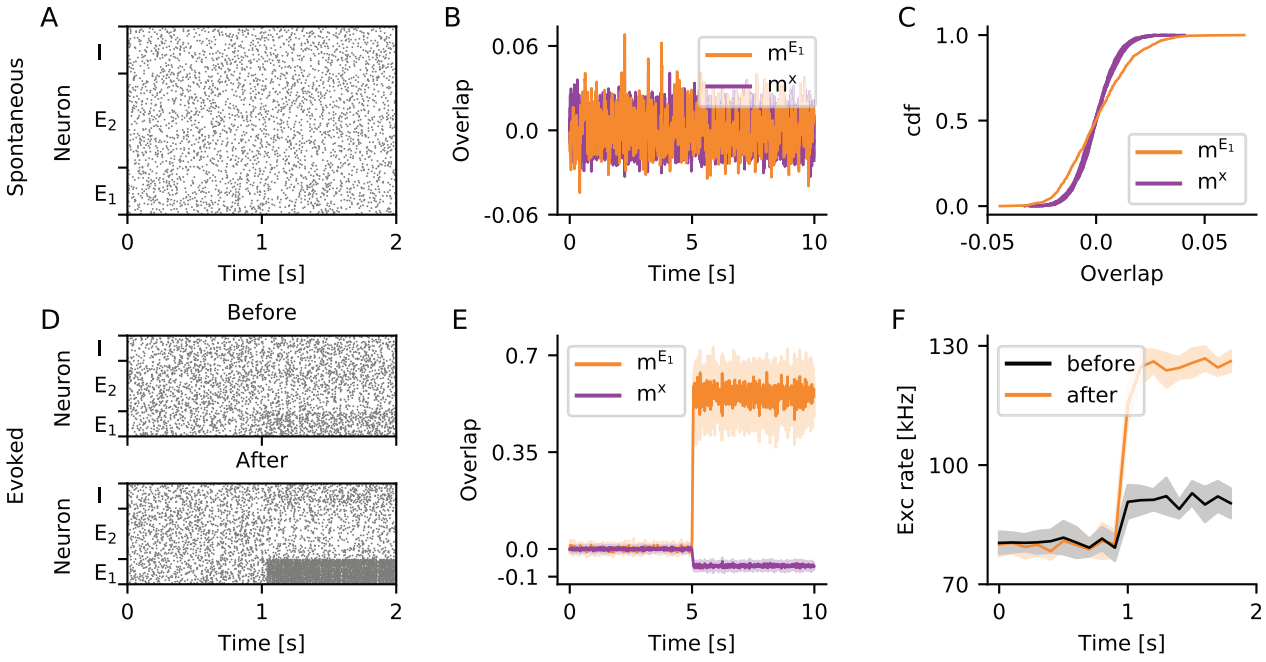


Fig 2. Silent memory based on structural engrams. (A–C) The ongoing activity of neurons belonging to the engram E_1 can hardly be distinguished from the activity of the rest of the network. (A) Raster plot showing the spontaneous activity of 50 neurons randomly selected from E_1, 100 neurons randomly selected from E_2 but not belonging to E_1, and 50 neurons randomly selected from the pool I of inhibitory neurons. (B) Overlap of ongoing network activity with the learned engram E_1 (m^{E_1}, orange), and separately for 10 different random ensembles x disjoint with E_1 (m^x, purple). (C) Cumulative distribution of m^{μ} shown in (B). (D–F) The activity evoked upon stimulation of E_1 is higher, if the within-engram connectivity is large enough ($\bar{C}_{E_1E_1} > 0.1$) as a consequence of learning. (D) Same as A for evoked activity, the stimulation starts at $t = 1$ s. The neurons belonging to engram E_1 are stimulated before (top, $\bar{C}_{E_1E_1} \approx 0.1$) and after (bottom, $\bar{C}_{E_1E_1} \approx 0.21$) engram encoding. (E) Overlap with the learned engram (m^{E_1}, orange) and with random ensembles (m^x, purple) during specific stimulation of engram E_1. (F) Population rate of all excitatory neurons during stimulation of E_1 before (black) and after (orange) engram encoding. (E, F) Solid line and shading depict mean and standard deviation across 10 independent simulation runs, respectively. In all panels, the bin size for calculating overlaps is 10 ms, and the bin size for calculating population rates is 100 ms.

https://doi.org/10.1371/journal.pcbi.1009836.g002

Table 1. List of symbols.

Symbol	Description
$\phi(t)$	Calcium trace
τ_{Ca}	Calcium time constant
$S(t)$	Spike train
$r(t)$	Instantaneous firing rate
ν	Target rate
$a(t)$	Number of axonal elements
$d(t)$	Number of dendritic elements
β_d	Dendritic growth parameter
β_a	Axonal growth parameter
$C_{ij}(t)$	Number of synaptic connections from neuron j (presynaptic) to neuron i (postsynaptic)
$\bar{C}_{ij}(t)$	Expected value of number of synaptic connections $C_{ij}(t)$
E	Excitatory neurons
E_1	Stimulated excitatory neurons
E_2	Non-stimulated excitatory neurons
I	Inhibitory neurons
τ_{drift}	Effective time constant of encoding
$\tau_{diffusion}$	Effective time constant of forgetting
$m^x(t)$	Overlap of network activity with pattern x
τ_{rate}	Relaxation time of rate dynamics

https://doi.org/10.1371/journal.pcbi.1009836.t001

the similarity of the neuron-by-neuron activity in a given time bin with a specified pattern, for example a pattern with all stimulated neurons being active and the remaining ones being silent (see Section Overlap measure for a more detailed explanation of the concept). Fig 2B depicts the time-dependent (bin by bin) overlap of ongoing network activity with the engram E_1 (m^{E_1}, see Table 1 for a complete list of symbols). It also shows the overlap with 10 different random ensembles x (m^x), which are of the same size as E_1 but have no neurons in common with it. The variance of m^{E_1} is slightly larger than that of m^x (Fig 2C). This indicates that the increased connectivity also increased the tendency of neurons belonging to the same learned engram to synchronize their activity, in comparison to other pairs of neurons. An increase in pairwise correlations within the engram, however, leads to increased fluctuations of the population activity [30], which also affects population measures such as the overlap used here.

During specific stimulation, the differences between the evoked activity of learned engrams and random ensembles were more pronounced. Fig 2D shows raster plots of network activity during stimulation of E_1 before and after the engram was encoded. The high recurrent connectivity within the E_1 assembly after encoding amplified the effect of stimulation, leading to much higher firing rates of E_1 neurons. This effect could even be seen in the population activity of all excitatory neurons in the network (Fig 2F). During stimulation, the increase in firing rate of engram neurons was accompanied by a suppression of activity of all other excitatory neurons not belonging to the engram. This was what underlied the conspicuous decrease in the overlap m^x with random ensembles x during stimulation (Fig 2E).

How does the evoked response of an engram depend on the connectivity within? To answer this question, we looked into evoked activity at different points in time during stimulation. The within-engram connectivity increased with every stimulation cycle (Fig 3A), and so did the population activity of excitatory neurons during stimulation (Fig 3B). The in-degree of excitatory neurons, in contrast, was kept at a fixed level by the homeostatic controller, even after engram encoding (Fig 4C). This behavior was well captured by a simple mean-field firing

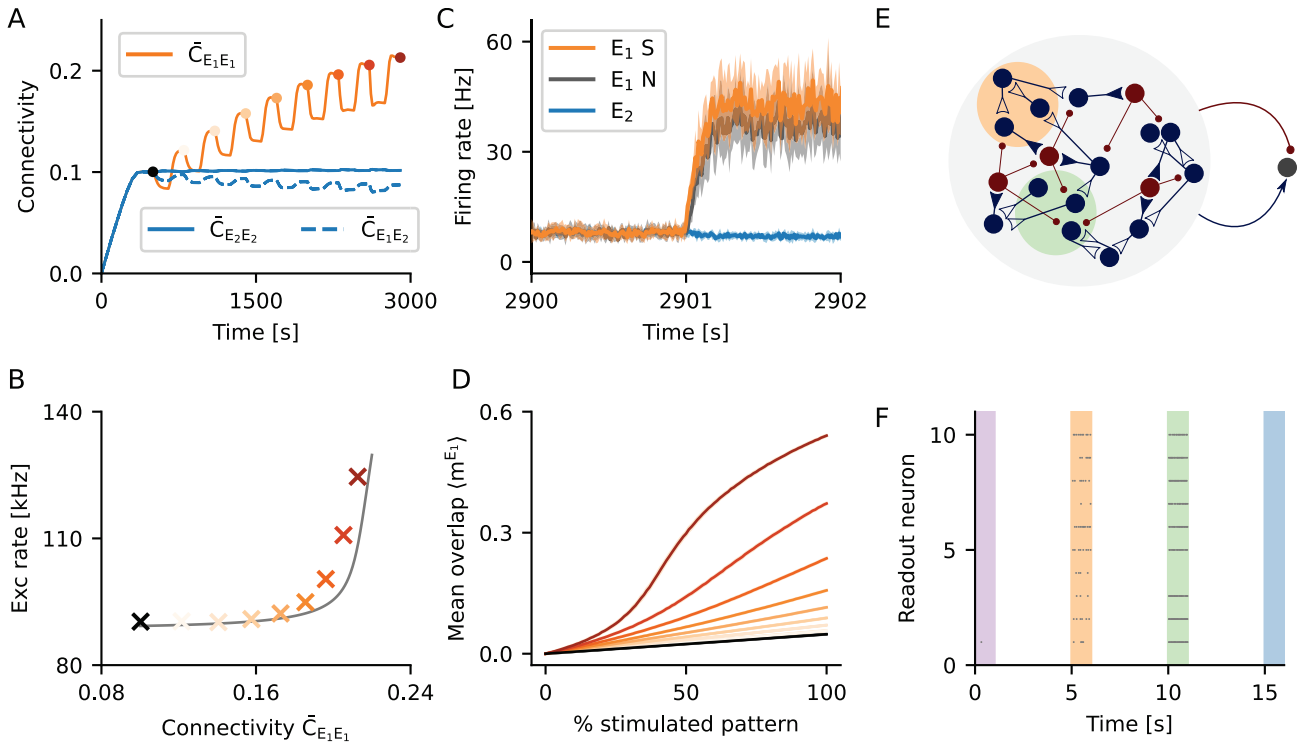


Fig 3. Evoked activity depended on the strength of a memory. (A) Starting from a random network grown under the influence of unstructured stimulation (black dot), we repeatedly stimulated the same ensemble of excitatory neurons E_1 to eventually form an engram. Multiple stimulation cycles increased the recurrent connectivity within the engram. (B) Population activity of all excitatory neurons upon stimulation of E_1, for different levels of engram connectivity $\bar{C}_{E_1E_1}$. Crosses depict the population rate observed in a simulation. Colors indicate engram connectivity $\bar{C}_{E_1E_1}$, matching the colors used in panels (A) and (D). The grey line outlines the expectation from a simple mean-field theory. (C) Firing rate of E_1 engram neurons upon stimulation of 50% its neurons. Shown is the mean firing rate of the stimulated engram neurons (E_1S, orange), the mean firing rate of the non-stimulated engram neurons (E_1N, grey), and the mean firing rate of excitatory neurons not belonging to the engram (E_2, blue). Solid line and shading depict mean and standard deviation for 10 independent simulation runs, respectively. (D) Time-averaged overlap $\langle m^{E_1} \rangle$, for different fractions of E_1 being stimulated. The recurrent nature of memory engrams enabled them to perform pattern completion. The degree of pattern completion depended monotonically on engram strength. (E) Two engrams (orange and green) were encoded in a network. Both engrams had a different strength with regard to their within-engram connectivity (green stronger than orange). A simple readout neuron received input from a random sample comprising 9% of all excitatory and 9% of all inhibitory neurons in the network. (F) Raster plot for the activity of 10 different readout neurons during the stimulation of learned engrams and random ensembles, respectively. Readout neurons were active when an encoded engram was stimulated (orange and green), and they generally responded with higher firing rates for stronger engrams (green). The activity of a readout neuron was low in absence of a stimulus (white), or upon stimulation of a random ensemble of neurons (purple and blue).

https://doi.org/10.1371/journal.pcbi.1009836.g003

rate model (grey line in Fig 3B), in which the within-engram connectivity was varied and all the remaining excitatory connections were adjusted to maintain a fixed in-degree of excitatory neurons.

We were interested to know whether engrams performed pattern completion, and how this feature depended on the connectivity within. Following the formation of the engram E_1 through 8 stimulation cycles (Fig 3A), we tested pattern completion by stimulating only 50% of the neurons in E_1 (E_1S), while the remaining neurons in E_1 (E_1N) were not directly stimulated. Fig 3C shows that, as expected, stimulation led to an increase in the activity of the directly stimulated neurons (E_1S, orange). It also shows, however, a very similar increase of the activity of the remaining non-stimulated neurons (E_1N, grey), indicating pattern completion. In order to quantitatively assess pattern completion at different levels of the connectivity within, we measured how the overlap of network activity within the engram, m^{E_1}, depended on partial stimulation. For an unstructured random network, m^{E_1} increased at a certain rate with the fraction of stimulated neurons (Fig 3C, black line). We speak of "pattern completion",

if m^{E_1} was systematically larger than in an unstructured random network. Fig 3C demonstrates clearly that the degree of pattern completion associated with a specific engram increased monotonically with the within-engram connectivity.

The evoked activity in learned engrams and in random ensembles of the same size is different. This feature can be taken as a marker for the existence of a stored memory. To demonstrate the potential of this idea, we employed a simple readout neuron for this task (Fig 3D). This neuron had the same properties as any other neuron in the network, and it received input from a random sample comprising a certain fraction (here 9%) of all excitatory and the same fraction of all inhibitory neurons in the network. We encoded two engrams in the same network, one being slightly stronger ($\bar{C}_{E_1E_1} \approx 0.19$, green) than the other one ($\bar{C}_{E_2E_2} \approx 0.18$, orange). We recorded the firing of the readout neuron during spontaneous activity, during specific stimulation of the engrams, and during stimulation of random ensembles of the same size. Fig 3E shows a raster plot of the activity of 10 different readout neurons, each of them sampling a different random subset of the network. With the parameters chosen here, the activity of a readout neuron was generally very low, except when a learned memory engram was stimulated. Due to the gradual increase in population activity with memory strength (Fig 3B), readout neurons responded with higher rates upon the stimulation of stronger engrams (green).

Homeostatic structural plasticity enables memories based on neuronal ensembles with increased within-ensemble connectivity, or engrams. Memories are acquired quickly and can persist for a long time. Moreover, the specific network configuration considered here admits a gradual response to the stimulation of an engram reflecting the strength of the memory. As we will show later, the engram connectivity lies on a line attractor, which turns into a slow manifold if fluctuations are taken into consideration. This configuration allows the network to simultaneously learn to recognize a stimulus ("Does the current stimulation correspond to a known memory?") and to assess its confidence of the recognition ("How strong is the memory trace of this pattern?"). Such behavior would be absolutely impossible in a system that relies on bistable firing rates (attractors) to define engrams. Details of our analysis will be explained later in Section Fluctuation-driven decay of engrams.

The mechanism of engram formation

We have shown how homeostatic structural plasticity created and maintained memory engrams, and we have then further elucidated the mechanisms underlying this process. We considered a minimal stimulation protocol [19] to study the encoding process for a single engram E_1. We performed numerical simulations and developed a dynamical network theory to explain the emergence of associative (Hebbian) properties. In Fig 4, the results of numerical simulations are plotted together with the results of our theoretical analysis (see Section Mean-field approximation of population dynamics). Upon stimulation, the firing rate followed the typical homeostatic dynamics [31]. In the initial phase, the network stabilized at the target rate (Fig 4A). Upon external stimulation, it transiently responded with a higher firing rate. With a certain delay, the rate was down-regulated to the set-point. When the stimulus was turned off, the network transiently responded with a lower firing rate, which was eventually up-regulated to the set-point again.

Firing rate homeostasis relied on the intracellular calcium concentration $\phi_i(t)$ of each neuron i (Fig 4D), which can be considered as a proxy for its firing rate. In our simulations, it was obtained as a low-pass filtered version of the spike train $S_i(t)$ of the neuron

$$\tau_{\mathrm{Ca}}\dot{\phi}_i(t) = -\phi_i(t) + S_i(t), \tag{1}$$

with time constant τ_{Ca}. Each excitatory neuron i used its own calcium trace $\phi_i(t)$ to control its

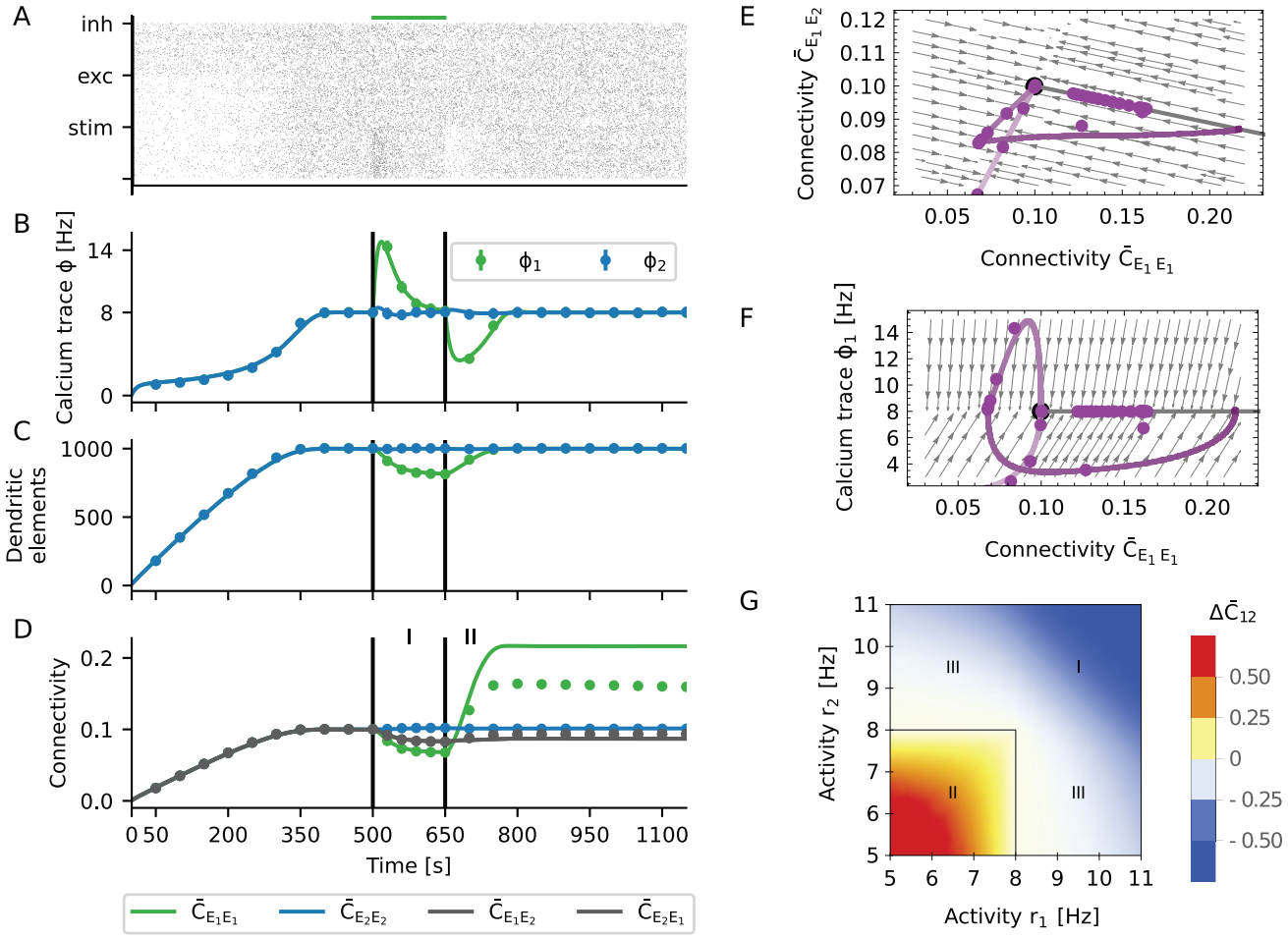


Fig 4. Hebbian properties emerge through the interaction of selective input and homeostatic control. (A) The activity of the neuronal network was subject to homeostatic control. For increased external input, it transiently responded with a higher firing rate. With a certain delay, the rate was down-regulated to the imposed set-point. When the stimulus was turned off, the network transiently responded with a lower firing rate, which was eventually up-regulated to the set-point again. The activity was generally characterized by irregular and asynchronous spike trains. (B) It was assumed that the intracellular calcium concentration followed the spiking dynamics, according to a first-order low-pass characteristic. Dots correspond to numerical simulations of the system, and solid lines reflect theoretical predictions from a mean-field model of dynamic network remodeling. (C) Dendritic elements (building blocks of synapses) were generated until an in-degree of $K_{in} = 1000$ was reached. This number slightly decreased during specific stimulation, but then recovered after the stimulus was removed. (D) Synaptic connectivity closely followed the dynamics of dendritic elements until the recovery phase, when the recurrent connectivity within the stimulated group E_1 overshooted. (A–D) Black vertical lines indicate beginning and end of the stimulation. (E, F) Phase space representation of the activity. The purple lines are projections of the full, high-dimensional dynamics to different two-dimensional subspaces: (E) within-engram connectivity vs. across-ensemble connectivity and (F) within-engram connectivity vs. engram calcium trace. The dynamic flow was represented by the gray arrows. The steady state of the plastic network was characterized by a line attractor (thick gray line), defined by a fixed total in-degree and out-degree. The ensemble of stimulated neurons formed a stable engram, and the strength of the engram was encoded by its position on the line attractor. (G) The effective "instantaneous" learning rule for the expected connectivity between a pair of neurons is homeostatic in nature. It could also be viewed as an "inverse covariance rule" with baseline at the set-point of the homeostatic controller. The emerging Hebbian properties results from the more long-term combinatorial properties of rewiring across the whole network.

https://doi.org/10.1371/journal.pcbi.1009836.g004

number of synaptic elements. Deviations of the instantaneous firing rate (calcium concentration) $\phi_i(t)$ from the target rate ν_i triggered either creation or deletion of elements according to

$$\beta_d \dot{d}_i(t) = \nu_i - \phi_i(t), \qquad \beta_a \dot{a}_i(t) = \nu_i - \phi_i(t), \tag{2}$$

where $a_i(t)$ and $d_i(t)$ are the number of axonal and dendritic elements, respectively. The parameters β_a and β_d are the associated growth parameters (see Section Plasticity model for

more details). These equations are describing homeostatic control in our model. When activity is larger than the set-point, excitatory synapses are deleted. This tends to reduce activity. When activity is lower than the set-point, synaptic elements are created and new excitatory synapses are formed. This tends to increase activity. The number of elements are represented by continuous variables. A floor operation $\lfloor x \rfloor$ is used to create the integers required for simulations.

During the initial growth phase, the number of elements increased to values corresponding to an in-degree $K^{\text{in}} = \epsilon N$, which was the number of excitatory inputs to the neuron that are necessary to sustain firing at the target rate (Fig 4D). Upon stimulation (during the learning phase), the number of connections was down-regulated due to the transiently increased firing rate of neurons. After the stimulus was turned off, the activity returned to its set-point. During the growing and learning phases, connectivity closely followed the dynamics of synaptic elements (Fig 4D), and connectivity was proportional to the number of available synaptic elements (Fig 4D). After removal of the stimulus, however, in the reconsolidation phase, the recurrent connectivity within the stimulated group E_1 was found to overshoot (Fig 4D). Then the average connectivity in the network returned to baseline (Fig 4C). While recurrent connectivity $\bar{C}_{E_1E_1}$ of the engram E_1 increased, both the connectivity to the rest of network $\bar{C}_{E_2E_1}$ and from rest of the network $\bar{C}_{E_1E_2}$ decreased, keeping the mean input to all neurons fixed. This indicates that although the network was globally subject to homeostatic control, local changes effectively exhibit associative features, as already pointed out in [19]. Our theoretical predictions generally match the simulations very well (Fig 4), with the exception that it predicted a larger overshoot. This discrepancy will be resolved in Section Fluctuation-driven decay of engrams.

Deriving a theoretical framework of network remodeling (see Section Mathematical re-formulation of the algorithm) for the algorithm suggested by [25] posed a great challenge due to the large number of variables of both continuous (firing rates, calcium trace) and discrete (spike times, number of elements, connectivity, rewiring step) nature. The dimensionality of the system was effectively reduced by using a mean-field approach, which conveniently aggregated discrete counting variables into continuous averages (see Section Time-continuous limit and Section Mean-field approximation of population dynamics for more details of derivation).

Synaptic elements are accounted for by the number of free axonal $a^+(t)$ and free dendritic $d^+(t)$ elements, while the number of deleted elements is denoted by $a^-(t)$ and $d^-(t)$, respectively. Free axonal elements are paired with free dendritic elements in a completely random fashion to form synapses. The deletion of dendritic or axonal elements in neuron i automatically induces the deletion of incoming or outgoing synapses of that neuron, respectively. We employed a stochastic differential equation (Section Time-continuous limit) to describe the time evolution of connectivity $\bar{C}_{ij}(t)$ from neuron j to neuron i

$$\frac{dC_{ij}(t)}{dt} = \underbrace{\frac{\rho'^{+}_{di}(t)\rho'^{+}_{aj}(t)}{\rho(t)} - C_{ij}(t)\left(\frac{\rho^{-}_{d_i}(t)}{K^{in}_i(t)} + \frac{\rho^{-}_{a_j}(t)}{K^{out}_j(t)}\right)}_{\text{deterministic drift}} + \underbrace{\frac{dW_{\text{spike noise}}}{dt} + \frac{dW_{\text{structural noise}}}{dt}}_{\text{stochastic noise}}. \tag{3}$$

In this equation, $\rho^{\pm}_{a_i}(t)$, $\rho^{\pm}_{d_i}(t)$ is the rate of creation/deletion of the axonal and dendritic elements of neuron i, respectively, and $\rho(t)$ is the rate of creation of elements in the whole network. Note that ρ'^{+} is a corrected version of ρ^{+} (see Section Time-continuous limit for details).

The stochastic process described by Eq 3 decomposes into a deterministic drift process and a diffusive noise process. The noise process has two sources. The first source derives from the stochastic nature of the spike trains, and the second source is linked with the stochastic nature of axon-dendrite bonding. In this section, we ignore the noise and discuss only the deterministic part of the equation. This is equivalent to reducing the spiking dynamics to a firing rate model and, at the same time, treat connectivity in terms of its expectation values.

Stable steady-state solutions of the system described by Eq 3 represent a hyperplane in the space of connectivity (see Eq 17 in Section Line attractor of the deterministic system). These steady-state solutions are (random) network configurations with a fixed in-degree K_i^{in} and out-degree K_i^{out} such that $C^*_{i,j} \propto K_i^{\text{in}} K_j^{\text{out}}$. We will show that these states are attractive in Section Network stability and constraints on growth parameters and Section Linear stability analysis. The hyperplane solution in our mean-field approach also generalizes to population variables (Section Mean-field approximation of population dynamics and Section Line attractor of the deterministic system). In the case of only one engram stored in the network, there is the stimulated population and the "rest" of the network. In this case, the hyperplane from Eq 17 reduces to a line attractor of Eq 16 (Fig 4E and 4F, gray line). Memories are stored in the network in the following way: When a group of neurons is stimulated, the network diverges from the line attractor and takes a different path back during reconsolidation. The new position on the line encodes the strength of the memory. Stronger memories increase their connectivity within $\bar{C}_{E_1E_1}$ at expense of other connections. Furthermore, as the attractor is a skewed hyperplane in the space of connectivity, the memory is distributed across the whole neural network, and not only in recurrent connections among stimulated neurons in E_1. As a reflection of this, other connectivity parameters ($\bar{C}_{E_2E_2}$, $\bar{C}_{E_1E_2}$, $\bar{C}_{E_2E_1}$, cf. Fig 4D) are also slightly changed.

To understand why changes in recurrent connectivity $\bar{C}_{E_1E_1}$ are associative, we note that the creation part of Eq 3 is actually a product $\rho^+_{a_i}(t)$ and $\rho^+_{d_i}(t)$, similar to a pre-post pair in a typical Hebbian rule. The main difference to a classical Hebbian rule is that only neurons firing below their target rate are creating new synapses. The effective rule is depicted in Fig 4G. Only neurons with free axonal or dendritic elements, respectively, can form new synapses, and those neurons are mostly the ones with low firing rates. The deletion part of Eq 3 depends linearly on $\rho^-_{a_i}(t)$ and $\rho^-_{d_i}(t)$, reducing to a simple multiplicative homeostasis. Upon excitatory stimulation, the homeostatic part is dominant and the number of synaptic elements decreases. Once the stimulus is terminated, the Hebbian part takes over, inducing a post-stimulation overshoot in connectivity. This leads to a peculiar bimodal dynamics of first decreasing connectivity and then overshooting, an important signature of this rule. We summarize this process in an effective rule

$$\Delta c_{ij} \quad \propto \quad \lambda \underbrace{\Delta I_i \Delta I_j}_{\text{effective Hebbian}} - \gamma_a \underbrace{c_{ij}(t)\Delta I_i}_{\text{dendritic homeostasis}} - \gamma_d \underbrace{c_{ij}(t)\Delta I_j}_{\text{axonal homeostasis}}, \tag{4}$$

where ΔI_i is the input perturbation of neuron i. The term $\Delta I_i \Delta I_j$ is explicitly Hebbian with regard to input perturbations. Eq 4 only holds, however, if the stimulus is presented for a long enough time such that the calcium concentration tracks the change in activity and connectivity drops.

Fluctuation-driven decay of engrams

The qualitative aspects of memory formation have been explored in Section The mechanism of engram formation. Now we investigated the process of memory maintenance. A noticeable discrepancy between theory and numerical simulations was pointed out in Fig 4D. The

overshoot is exaggerated and memories last forever. We found that this discrepancy was resolved when we take the spiking nature of neurons into account (Section Spiking noise).

Neurons use discrete spike trains $S_i(t) = \sum_k \delta(t - t_k^i)$ for signaling, and we conceived them here as stochastic point processes. We found that Gamma processes (Section Spiking noise) could reproduce the first two moments of the spike train statistics of the simulated networks with sufficient precision. The homeostatic controller in our model uses the trace of the calcium concentration $\phi_i(t)$ as a proxy for the actual firing rate of the neuron. As the calcium trace $\phi_i(t)$ is just a filtered version of the stochastic spike train $S_i(t)$, it is a stochastic process in its own respect. In Fig 5A we showed the stationary distribution of the time-dependent calcium concentration $\phi_i(t)$. Apparently, a filtered Gamma process (purple line) provided a better fit to the simulated data than a filtered Poisson process (red line). The reason is that Gamma processes have an extra degree of freedom to match the irregularity of spike trains (coefficient of variation, here CV ≈ 0.7) as compared to Poisson processes (always CV = 1).

The homeostatic controller (Eq 2) strived to stabilize $\phi_i(t)$ at a fixed target value ν, but $\phi_i(t)$ fluctuates (Fig 5A) due to the random nature of the spiking. These fluctuations resulted in some degree of random creation and deletion of connections. Our theory (Section Spiking noise) reflected this aspect by an effective rule (Eq 12), which was obtained by averaging Eq 3

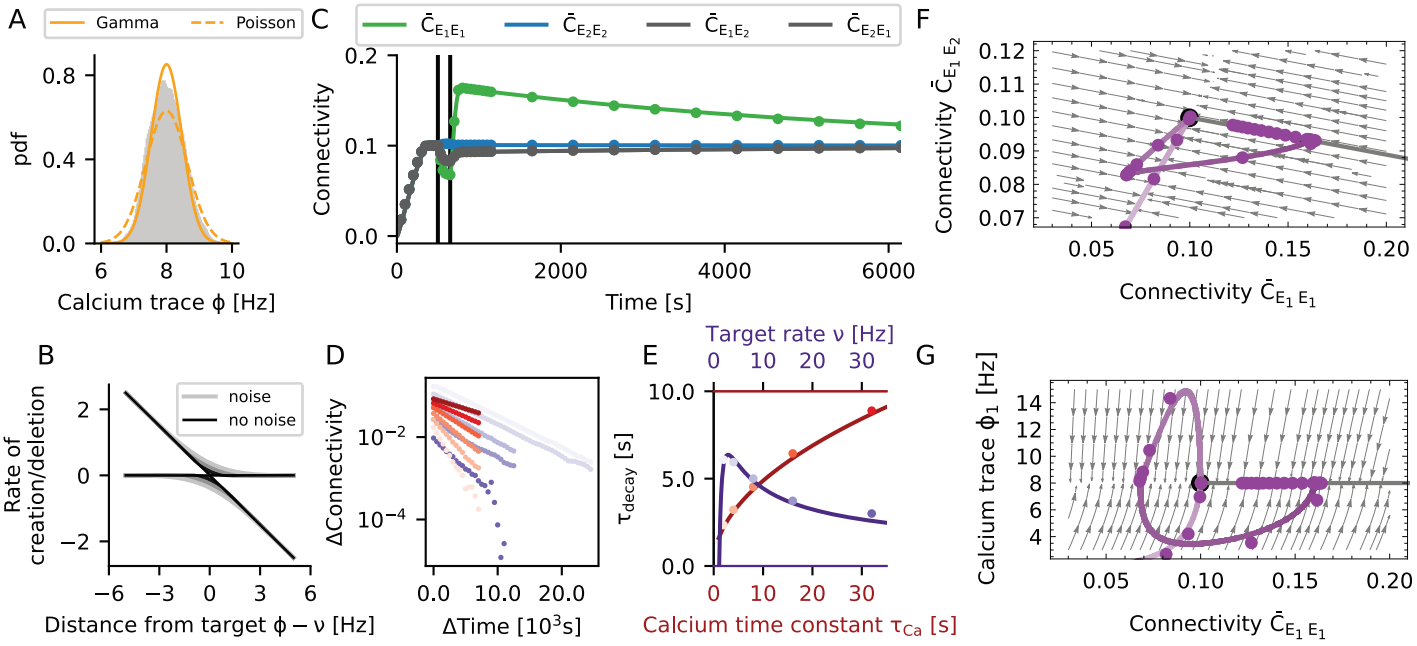


Fig 5. Noisy spiking induces fluctuations that lead to memory decay. (A) The gray histogram shows the distribution of calcium levels for a single neuron across 5000 s of simulation. The yellow lines resulted from modeling the spike train as a Poisson process (dotted line) or a Gamma process (solid line), respectively. (B) The rate of creation or deletion of synaptic elements depends on the difference between the actual firing rate from the target rate (set-point), for different levels of spiking noise. The negative gain (slope) of the homeostatic controller in presence of noise is transformed into two separate processes of creation and deletion of synaptic elements. In the presence of noise (grey lines, lighter colors correspond to stronger noise), even when the firing rate is on target, residual fluctuations of the calcium signal induced a continuous rewiring of the network, corresponding to a diffusion process. (C) If noisy spiking and the associated diffusion was included in the model, our mean-field theory matched the simulation results very well. This concerns the initial decay, the overshoot and subsequent slow decay. Vertical lines indicate the beginning and the end of the stimulation. (D) Change in connectivity during the decay period, for different values of the calcium time constant (different shades of red, from light to dark $\tau_{Ca} = 1, 2, 4, 8, 16, 32$ s, $\rho = 8$ Hz) and the target rate (different shades of purple, from light to dark $\rho = 2, 4, 8, 16, 32$ Hz, $\tau_{Ca} = 10$ s). We generally observed exponential relaxation as a consequence of a constant rewiring rate. (E) Time constant of the diffusive decay as a function of the calcium time constant and the target rate. Lines show our predictions from theory, and dots represent the values extracted from numerical simulations of plastic networks. The decay time τ_{decay} increases with $\sqrt{\tau_{Ca}}$. The memory was generally more stable for small target rates ν, but collapsed for very small rates. This indicates an optimum for low firing rates, at about 3 Hz. (F, G) Same phase diagrams as shown in Fig 4E and 4F, but taking noise into consideration. (F) The spiking noise compromises the stability of the line attractor, which turns into a slow manifold. (G) The relaxation to the high-entropy connectivity configuration during the decay phase is indeed confined to a constant firing rate manifold.

https://doi.org/10.1371/journal.pcbi.1009836.g005

over the spiking noise (Fig 5B, red lines). The shape of this function indicates that connections were randomly created and deleted even when neurons were firing at their target rate. The larger the amplitude of the noise, the larger is the asymptotic variance of the process and the amount of spontaneous rewiring taking place.

We now extended our mean-field model of the rewiring process (see Section The mechanism of engram formation) to account for the spiking noise (see Section Spiking noise). According to the enhanced model, both the overshoot and the decay now matched very well with numerical simulations of plastic networks (Fig 5C). The decay of connectivity following its overshoot was exponential (Fig Fluctuation-driven decay of engramsD), and Eq 3 revealed that the homeostasis was multiplicative and that the decay rate should be constant. The exponential nature of the decay is best understood by inspecting the phase space (Fig 5F and 5G). In terms of connectivity, the learning process was qualitatively the same as in the noise-less case (Fig 4E), where a small perturbation led to a fast relaxation to the line attractor (see Eq 16 in Section Line attractor of the deterministic system). In the presence of spiking noise (Fig 5F), however, the line attractor was deformed into a slow stochastic manifold (see Eq 18 in Section Slow manifold and diffusion to the global fixed point). The process of memory decay corresponded to a very slow movement toward the most entropic stable configuration compatible with firing rates clamped at their target value (Fig 5G). In our case, this led to a constraint on the in-degree $K^{\text{in}} = \epsilon N_E$, where ϵ is the mean equilibrium connectivity of excitatory neurons and N_E is the number of excitatory neurons. In equilibrium, the connectivity matrix relaxed to a uniform connection probability $\mathbb{P}_L[C^*_{ij}] = \epsilon$. The realized connections, however, were constantly fluctuating. In every moment, the network had a configuration as for the Erdős-Rényi model. We refer to this state as "the" most entropic configuration. It will only change as a result of external stimulation. We summarize the memory decay process by the equation

$$\begin{aligned} \mathbb{P}_L[C^*_{ij}] &\xrightarrow{\tau_{\text{diffusion}}} \text{const.} \times K^{in}_i K^{out}_j = \frac{K^{in}}{N_E}, \\ \mathbb{P}_L[C^*_{ij}] &\xleftarrow{\tau_{\text{drift}}} C^*_{ij} + \text{Stimulus}. \end{aligned} \tag{5}$$

The fast "drift" process of relaxation back toward the slow manifold corresponds to the deterministic part of Eq 3. In contrast, the slow "diffusion" process of memory decay along the slow manifold corresponds to the stochastic part of Eq 3. In the equation above, τ_{drift} represents the time scale of the fast drift process, and $\tau_{\text{diffusion}}$ is the time scale of the slow diffusion process.

The drift process is strongly non-linear, and its bandwidth is limited by the time constant of the calcium filter τ_{Ca}, but also by the growth parameters of the dendritic elements β_d and axonal elements β_a. The diffusion process, on the other hand, is essentially constrained to the slow manifold of constant in-degree and firing rate (Section Slow manifold and diffusion to the global fixed point). Analytic calculations yield the relation

$$\tau_{\text{diffusion}} = \sqrt{\frac{4\pi\tau_{\text{Ca}}}{\eta^2 \nu}} \frac{N_E c}{\frac{1}{\beta_d} + \frac{1}{\beta_a}} \tag{6}$$

where N_E is number of excitatory neurons, c is the average connectivity between excitatory neurons, and η is a correction factor to account for the reduced irregularity of spike trains as compared to a Poisson process. Both size and connectivity of the network increase the longevity of stored memories. Assuming that neurons rewire at a constant speed, it takes more time to rewire more elements. There is an interesting interference with the noise process, as memory longevity depends on the time constant of calcium in proportion to $\sqrt{\tau_{\text{Ca}}}$ (Fig 5E). On the other hand, the time scale of learning τ_{drift} is limited by the low-pass characteristics of calcium,

represented by τ_{Ca}. Increasing τ_{Ca} leads to more persistent memories, but it also makes the learning slower. In the extreme case of $\tau_{Ca} \to \infty$ the system is unresponsive and exhibits no learning. As the calcium time constant, however, is finite, we don't consider this possibility. We also find that the longevity of memories depends as $\frac{1}{\sqrt{\nu}}$ on the target rate. Therefore, making the target rate small enough should lead to very persistent memories. This path to very stable memories is not viable, though, because the average connectivity c implicitly changes with target rate. In Fig 5E longevity of the memory is depicted as a function of the target rate with corrected connectivity $c = c(\nu)$, and we find that for very small target rates memory longevity tends to zero instead. This suggests that there is an optimal range of target rates centered at a few spikes per second, fully consistent with experimental recordings from cortical neurons [32].

To summarize, the process of forming an engram ("learning", see Section The mechanism of engram formation) exploits the properties of a line attractor. Taking spiking noise into account, the structure of the line attractor is deformed into a slow stochastic manifold. This still allows learning, but introduces controlled "forgetting" as a new feature. Forgetting is not necessarily an undesirable property of a memory system. In a dynamic environment, it might be an advantage for the organism to forget non-persistent or unimportant aspects of it. Homeostatic plasticity implements a mechanism of forgetting with an exponential time profile. However, psychological forgetting curves are often found to follow a power law [33]. The presented model cannot account for this directly, and additional mechanisms (e.g., replay, repeated stimulation, or alternative forms of plasticity) might be necessary to get this. Strong memories are sustained for a longer period, but they will eventually also be forgotten. In the framework of this model, the only way to keep memories forever is to repeat the corresponding stimulus from time to time, as illustrated in Fig 1D. If we think of the frequency of occurrence as a measure of the relevance of a stimulus, this implies that irrelevant memories decay but the relevant ones remain. Memories are stored in a way where the increase in connectivity within the engram is accompanied by a decrease of connectivity with non-engram neurons. This means that storage is linked with distributed changes in the connectivity matrix, instead of being stored in specific and localized synaptic connections [11]. Forgetting is reflected by a diffusion to the most entropic network configuration along the slow manifold. As a result, the system performs continuous inference from a persistent stream of information about the environment. Already stored memories are constantly refreshed in terms of a movement in directions away from the most entropic point of the slow manifold (novel memories define new directions), while diffusion pushes the system back to the most entropic configuration.

Network stability and constraints on growth parameters

So far, we have described the process of forming and maintaining memory engrams based on homeostatic structural plasticity. We have explained the mechanisms behind the striking associative properties of the system. Now, we will explore the limits of stability of networks with homeostatic structural plasticity and derive meaningful parameter regimes for a robust memory system. A homeostatic controller that operates on the basis of firing rates can be expected to be very stable by construction. Indeed we find that, whenever parameters are assigned meaningful values, the Jacobian $\mathcal{J} = \mathcal{J}(\epsilon)$ obtained by linearization of the system around the stable mean connectivity ϵ has only eigenvalues λ with non-positive real parts $\mathrm{Re}[\lambda] \leq 0$ (see Section Linear stability analysis). As a demonstration, we consider the real part of the two "most unstable" eigenvalues as a function of two relevant parameters, the dendritic/axonal growth parameter β_d and the calcium time constant τ_{Ca} (Fig 6B, upper panel). The real part of these eigenvalues remains negative for any meaningful choice of time constants. It should be noted at this point, however, that the system under consideration is strongly non-linear, and

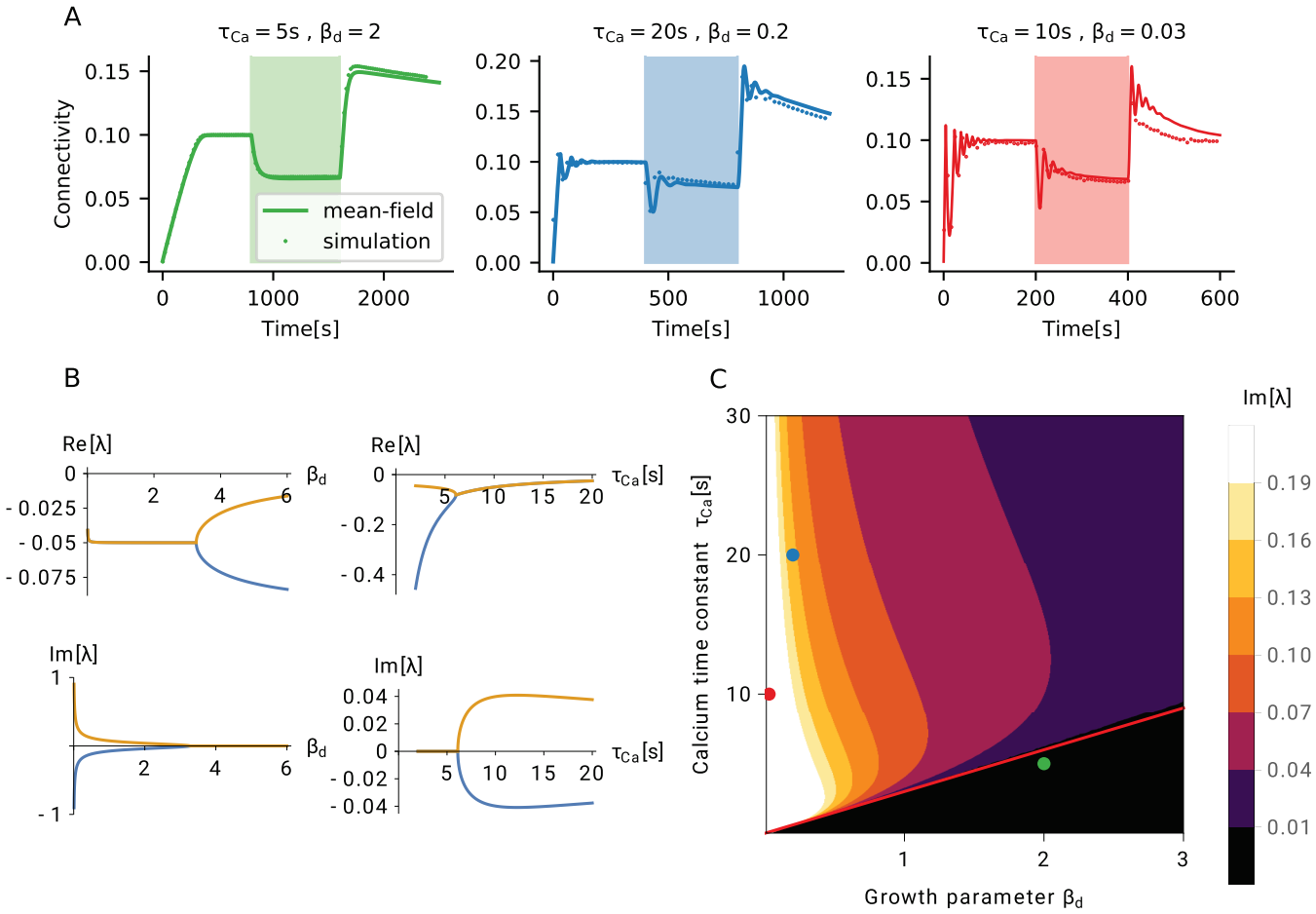


Fig 6. Linear stability of a network with homeostatic structural plasticity. (A) For a wide parameter regime, the structural evolution of the network has a single fixed point, which is also stable. Three typical types of homeostatic growth responses are depicted for this configuration: non-oscillatory (left), weakly oscillatory (middle), and strongly oscillatory (right) network remodeling. (B) All eigenvalues of the linearized system have a negative real part, for all values of the growth parameters of dendritic (axonal) elements β_d and calcium τ_{Ca}. In the case of fast synaptic elements (small β_d) or slow calcium (large τ_{Ca}), the system exhibits oscillatory responses. Shown are real parts and imaginary parts of the two "most unstable" eigenvalues, for different values of β_d (left column, $\tau_{Ca} = 10$ s) and τ_{Ca} (right column, $\beta_d = 2$). (C) Phase diagram of the linear response. The black region below the red line indicates non-oscillatory responses, which corresponds to the configuration $\tau_{Ca} \leq 3\beta_d$ s. Dots indicate the parameter configurations shown in panel (A), with matching colors.

https://doi.org/10.1371/journal.pcbi.1009836.g006

linear stability alone does not guarantee global stability. We will discuss an interesting case of non-linear instability in Section Loss of control leads to bursts of high activity. Oscillatory transients represent another potential issue in general control systems, and we will now explore the damped oscillatory phase of activity in more detail.

In Fig 6A we depict three typical cases of homeostatic growth responses: non-oscillatory (left, green), weakly oscillatory (middle, blue), and strongly oscillatory (right, red) network remodeling. The imaginary parts of the two eigenvalues shown in Fig 6B (bottom left), which are actually responsible for the oscillations, become non-zero when the dendritic and axonal growth parameters are too small (for other parameters, see Section The mechanism of engram formation), and both creation and deletion of elements are too fast. Oscillations occur, on the other hand, for large values of the calcium time constant τ_{Ca} (Fig 6B, bottom right). The system oscillates, if the low-pass filter is too slow as compared to the turnover of synaptic elements. The combination of β_d and τ_{Ca} that leads to the onset of oscillations can be derived from the condition $\mathrm{Im}[\lambda] = 0$. We can further exploit the fact that two oscillatory eigenvalues are

complex conjugates of each other, and that the imaginary part is zero, when the real part bifurcates.

To elucidate the relative importance of the two parameters β_d and τ_{Ca}, we now explore how they together contribute to the emergence of oscillations (Fig 6C). The effect of parameters on oscillations is a combination of the two mechanisms discussed above: low-pass filtering and agility of control. We use the bifurcation of the real part of the least stable eigenvalues as a criterion for the emergence of $\mathrm{Im}[\lambda] = 0$, which yields the boundary between the oscillatory and the non-oscillatory region (Fig 6C, red line). The black region of the phase diagram corresponds to a simple fixed point with no oscillations, the green point corresponds to the case shown in Fig 6A, left. From Fig 6B, bottom, we conclude that the fastest oscillations are created when β_d is very small, as the oscillation frequency then exhibits a $\frac{1}{\beta_d}$ asymptotic dependence. The calcium dependence is a slowly changing function. The specific case indicated by a red point in Fig 6C corresponds to the dynamics shown in Fig 6A, right. Although intermediate parameter values result in damped oscillations (Fig 6C, blue dot; Fig 6A, middle), its amplitude remains relatively small. This link between two parameters can be used to predict a meaningful range of values for β_d. Experimentally reported values for τ_{Ca} are combined with the heuristic of not exhibiting strong oscillations.

The analysis outlined in the previous paragraph clearly suggests that, in order to avoid excessive oscillations, the calcium signal (a proxy for neuronal activity) has to be faster than the process which creates elements. Oscillations in network growth are completely suppressed, if it is at least tree times faster. Strong oscillations can compromise non-linear stability, as we will show in Section Loss of control leads to bursts of high activity. A combination of parameters that leads to oscillations could also imply a loss of stability for certain stimuli, for reasons explained later. We assume that the calcium time constant is in the range between 1 s and 10 s, in line with the values reported for somatic calcium transients in experiments [34–37]. This indicates that homeostatic structural plasticity should not use element growth parameters smaller than around 0.4. Faster learning must be based on other types of synaptic plasticity (e.g. spike-timing dependent plasticity, or fast synaptic scaling).

We use this analysis framework now to compute turnover rates (TOR) and compare them to the values typically found in experiments. In Sections The mechanism of engram formation and Fluctuation-driven decay of engrams we use a calcium time constant of 10 s and a growth parameters for synaptic elements of $\beta_a = \beta_d = 2$, which results in TOR of around 18% per day (see Methods). Interestingly, [38] measured TOR in the barrel cortex of young mice and found TOR values of around 20% per day. After sensory deprivation, the TOR increased to a maximum of around 30% per day in the barrel cortex (but not elsewhere). In our model, stimulus-dependent rewiring is strongest in the directly stimulated engram E_1 (Fig 4C). This particular ensemble rewires close to 25% of its dendritic elements per stimulation cycle. This very large turnover is a result of our experimental design involving a very strong stimulus. In principle, we could choose weaker stimuli, but then we would need many more encoding episodes, as described in [20]. The main results of the paper, however, would remain the same.

Loss of control leads to bursts of high activity

A network the connectivity of which is subject to homeostatic regulation generally exhibits robust linear stability around the fixed point of connectivity ϵ, as explained in detail in Section Network stability and constraints on growth parameters and Section Linear stability analysis. But what happens, if the system is forced far away from its equilibrium? To illustrate the new phenomena arising, we repeat the stimulation protocol described in Section The mechanism of engram formation with one stimulated ensemble E_1. However, we now increase both the

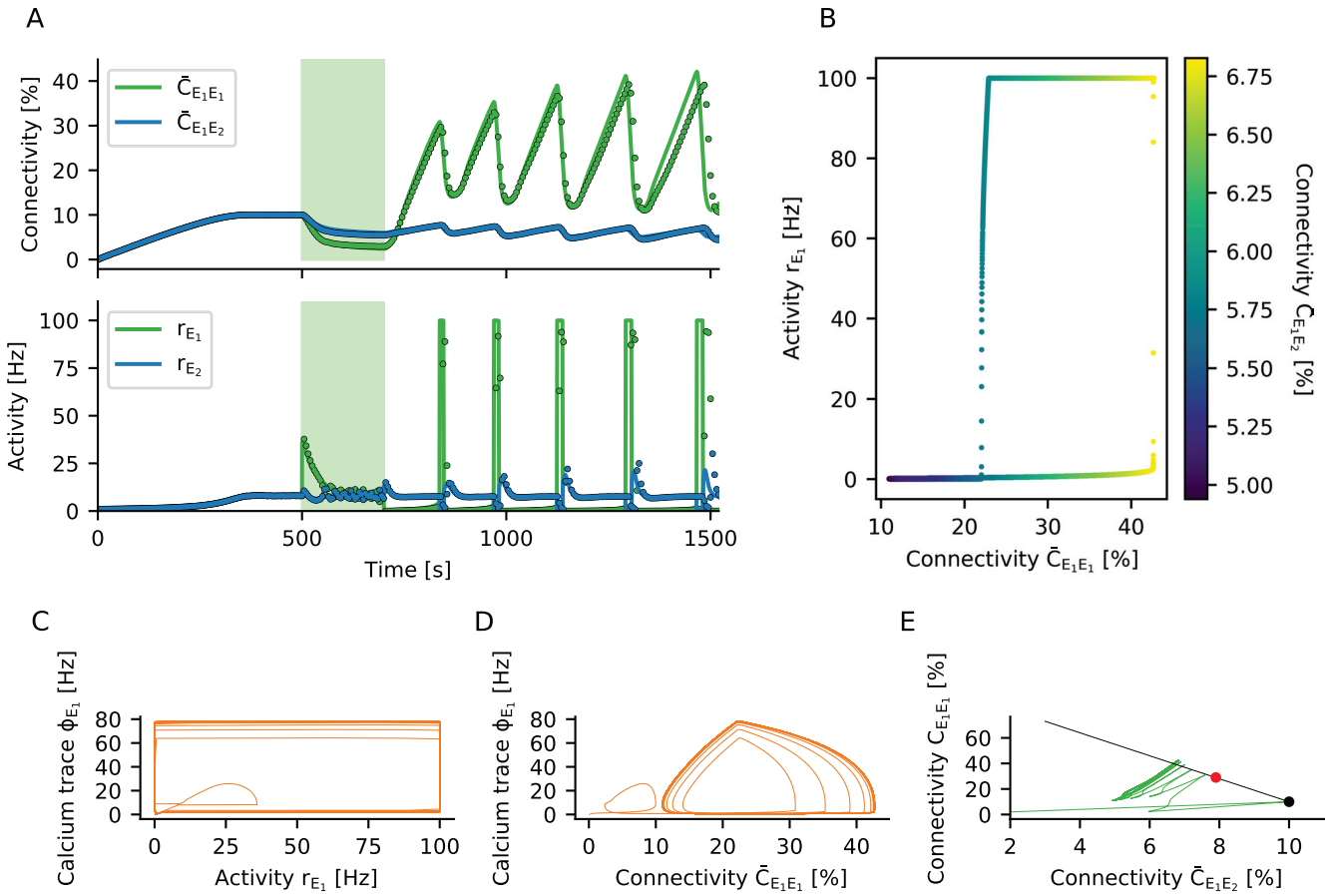


Fig 7. Non-linear stability of a network with homeostatic structural plasticity. (A) The engram E_1 is stimulated with a very strong external input. As the homeostatic response triggers excessive pruning of recurrent connections, the population E_1 is completely silenced after the stimulus is turned off. This, in turn, initiates a strong compensatory overshoot of connectivity and consecutive runaway population activity. The dots with corresponding color show the results of a plastic network simulation, and the solid lines indicate the corresponding predictions from our theory. The theoretical instantaneous firing rate is clipped at 100 Hz. (B) The network settles in a limit cycle of connectivity dynamics. The hysteresis-like behavior is caused by the faster growth of within-engram connectivity $\bar{C}_{E_1E_1}$ as compared to connectivity from the non-engram ensemble $\bar{C}_{E_1E_2}$. During the initial phase of the cycle, the increase of $\bar{C}_{E_1E_1}$ has no effect on the activity of population E_1 yet, as its neurons are not active. Only when the input from population E_2 through $\bar{C}_{E_1E_2}$ gets large enough, the rate r_{E_1} becomes non-zero and rises to very high values quickly due to already large recurrent $\bar{C}_{E_1E_1}$ connectivity. (C) The calcium signal ϕ adds an additional delay to the cycle. (D) This leads to smoother trajectories when scattering calcium concentration against connectivity. (E) Connectivity within the stimulated group $\bar{C}_{E_1E_1}$ plotted against input connectivity from the non-engram population $\bar{C}_{E_1E_2}$. The black line shows configurations with constant in-degree, of which the black dot represents the most entropic one. The red dot corresponds to critical connectivity, beyond which the limit cycle behavior is triggered. The limit cycle transients in connectivity space are orthogonal to the line attractor, indicating that the total in-degree is oscillating and no homeostatic equilibrium can be established.

https://doi.org/10.1371/journal.pcbi.1009836.g007

strength and the duration of the stimulation (Fig 7A). The network behaves as before during the growth and the stimulation phase (Fig 7A, upper panel), but during the reconsolidation phase connectivity gets out of control. New recurrent synapses are formed at a very high rate until excessive feedback of activity triggered by input from the non-stimulated ensemble causes an explosion of firing rates (Fig 7A bottom panel). The homeostatic response of the network to such seizure-like activity can only be a brisk decrease in recurrent connectivity. As a consequence, the activity r_{E_1} quickly drops to zero and the deregulated growth cycle starts all over.

We explore the mechanism underlying this runaway process by plotting the long-term dynamics in a phase plane spanned by recurrent connectivity $\bar{C}_{E_1E_1}$ and the activity of the

engram r_{E_1} (Fig 7B). A special type of limit cycle emerges, and we can track it using the input connectivity from the rest of the network to the learned engram ($\bar{C}_{E_1E_2}$, see Fig 7B). The cycle is started when the engram E_1 is stimulated with a very strong external input. As the homeostatic response of the network triggers excessive pruning of its recurrent connections, the population E_1 is completely silenced, $r_{E_1} = 0$, after the stimulus has been turned off. Then, homeostatic plasticity sets in and tries to compensate the activity below target by increasing the recurrent excitatory input to the engram E_1. The growth of intra-ensemble connectivity $\bar{C}_{E_1E_1}$ is faster than the changes in inter-ensemble connectivity $\bar{C}_{E_1E_2}$, as the growth rate of $\bar{C}_{E_1E_1}$ is quadratic in the rate r_{E_1} (Fig 4G), but $\bar{C}_{E_1E_2}$ depends only linearly on it. However, while there are no recurrent spikes, $r_{E_1} = 0$, the increase in intra-ensemble connectivity cannot restore the activity to its target value. As soon as input from the rest of the network via $\bar{C}_{E_1E_2}$ is strong enough to increase the rate r_{E_1} to non-zero values, engram neurons very quickly increase their own rate by activating recurrent connectivity $\bar{C}_{E_1E_1}$. At this point, however, the network has entered a state in which the population activity is bistable, and there is a transition to high activity state, coinciding with a pronounced outbreak of population activity r_{E_1}. The increase in rate is then immediately counteracted by the homeostatic controller. Due to the seizure-like activity burst, a high amount of calcium is accumulated in all participating cells. As a consequence, neurons delete many excitatory connections, and the firing rates are driven back to zero. This hysteresis-like cycle of events is repeated over and over again (Fig 7B), even if the stimulus has meanwhile been turned off. The period of the limit cycle is strongly influenced by the calcium variable, which lags behind activity (Fig 7C). Replacing activity r_{E_1} by recurrent connectivity $\bar{C}_{E_1E_1}$, a somewhat smoother picture emerges (Fig 7D).

Two aspects are important for the emergence of the limit cycle. Firstly, the specific relation between the time scales of calcium and synaptic elements gives rise to different types of instabilities (see Figs 6, 7C and 7D). Secondly, the rates of creation and deletion of elements do not have the same bounds. While the rate with which elements are created ρ^+ is limited by $\frac{1}{\beta_d}\nu$, the rate of deletion ρ^- is limited by $\frac{1}{\beta_d}\frac{1}{\tau_{\text{ref}}}$. This peculiar asymmetry causes the observed brisk decrease in connectivity after an extreme seizure-like burst of activity. An appropriate choice of the calcium time constant, in combination with a strict limit on the rate of deletion, might lead to a system without the (pathological) limit cycle behavior observed in simulations. Finally, we have derived a criterion for bursts of population activity to arise, related to the loss of stability due to excessive recurrent connectivity (see Section Linear stability analysis, Eq 22). Indeed, a network with fixed in-degrees becomes dynamically unstable, if the connectivity of subpopulation E_1 exceeds the critical value $\bar{C}^{\text{crit}}_{E_1E_1} = \epsilon\left(1 + \frac{N_{E2}}{N_{E1}j_E}\right)$, which for our parameters is at about 29%. The neurons comprising the engram E_1 receive too much recurrent input (Fig 7E), and the balance of excitation and inhibition brakes. In this configuration, only one attractive fixed point exists for high firing rates, and a population burst is inevitable. In simulations, the stochastic nature of the system tends to elicit population bursts even earlier, at about 22% connectivity in our hands. We conclude that 22–29% connectivity is a region of bi-stability, with two attractive fixed points coexisting. Early during limit cycle development, the total in-degree is less than ϵN (the connectivity is in the region below the black line in Fig 7E), and the excitation-inhibition balance is broken by positive feedback. Later, the limit cycle settles into a configuration, where the total in-degree exceeds ϵN (the black line is crossed from below in Fig 4E). We have shown before that stable learning leads to silent memories in the network (Section Formation of memory engrams by homeostatic structural plasticity and Engrams

represent silent memories, not attractors), but in the case discussed here, sustained high activity is at odds with stable homeostatic control of network growth.

Discussion

We have demonstrated by numerical simulations and by mathematical analysis that structural plasticity controlled by firing rate homeostasis can implement a memory system based on the emergence and the decay of engrams. Input patterns are defined by a stimulation of the corresponding ensemble of neurons in a recurrent network. Presenting two patterns concurrently leads to their association by newly formed synaptic connections between the involved neurons. This mechanism can be used, among other things, to effectively implement classical conditioning. The memories are stored in network connectivity in a distributed fashion, defined by the engram as a whole and not by isolated individual connections. The memories are dynamic. They decay if previously learned stimuli are no longer presented, but they get stronger with every single recall. The memory is not affecting the firing rates during spontaneous activity, but even weak memory traces can be identified by the correlation of activity. Memories become visible as a firing rate increase of a specific pattern upon external stimulation, though. The embedding network is able to perform pattern completion, if a partial cue is presented. Finally, we have devised a simple recognition memory mechanism, in which downstream neurons respond with a higher firing rate, if any of the previously learned patterns is stimulated.

Memory engrams emerge because the homeostatic rule acts as an effective Hebbian rule with associative properties. This unexpected behavior is achieved by an interaction between the temporal dynamics of homeostatic control and a network-wide distributed formation of synapses. Memory formation is a fast process, exploiting degrees of freedom orthogonal to a line attractor while it reacts to the stimulus, and storing memories as positions on the line attractor. The spiking of neurons introduces fluctuations, which lead to the decay of memory on a slow time scale through diffusion along the line attractor. In absence of specific stimulation, the network slowly relaxes to the most entropic configuration of uniform connectivity across all pairs of neurons. In contrast, multiple repetitions of a stimulus push the system to states of lower entropy, corresponding to stronger memories. The dynamics of homeostatic networks is, by construction, very robust for a wide parameter range. Instabilities occur when the time scales of creating and deleting synaptic elements are much smaller than the time scales of the calcium trace, which feeds the homeostatic controller. Under these conditions, the network displays oscillations of fast decaying amplitudes, but it remains linearly stable. Stability is lost, though, when the stimulus is too strong. In this case, the compensatory forces lead to a limit cycle dynamics with pathologically large amplitudes.

Experiments involving engram manipulation have increased our current knowledge about this type of memory [28], and some recent findings are actually in accordance with our model. For example, memory re-consolidation was disrupted if a protein synthesis inhibitor was administered immediately after the retrieval cue during an auditory fear conditioning experiment [39]. In our model, engram connectivity initially decreases upon stimulation, and memories are shortly destabilized and consolidated again after every retrieval. Interfering with plasticity during or after retrieval could, therefore, also lead to active forgetting. Our model also predicts a decreased connectivity following cue retrieval, which should be visible if the disruption is imposed at the right point in time. We are, however, not aware of any experiments demonstrating this decreased connectivity. It is conceivable, therefore, that additional mechanisms exist in biological networks that actively counteract such decay. Another example of experiments that are in accordance to our model refers to excitability and engram allocation. Recent experiments have shown that neurons are more likely to be allocated to an engram, if

they are more excitable before stimulation [40, 41]. In our model, more excitable neurons would fire more during stimulation, making them more likely to become part of an engram as a result of the increased synaptic turnover. Moreover, our analysis of the model suggests that decreasing excitability of some neurons soon after stimulation should also increase the likelihood that they become part of the engram. Further research with our model of homeostatic engram formation might include even more specific predictions for comparison with experiments involving engram manipulation. This would also help to better characterize and understand the process of engram formation in the brain.

The exact rules governing the sprouting and pruning of spines and boutons, and how these depend on neuronal activity, are unknown. While some studies show a constant turnover of spines, but fixed spine density [42], [38], other studies report either an increased spine density after stimulation in LTP protocols [43], [44], [45], or a loss of spines after persistent depolarization [46], [47], see [48] for a review on activity dependent structural plasticity. These different results might seem contradictory at first, but our model suggests that they could all be related. This is because the specific connectivity within the stimulated population undergoes a bimodal change over time (Fig 4D), and whether one encounters an increase or a decrease in spine density would depend on when the measurements were performed. More specifically, our model predicts that the overall spine density remains unchanged before and after stimulation (Fig 4C), notwithstanding a small temporary decrease during stimulation. Most importantly, however, our model emphasizes a reallocation of synapses, resulting in higher connectivity between stimulated neurons and reduced connectivity to non-stimulated neurons. This means that the overall number of spines is the same before and after stimulation (same number of dendritic elements before and after stimulation on Fig 4). However, the number of spines connecting to other stimulated neurons is higher after stimulation (green line on Fig 4D) and the number of spines connecting to non-stimulated neurons is smaller (grey lines on Fig 4D). Demonstrating such specific effects experimentally would require to establish for each synapse separately, whether the presynaptic and the postsynaptic neuron was stimulated, or not. To our knowledge, there is currently no experimental study available reporting such labeled connectivity data. As the model is inherently stochastic, it is not necessary that a synapse is deleted each time their presynaptic or postsynaptic neurons are stimulated. Instead, random deletions between any pair of stimulated neurons might lead to a small change in overall engram connectivity. Although small, these decrements in connectivity accumulate over time, as a result of repetitive stimulation [20].

High turnover rates of synapses increase the volatility of network structure. This, in turn, poses a grand challenge to any synaptic theory of memory [11], and it is not yet clear how memories can at all persist in a system that is constantly rewiring [49]. In our model, the desired relative stability of memories is achieved by storing them with the help of a slow manifold mechanism. An estimation of turnover rates in our model amounts to about 18% per day, which is comparable to the 20% per day that have been measured in mouse barrel cortex [38]. In general, however, adult mice have more persistent synapses with much lower turnover rates as low as 4% per month [10, 50, 51]. This can be accounted for in our model, as increased growth parameters of axonal, β_a, and dendritic, β_d, elements would lead to smaller synaptic turnover rates and, consequently, to more persistent spines (see Methods). The downside of increasing the growth parameters is that the learning process becomes slower. The turnover rate of 18% per day corresponds to a specific value of the parameter β_d. It is conceivable, however, to implement an age-dependent parameter β_d. For example, one could have a high turnover rate in the beginning and let the growth become slower with time. This would reflect the idea that the brains of younger animals are more plastic than the brains of older ones. As animals grow older, synapses become more persistent. Similar to certain machine learning

strategies ("simulated annealing"), this could be the optimal strategy for an animal, which first explores a given environment and then exploits the acquired adaptations to thrive in it.

Recently, [12] showed that Hebbian structural plasticity could be the force behind memory consolidation through a process of stabilization of connectivity, which is based on the existence of an attractive fixed point in the plastic network structure. In our model, because of the decay along the slow manifold, memories are never permanent, and repeated stimulation is necessary to refresh them. We would argue, though, that forgetting is an important aspect of any biological system. In our case, we observe an exponential decay, if the stimulus is no longer presented. Furthermore, [25] has shown that a network can repair itself after lesion using a structural plasticity model similar to the one used in our current paper. Together with our results, this suggests that a structural perturbation of engrams (e.g. by removing connections or deleting neurons) could actually trigger a "healing" process and eventually rescue the memory. In the case of unspecific lesions, however, such perturbation might also lead to the formation of "fake" memories, or to the false association of actually unconnected memory items.

It appears that the attractor metaphor of persistent activity is not consistent with our model of homeostatic plasticity. As explained in Section Loss of control leads to bursts of high activity, homeostatic control tends to delete connections between neurons which are persistently active, and in extreme cases could even lead to pathological oscillations. In the case discussed in Section Engrams represent silent memories, not attractors, in contrast, the memories formed are "silent" (elsewhere classified as "transient" [52]), very different from the persistent activity usually considered in working memory tasks (elsewhere classified as "persistent", or "dynamic" [52]). It was previously shown that silent memories can emerge in networks with both excitatory and inhibitory plasticity [4–6]. In all these cases, inhibitory plasticity allows the memory to be silent, but memory formation still relies on explicit Hebbian plasticity rules of excitatory connections. Some authors [5, 6] apply STDP to the excitatory connections, and others [4] just impose an increase in weight of excitatory-to-excitatory connections between neurons forming the assembly. As we understand it, such an increase could be achieved by Hebbian plasticity of excitatory connections. In contrast, our work shows that silent memories could potentially also be formed without correlation-based Hebbian plasticity on excitatory connections. Persistent activity, on the other hand, seems to be exclusively consistent with Hebbian plasticity models [2, 3, 12]. One consequence of the lack of persistent activity in our model is that pattern completion is restricted to the stimulation period. As previously seen, in our model, stimulating partial patterns leads to the subsequent activation of the full pattern (Fig 3C and 3D). Given that there is no persistent activity, however, the firing rates of all neurons go back to their baseline values as soon as the stimulus is gone. In models with persistent activity, on the other hand, a brief stimulation of partial patterns leads to activation of the full pattern, which remains active even after the stimulus is removed.

One possible way to integrate both mechanisms in a single network would be to keep their characteristic time scales separate. This could be accomplished, for example, by choosing faster time constants for Hebbian functional plasticity, and slower ones for homeostatic structural plasticity. An effective separation of time scales could also be obtained, if homeostatic structural plasticity would use somatic calcium as a signal, but not exert any control of the intermediate calcium levels in dendritic spines [53]. This might eliminate the need to specify a target rate in the model, and fast functional plasticity would shape connectivity in the allowed range of values where neurons have a distribution of firing rates reflecting previous experience. This induces a natural separation of time scales, where memories encoded by homeostatic plasticity would last much longer than in the present model, as only extreme transients would trigger rewiring. Homeostatic plasticity would perform Bayesian-like inference similar to structure learning, while functional plasticity would perform fast associative learning, similar to the

system proposed by [54]. Integrating both functional and structural plasticity opens the possibility that different information is represented by the number of synapses and by the synaptic weights between pairs of neurons. This would allow for the encoding of more complex patterns. If the mean connectivity among neurons establishes the engram, the synaptic weights would offer additional degrees of freedom to encode information and to modulate the activity of individual neurons within the engram in an independent way. Different plasticity rules could thus be used to encode different temporal aspects of neuronal activity in either synaptic weights or synaptic connectivity.

Synaptic plasticity also influences the joint activity dynamics of neurons, which can be assessed with appropriate data analysis methods. Functional and effective connectivity, for example, are inferred from measured neuronal activity [55], [56], [57]. One implication of our results is that different activity dependent plasticity rules can lead to the same changes in functional and effective connectivity. Specifically, changes in network connectivity triggered by homeostatic plasticity also lead to changes in correlation. Therefore, these changes appear to be driven by correlation, and a correlation detector for each synapse might be postulated to implement it. This is, however, a wrong conclusion. The increased correlation and the Hebbian properties of the homeostatic model emerge as a network property due to availability of free elements and random self-organization. It remains a challenge to devise experiments that can differentiate between these fundamentally different possibilities.

We showed that very strong stimulation can damage the network by deleting too many synapses in a short time. The compensatory processes, which normally guarantee stability, get out of control and lead to seizure-like bursts of very high activity. This pathological behaviour of the overstimulated system could contribute to the etiology of certain brain diseases, such as epilepsy. The disruption of healthy stable activity is caused by a broken excitation-inhibition balance due to the high activity of one subgroup (Fig 7A). This, in turn, leads to the emergence of a runaway connectivity cycle (Fig 7E). Strategies for intervention in this case must take the whole cycle into account, and not just the phase of extreme activity. Inhibiting neurons during the high-activity phase, for example, could have an immediate effect, but it would not provide a sustainable solution to the problem of runaway connectivity. Our results suggest, against intuition maybe, that additional excitation of the highly active neurons could actually terminate the vicious cycle quite efficiently. It is important to note, however, that our system has not been designed as a model of epilepsy, and therefore does not reproduce all features of it [58]. In particular, seizure occurrence is stochastic in nature, but the limit cycle we describe here implies periodic activity and a periodic dynamics of connectivity. Although increased mossy-fibre connectivity among granule cells is known to be one of the structural changes related to epilepsy [59], there is currently no evidence for cyclic changes of this recurrent connectivity. Furthermore, the process of epileptogenesis in real brains is accompanied by other structural changes, such as neuronal death and glia-related tissue reorganization. In any case, our results shed light on a novel mechanism of pathological structural overcompensation and could potentially instruct alternative approaches in future epilepsy research.

René Descartes already proposed a theory of memory, paraphrased in [11]: Putting needles through a linen cloth would leave traces in the cloth that either stay open, or can more easily be opened again. Richard Semon, who originally coined the term "engram" in his book [28, 60], proposes that an engram is a "permanent record" formed after a stimulus impacts an "irritable substance". Putting these two ideas together, we can think of Descartes' needles not to penetrate an inanimate linen cloth, but a living brain, the irritable substance. In this case, we should expect the formation not of permanent holes, but of scar tissue, which grows further by repeating the procedure. Therefore, the memory of the system nothing else but a "scar" left by sensory experience. We think that this is a good metaphor for the type of memories described

in our paper, formed through homeostatically controlled structural plasticity. Interestingly, this model was originally meant as a model for rewiring after lesion [25]. In our work, however, the "lesion" is imposed by stimulation, which induces phenomena similar to scar formation. The dynamics of this healing process is very universal, where resources from the whole network are used to fix a local problem leading to a scar. A perturbation introduces heterogeneity in a previously homogeneous organic substrate.

Methods

Network model

The neuronal network consists of $N_E = 10000$ excitatory and $N_I = 2500$ inhibitory current-based leaky integrate-and-fire (LIF) neurons. The sub-threshold dynamics of the membrane potential V_i of neuron i obeys the differential equation

$$\tau_m \frac{dV_i}{dt} = -V_i + \tau_m \sum_j C_{ij} J_{ij} S_j(t - D) + RI_i(t). \tag{7}$$

The membrane time constant τ_m is the same for all neurons. The number of synaptic contacts between a presynaptic neuron j and a postsynaptic neuron i is denoted by C_{ij} (see Table 1 for a complete list of symbols). The synaptic weights of individual contacts J_{ij} is the peak amplitude of the postsynaptic potential and depends only on the type of the presynaptic neuron. Excitatory connections have a strength of $J_E = J = 0.1$ mV. Inhibitory connections are stronger by a factor $g = 8$ such that $J_I = -gJ = -0.8$ mV. A spike train $S_j(t) = \sum_k \delta(t - t_j^k)$ consists of all spikes produced by neuron j. The external input $RI_i(t)$ to a given neuron in the network is conceived as a Poisson process of rate $\nu_{\text{ext}} = 15$ kHz and amplitude $\tau_m J$. The external input to different neurons is assumed to be independent. All synapses have a constant transmission delay of $D = 1.5$ ms. When the membrane potential reaches the firing threshold $V_{\text{th}} = 20$ mV, the neuron emits a spike that is transmitted to all postsynaptic neurons. Its membrane potential is then reset to $V_r = 10$ mV and held there for a refractory period of $t_{\text{ref}} = 2$ ms.

The number of input synapses is fixed at $0.1N_I$ for inhibitory-to-inhibitory and inhibitory-to-excitatory connections, and at $0.1N_E$ for excitatory-to-inhibitory synapses. Once synaptic connections of these three types are established, they remain unchanged throughout the simulation. In contrast, excitatory-to-excitatory connections are initially absent and grow only under the control of a structural plasticity rule.

Plasticity model

Growth and decay of excitatory-to-excitatory (EE) connections follow a known model of structural plasticity regulated by firing rate homeostasis [19, 25]. In this model, each neuron i has a certain number of synaptic elements of two kinds available, axonal elements $a_i(t)$ and dendritic elements $d_i(t)$. These elements are bonded together to create functional synapses. Synaptic elements that have not yet found a counterpart are called free elements, denoted by a_i^+ and d_j^+, respectively. If $K_i^{\text{in}} = \sum_j C_{ij}(t)$ denotes the in-degree and $K_i^{\text{out}} = \sum_j C_{ji}(t)$ the out-degree of neuron i, the number of free elements in every moment is given by $a_i^+(t) = [a_i(t) - K_i^{\text{out}}(t)]_+$ and $d_i^+(t) = [d_i(t) - K_i^{\text{in}}(t)]_+$, with $[x]_+ = \max(x, 0)$.

Firing rate homeostasis is implemented by allowing each neuron to individually control the number of its synaptic elements. We assume that each neuron i maintains a time-dependent estimate of its own firing rate, using its intracellular calcium concentration $\phi_i(t)$ as a proxy.

This variable reflects the spikes $S_i(t)$ the neuron has generated in the past, according to

$$\tau_{\mathrm{Ca}}\dot{\phi}_i(t) = -\phi_i(t) + S_i(t),$$

with a time constant of $\tau_{\mathrm{Ca}} = 10$ s for all neurons. This implements a first-order low-pass filter. In this model, more weight is given to more recent spikes, following a decaying exponential. The calcium trace $\phi_i(t)$ of individual neurons is used as a control signal for the number of axonal elements $a_i(t)$ and dendritic elements $d_i(t)$ according to the homeostatic equations

$$\beta_d \dot{b}_i(t) = \nu_i - \phi_i(t) \qquad \beta_a \dot{a}_i(t) = \nu_i - \phi_i(t),$$

where β_d is the dendritic and β_a is the axonal growth parameter. Both have a value of 2 in our default setup. The parameter ν_i is called the target rate. Whenever the firing rate estimate (in fact, the calcium concentration) is below the target rate, the neuron creates new axonal and dendritic elements, from which new synapses can be formed. Whenever the estimated firing rate is larger than the target rate, the neuron deletes some of its elements, removing the synapses they form. The (negative) decrements a_i^- and d_i^- in the number of synaptic elements, respectively, are in each moment given by $a_i^-(t) = [a_i(t) - K_i^{\mathrm{out}}(t)]_-$ and $d_i^-(t) = [d_i(t) - K_i^{\mathrm{in}}(t)]_-$ for $[x]_- = \min(x, 0)$.

After time intervals of duration ΔT_s, all negative elements are collected, and a_i^- of the existing outgoing connections of neuron i are randomly deleted. Similarly d_i^- connections are randomly deleted. The deletion of bonded elements of one type frees their counterparts that were previously connected to the deleted element. Then, all free dendritic $|\mathbf{d}^+|$ and axonal elements $|\mathbf{a}^+|$ are collected and randomly combined into pairs, creating $n = \min(|\mathbf{a}^+|, |\mathbf{d}^+|)$ new synaptic connections. This algorithm has originally been devised by [25], an efficient implementation of it in NEST exists [61] and has been employed for all our simulations.

Mathematical re-formulation of the algorithm

The algorithm of homeostatically controlled structural plasticity can be expressed as a discrete-time stochastic process. Rewiring takes place at regular intervals of duration ΔT_s

$$T_{\mathrm{rewire}} = \{t_0 = 0, t_1 = \Delta T_s, t_2 = 2\Delta T_s, \cdots, t_n = n_T \Delta T_s\}.$$

Between any two rewiring events, for $t \in (t_k, t_k + \Delta T_s)$, the neuron just accumulates synaptic elements

$$\begin{aligned} \Delta C_{ij}(t) &= 0 \\ a_i(t) &= \int_{t_k}^{t} \frac{1}{\beta_a}(\nu_i - \phi(t'))\, dt' \\ d_i(t) &= \int_{t_k}^{t} \frac{1}{\beta_d}(\nu_i - \phi_i(t'))\, dt' \end{aligned}$$

while the already established connectivity remains unchanged. At every rewiring step, the rearrangement of connectivity is completely random, driven by the probabilities $P(\Delta_{\pm} C_{ij}(t_k) = c \mid \mathbf{a}(t_k), \mathbf{d}(t_k), \mathbf{C}(t_{k-1}))$ of creating or deleting c connections from neuron j to neuron i at time t_k, during a time step of duration ΔT_s. This gives rise to a discrete-time stochastic process

$$C_{ij}(t_k) = C_{ij}(t_{k-1}) + \Delta_+ C_{ij}(t_k) - \Delta_{a^-} C_{ij}(t_k) - \Delta_{d^-} C_{ij}(t_k),$$

for $t_k \in T_{\mathrm{rewire}}$. Here, we define $\Delta_+ C_{ij}(t_k)$ as the random variable representing the creation of

synapses, while $\Delta_{a^-}C_{ij}(t)$ and $\Delta_{d^-}C_{ij}(t)$ are random variables describing the deletion of synapses by removing their corresponding axonal and dendritic elements, respectively.

We first calculate the probability to create just one new connection $p_{ij}^{+}(t_k) = P(\Delta_{+}C_{ij}(t_k) = 1|\mathbf{a}(t_k), \mathbf{d}(t_k), \mathbf{C}(t_{k-1}))$. We can express this as a process of selecting a presynaptic partner j with probability $P(\text{presynaptic neuron} = j) = \frac{a_j^+(t_k)}{|\mathbf{a}^+(t_k)|}$ and then a postsynaptic partner i with probability $P(\text{postsynaptic neuron} = i) = \frac{d_i^+(t_k)}{|\mathbf{d}^+(t_k)|}$. We then connect the pair with the product of both probabilities

$$p_{ij}^{+}(t_k) = \underbrace{\frac{d_i^+(t_k)}{|\mathbf{d}^+(t_k)|}}_{\text{prob. of choosing a post neuron}} \times \underbrace{\frac{a_j^+(t_k)}{|\mathbf{a}^+(t_k)|}}_{\text{prob. of choosing a pre neuron}}$$

as they are independent random variables. We now have the probability of creating individual connections, but the full probability of an increment for the whole network is hard to calculate. This is because of statistical dependencies that arise from the fact that rewiring affects all neurons in the network simultaneously. First, the total number of new connections is $n = \min(|\mathbf{a}^+|, |\mathbf{d}^+|)$. Second, the number of new connections for a given pair of neurons is bounded by the number of free axonal and dendritic elements in the two neurons, respectively, $n_{ij} = \min(d_i^+, a_j^+)$. Finally, we cannot delete more connections than we actually have. This indicates that independent combinations of individual probabilities cannot be expressed by simple binomial distributions, but a hypergeometric distribution arises instead. We obtain for the probability of creating c synapses from neuron j to neuron i

$$P(\Delta_{+}C_{ij}(t_k) = c|\mathbf{a}(t_k), \mathbf{d}(t_k), \mathbf{C}(t_{k-1})) = \frac{\binom{d_i^+(t_k)a_j^+(t_k)}{c}\binom{|\mathbf{a}^+(t_k)||\mathbf{d}^+(t_k)| - d_i^+(t_k)a_j^+(t_k)}{n(t_k) - c}}{\binom{|\mathbf{a}^+(t_k)||\mathbf{d}^+(t_k)|}{n(t_k)}}.$$

This probability is easy to understand: We divide the ensemble of all possible new synapses (which has size $|\mathbf{a}^+||\mathbf{d}^+|$) into the ensemble of potential synapses between the pair (i, j) (which has size $d_i^+ a_j^+$) and all the rest. We then choose $c_{ij} < n_{ij}(t_k)$ connections from the preferred ensemble, and the rest of connections from the remaining pool. This is "sampling without replacement" as there is a fixed number of new connections $n(t_k)$ in each time step.

We can now calculate the probability to delete one connection using axonal "negative" elements as $p_{ij}^{a^-}(t_k) = P(\Delta_{a^-}C_{ij}(t_k) = 1|\mathbf{a}(t_k), \mathbf{C}(t_{k-1}))$. This is the probability to choose one to-be-deleted element from all existing elements $p_{ij}^{-a}(t_k) = \frac{a_i^-(t_k)}{a_i(t_k)}$ and $p_{ij}^{d^-}(t_k) = P(\Delta_{d^-}C_{ij}(t_k) = 1|\mathbf{d}(t_k), \mathbf{C}(t_{k-1})) = \frac{d_i^-(t_k)}{d_i(t_k)}$. Out of $a_i(t_k)$ candidates for deletion, we select $a_i^-(t_k)$, subject to the condition not to delete more than $C_{ij}(t_k)$ for this particular pair of neurons. This constraint is reflected by the hypergeometric distribution. The preferred population is represented by the elements bonded into connections from neurons j to neuron i, and the other population is comprised by all remaining elements of neuron j. Finally, during

rewiring events in T_{rewire}, we obtain for the stochastic evolution of $C_{ij}(t_k)$

$$\begin{aligned} C_{ij}(t_k) &= C_{ij}(t_{k-1}) + \Delta_+ C_{ij}(t_k) - \Delta_-^a C_{ij}(t_k) - \Delta_-^b C_{ij}(t_k) \\ \Delta_+ C_{ij}(t_k) &\sim \text{Hypergeometric}(\Delta_+ C_{ij}(t_k) = c | d_i^+(t_k) a_j^+(t_k), |\mathbf{d}^+||\mathbf{a}^+|, \min(|\mathbf{a}^+|, |\mathbf{d}^+|)) \\ \Delta_-^a C_{ij}(t_k) &\sim \text{Hypergeometric}(\Delta_+ C_{ij}(t_k) = c | a_j^-(t_k), a_j(t_k), C_{ij}(t_{k-1})) \\ \Delta_-^b C_{ij}(t_k) &\sim \text{Hypergeometric}(\Delta_+ C_{ij}(t_k) = c | d_i^-(t_k), d_i(t_k), C_{ij}(t_{k-1})). \end{aligned} \tag{8}$$

To calculate the total change $\Delta C_{ij}(t_k)$ we have to account for all these contributions. The distribution of the total increment is not simply the convolution of the three distributions given above, as they are not independent. On the contrary, (negative) decrements occasionally influence (positive) increments, as we first delete connections and thereby create additional free elements. Therefore, the number of free elements has to be corrected as

$$a'^+_i(t_k) = a_i^+(t_k) + \sum_l \Delta_-^b C_{li}(t_k), \qquad b'^+_i(t_k) = d_i^+(t_k) + \sum_l \Delta_-^a C_{il}(t_k).$$

Eq 8 defines a complicated discrete-time stochastic process, but since we are at this point interested only in the expected change of connectivity, we can restrict ourselves to the evolution of expectations. We use $\mathbb{E}_s$ to denote the (linear) operator of expectation over the structural noise, i.e. over the realizations of the increments/decrements $\Delta C_{ij}(t_k)$. We have

$$\mathbb{E}_s[\Delta C_{ij}(t_k)] = \frac{b'^+_i(t_k) a'^+_j(t_k)}{\max(|\mathbf{a}'^+(t_k)|, |\mathbf{b}'^+(t_k)|)} + \mathbb{E}_s[C_{ij}(t_{k-1})] \left(\frac{d_i^-(t_k)}{d_i(t_k)} + \frac{a_j^-(t_k)}{a_j(t_k)} \right). \tag{9}$$

Further on in this paper, for notational convenience, we write $\bar{C}_{ij}(t_k)$ for the expectation $\mathbb{E}_s[C_{ij}(t_k)]$.

Time-continuous limit

We now switch over to continuous equations, which result from the limit $\Delta T_s \to 0$. In this case, we can express free elements and negative elements as

$$a_i^\pm(t) = \left[\frac{1}{\beta_a} (v_i - \phi_i(t)) dt \right]_\pm, \qquad d_i^\pm(t) = \left[\frac{1}{\beta_d} (v_i - \phi_i(t)) dt \right]_\pm,$$

which will assign infinitesimally small values to the numbers of free elements and negative elements. The bracket notation $[\cdot]_\pm$ corresponds to flooring/ceiling operations using Heaviside step function $[x]_\pm = \theta[\pm x]$. Since the rewiring takes place continuously, the numbers of elements are, up to infinitesimal correction, the same as the degrees, $a_i \approx K_i^{\text{in}}$ and $d_i \approx K_i^{\text{out}}$. We can now define the rates of creation and deletion of axonal elements as $\rho_{ai}^{\pm} = \lim_{dt \to 0} \frac{a_i^\pm}{dt}$ and dendritic of elements as $\rho_{di}^{\pm} = \lim_{dt \to 0} \frac{d_i^\pm}{dt}$ and we write explicitly

$$\rho_{ai}^{\pm}(t) = \frac{1}{\beta_a} [v_i - \phi_i(t)]_\pm, \qquad \rho_{di}^{\pm}(t) = \frac{1}{\beta_d} [v_i - \phi_i(t)]_\pm. \tag{10}$$

As noted above, we have omitted here the expectation over the structural noise $\mathbb{E}_s$ from the notation. We can now calculate the evolution of connectivity from Eq 9 as

$$\dot{\bar{C}}_{ij}(t) = \frac{\rho'^+_{di}(t) \rho'^+_{aj}(t)}{\rho(t)} - \bar{C}_{ij}(t) \left(\frac{\rho_{di}^-(t)}{K_i^{in}(t)} + \frac{\rho_{aj}^-(t)}{K_j^{out}(t)} \right), \tag{11}$$

where $\rho(t) = \max(|\boldsymbol{\rho}'^{+}_{a}(t)|, |\boldsymbol{\rho}'^{+}_{d}(t)|)$. The specific implementation requires that the deletion of elements takes place first, with the corresponding synaptic partner remaining available as a free element to form new connections. After an axonal element has been deleted, the previously bonded dendritic element becomes a free element. This requires a correction on the rate of free elements

$$\begin{aligned}\rho'^{+}_{ai}(t) &= \frac{1}{\beta_a}[\nu_i - \phi_i(t)]_+ + \sum_k \frac{\rho^{-}_{dk}(t)}{K^{in}_k(t)}\bar{C}_{ki}(t)\\ \rho'^{+}_{di}(t) &= \frac{1}{\beta_d}[\nu_i - \phi_i(t)]_+ + \sum_k \bar{C}_{ik}(t)\frac{\rho^{-}_{ak}(t)}{K^{out}_k(t)}.\end{aligned}$$

Spiking noise

We assume that spike trains $S_i(t) = \sum_k \delta(t - t^i_k)$ reflect an asynchronous-irregular state of the network, so they can be modeled as a stochastic point process. As described in [62, 63], the coefficient of variation of the spike trains generated by a leaky integrate-and-fire neuron driven by Gaussian white noise current is given by

$$\mathrm{CV}(\mu,\sigma) = 2\pi(\nu_0\tau_m)^2 \int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{x^2}\,dx \int_{-\infty}^{x} e^{y^2}(1+\mathrm{erf}(y))^2\,dy,$$

where μ and σ are the parameters of the current input, and τ_m, V_r and θ are the parameters of the neuron, respectively. The configuration used here yields CV ≈ 0.7. A good approximation for spike trains of a given rate r and irregularity CV is obtained with a specific class of renewal processes, so-called Gamma processes. These have an ISI distribution $f(t) = H(t)\frac{\rho}{\Gamma(\alpha)}(\rho t)^{\alpha-1}e^{-\rho t}$, with parameters $\alpha = \frac{1}{\mathrm{CV}^2}$ and $\rho = \frac{r}{\mathrm{CV}^2}$.

In our model, the fluctuating intracellular calcium concentration is a shotnoise, a continuous signal that arises from a point process through filtering. Here, the point process has a mean rate ν, and the calcium signal is a convolution with an exponential kernel $F(t) = H(t)\frac{1}{\tau_{\mathrm{Ca}}}e^{-\frac{t}{\tau_{\mathrm{Ca}}}}$. Campbell's theorem allows us compute the mean $\mu_{\mathrm{Ca}} = \frac{\mathbb{E}[N_T]}{T}\int_0^t F(s)\,ds = \nu$ and the variance $\sigma^2_{\mathrm{Ca}} = \frac{\mathrm{Var}[N_T]}{T}\int_0^t F(s)^2\,ds \approx \mathrm{CV}^2\frac{\nu}{2\tau_{\mathrm{Ca}}}$ of the calcium variable. Here we used the fact that spike count of Gamma process $N_T = \int_0^T S(t)\,dt$ has a mean $\mathbb{E}[N_T] = \nu T$ and a variance $\mathrm{Var}[N_T] \approx \frac{\sigma^2_{\mathrm{ISI}}}{\mu^3_{\mathrm{ISI}}}T = \frac{\nu}{\alpha}$, provided the observation time T is long enough [64]. As a consequence of the Central Limit Theorem, the amplitude distribution of the calcium signal is approximately Gaussian $\mathcal{N}(\mu_{\mathrm{Ca}}, \sigma^2_{\mathrm{Ca}})$, provided the mean spike rate is much larger than the inverse time constant of the calcium signal $\langle r_i(t)\rangle \gg \frac{1}{\tau_{\mathrm{Ca}}}$. In other words, if the mean firing rate is 8 Hz and the calcium constant is τ_{Ca} = 10 s, there are on average 80 spikes in the characteristic time interval τ_{Ca}.

We are actually interested in the equilibrium rates of free and negative elements in Eq 10. These rates are rectified versions of the shifted calcium trace, and in order to calculate them we resort to the ergodic theorem and to an adiabatic approximation. The former is a self-averaging property, according to which the time-averaged variable for a long observation is equal to the equilibrium mean value $\frac{1}{T}\int_0^T x_i(t)\,dt = \langle x_i\rangle_t$. The adiabatic property comes from

the fact that spiking dynamics of a network is much faster than network remodeling due to structural plasticity. At every point in time, plasticity is driven by the average adiabatic rate. This implies that, at every point in time t, the structural plasticity sees the equilibrium distribution of the calcium trace $P_{\text{eq}}(\phi_i(t)) = \mathcal{N}(\mu_{\text{Ca}}(t), \sigma^2_{\text{Ca}}(t))$. The right-hand side of Eq 10 becomes

$$\begin{aligned} \langle \rho_{ai}^{\pm}(t) \rangle &= \frac{1}{\beta_a} \langle [\nu_i - \phi_i(t)]_{\pm} \rangle = R^{\pm}_{\sigma_a}\left(\frac{\nu_i - \langle r_i(t) \rangle}{\beta_a}\right) \\ \langle \rho_{di}^{\pm}(t) \rangle &= \frac{1}{\beta_d} \langle [\nu_i - \phi_i(t)]_{\pm} \rangle = R^{\pm}_{\sigma_b}\left(\frac{\nu_i - \langle r_i(t) \rangle}{\beta_b}\right). \end{aligned} \tag{12}$$

The transfer function for synaptic elements is given by $R_\sigma(\mu) = \frac{1}{2}\left(\mu + \mu\ \text{erf}\left(\frac{\mu}{\sqrt{2}\sigma}\right) + \sqrt{\frac{2}{\pi}}\sigma e^{-\frac{\mu^2}{2\sigma^2}}\right)$, and the variance of the rate of elements is $\sigma_x^2 = \frac{\eta^2 \nu_i}{2\tau_{\text{Ca}}\beta_x^2}$. The parameter η is a correction factor, which accounts for the regularity of spike trains. In our case η = CV. Even if the mean number of free elements is zero, the noise will still drive the creation and deletion of elements with rate $\frac{\sigma_x}{\sqrt{2\pi}}$. Even if homeostatic control manages to drive all neurons to their target rates, the inherent calcium fluctuations and the associated uncertainty of firing rate inference will still induce random rewiring.

Mean-field approximation of population dynamics

We define population means for variables $x \in \{r, \phi, a, b\}$ and neuronal populations Y, Z as

$$\begin{aligned} x_Y(t) &= \frac{1}{N_Y}\sum_{i \in Y} x_i(t) \\ \bar{C}_{YZ}(t) &= \frac{1}{N_Y N_Z}\sum_{i \in Y}\sum_{j \in Z} \bar{C}_{ij}(t), \end{aligned}$$

where N_Y is the size of population Y. An individual neuron in population Y typically gets many inputs from every other neuronal population, which invites use of mean-field approximation due to the central limit theorem. The resulting currents aggregate to a Gaussian white noise process with mean and variances given as

$$\begin{aligned} \mu_Y(t) &= J\tau \sum_{Z \in \mathcal{E}} \bar{C}_{YZ}(t) N_Z r_Z(t - D) - gJ\tau \bar{C}_I r_I(t - D) + \tau J \bar{C}_E \nu_{ext}, \\ \sigma_Y^2(t) &= J^2\tau \sum_{Z \in \mathcal{E}} \bar{C}_{YZ}(t) N_Z r_Z(t - D) + g^2 J^2 \tau \bar{C}_I r_I(t - D) + \tau J^2 \bar{C}_E \nu_{ext}. \end{aligned}$$

The stationary firing rate of a leaky integrate-and-fire neuron driven by input with mean μ and variance σ^2 is [62]

$$r = f(\mu, \sigma) = \left(\tau_r + \tau_m\sqrt{\pi}\int_{\frac{V_r - \mu}{\sigma}}^{\frac{V_r - \theta}{\sigma}} e^{u^2}(1 + \text{erf}(u))\, du\right)^{-1}.$$

Solving the time-dependent self-consistency problem for multiple interacting plastic

populations is a challenging problem. We suggest here to use an adiabatic approximation, resorting to the fact that the firing rate dynamics is much faster than the plastic growth processes. Therefore, we employ a Wilson-Cowan type of the firing rate dynamics [65]

$$\tau_{\text{rate}}\dot{r}_Y(t) = -r_Y(t) + f(\mu_Y(t), \sigma_Y(t)). \tag{13}$$

The relaxation time is set to $\tau_{\text{rate}} < \tau_m$ to account for the fact that population response is generally much faster than the membrane potential dynamics [66]. The parameter τ_{rate} is the only free parameter in this model, but our results do not depend on its exact value as long as $\tau_{\text{rate}} \leq \tau_m$. The heuristic described by Eq 13 results in a tractable and numerically stable system to be analyzed with standard dynamical system tools.

As the equation is linear, the average calcium trace can be computed without approximation

$$\tau_{\text{Ca}}\dot{\phi}_Y(t) = r_Y(t) - \phi_Y(t). \tag{14}$$

The rate of element creation and deletion is also linear, therefore population averages are given by

$$\rho_{dY}(t) = \frac{1}{\beta_d}(\nu_Y - \phi_Y(t)) \qquad \rho_{aY}(t) = \frac{1}{\beta_a}(\nu_Y - \phi_Y(t)).$$

Finally we calculate the average connectivity

$$\dot{\bar{C}}_{Y,Z}(t) = \frac{\rho'^{+}_{dY}(t)\rho'^{+}_{aZ}(t)}{\rho(t)} + \bar{C}_{Y,Z}(t)\left(\frac{\rho^{-}_{dY}(t)}{k^i_Y(t)} + \frac{\rho^{-}_{aZ}(t)}{k^o_Y(t)}\right). \tag{15}$$

Here, $\rho^{\pm}_{dY}(t) = R_\sigma(\pm\rho^{\pm}_{dY}(t))$, $\rho^{\pm}_{aY}(t) = R_\sigma(\pm\rho^{\pm}_{aY}(t))$, and $\rho(t)$ is same as before, while the corrected rate for free elements is

$$\begin{aligned}\rho'^{+}_{aY}(t) &= \frac{1}{\beta_d}(\nu_Y - \phi_Y(t)) + \sum_{Z\in\mathcal{E}}\frac{\rho^{-}_{dZ}(t)}{K^{in}_Z(t)}N_Z\bar{C}_{\text{pop}_k,Y}(t)\\ \rho'^{+}_{dY}(t) &= \frac{1}{\beta_a}(\nu_Y - \phi_Y(t)) + \sum_{Z\in\mathcal{E}}\bar{C}_{YZ}(t)N_Z\frac{\rho^{-}_{ak}(t)}{K^{out}_Z(t)}.\end{aligned}$$

Line attractor of the deterministic system

We represent the state of the network consisting of one static inhibitory and two plastic excitatory ensembles as a vector

$$\mathbf{y}(t) = (\phi_{E_1}(t), \phi_{E_2}(t), \bar{C}_{E_1E_1}(t), \bar{C}_{E_1E_2}(t), \bar{C}_{E_2E_1}(t), \bar{C}_{E_2E_2}(t), r_{E_1}(t), r_{E_2}(t), r_I(t))^T,$$

the components of which adhere to the calcium dynamics Eq 14, the connectivity dynamics Eq 15, and the activity dynamics Eq 13. The joint ODE system $\frac{d\mathbf{y}}{dt} = \mathbf{F}(\mathbf{y}, t)$ defines the vector field $\mathbf{F}$. We first explore the stationary states of the deterministic system ($\sigma_x = 0$), setting the left hand side of all ODEs to zero. In this case, calcium concentration and excitatory firing rates are fixed at their target values, and inhibitory firing rates r^*_I can be obtained using the self-consistency Eq 13. Stationary connectivity is calculated from the condition that the rates of creation and deletion of synaptic elements are zero. This implies that the numbers of free axonal and dendritic elements are zero, $\mathbf{a}^{\pm} = \mathbf{0}$ and $\mathbf{d}^{\pm} = \mathbf{0}$, which in turn implies

that $\mathbf{a} = \mathbf{1}^T \cdot \mathbf{C}$ and $\mathbf{d} = \mathbf{C} \cdot \mathbf{1}$. The second condition is that all axonal elements are bonded with a dendritic element, and the total number of both types of elements are the same $|\mathbf{a}| = |\mathbf{d}|$.

Let us consider the case of two plastic excitatory ensembles E_1 and E_2 and one static inhibitory population I. Using the parameter $x = \bar{C}_{E_1E_1}$, the stationary state of the deterministic network is a line

$$\mathbf{L}(x) = \begin{bmatrix} \phi^*_{E1} \\ \phi^*_{E2} \\ \bar{C}^*_{E_1E_1} \\ \bar{C}^*_{E_1E_2} \\ \bar{C}^*_{E_2E_1} \\ \bar{C}^*_{E_2E_2} \\ r^*_{E1} \\ r^*_{E2} \\ r^*_I \end{bmatrix} = \begin{bmatrix} v \\ v \\ x \\ \frac{K^*_{E1}}{N_{E2}} - x\frac{N_{E1}}{N_{E2}} \\ \frac{K^*_{E1}}{N_{E2}} - x\frac{N_{E1}}{N_{E2}} \\ \left(\frac{K^*_{E2}}{N_{E2}} - \frac{N_{E1}}{N_{E2}}\frac{K^*_{E1}}{N_{E2}}\right) + x\left(\frac{N_{E1}}{N_{E2}}\right)^2 \\ v \\ v \\ r^*_I \end{bmatrix}. \tag{16}$$

Here, K^*_Y is the mean stationary in-degree, and we assume that the dendritic element growth factor β_d is less or equal to the axonal element growth factor β_a. The attractor is a line segment, as a consequence of linear conditions for element numbers and non-negativity of excitatory connections $\bar{C}_{ij} \geq 0$. This solution can be easily generalized to a solution for individual connections of N_E excitation neurons, or for the case of n_E excitatory populations. The minimal invariant changes of the stationary connectivity matrix $\mathbf{C}$ that keep in-degree and out-degree conditions valid are of the type

$$\begin{pmatrix} \ddots & & & \iddots \\ & 1 & \cdots & -1 & \\ & \vdots & \ddots & \vdots & \\ & -1 & \cdots & 1 & \\ \iddots & & & & \ddots \end{pmatrix},$$

where all missing entries are considered to be zero. This transformation defines a hyperplane

section invariant space. The stationary connectivity can be solved as

$$
\begin{aligned}
|\mathbf{a}| &= |\mathbf{d}| \\
\mathbf{C} &= \begin{pmatrix}
\bar{C}_{11} & \cdots & \bar{C}_{1(N_E-1)} & d_1 - \sum_{k=1}^{N_E-1} \bar{C}_{1k} \\
\vdots & \ddots & \vdots & \vdots \\
\bar{C}_{(N_E-1)1} & \cdots & \bar{C}_{(N_E-1)(N_E-1)} & d_{N_E-1} - \sum_{k=1}^{N_E-1} \bar{C}_{(N_E-1)k} \\
a_1 - \sum_{k=1}^{N_E-1} \bar{C}_{k1} & \cdots & a_{N_E-1} - \sum_{k=1}^{N_E-1} \bar{C}_{k(N_E-1)} & d_n - \sum_{k=1}^{N_E-1} a_k + \sum_{l=1,m=1}^{N_E-1} \bar{C}_{lm}
\end{pmatrix}
\end{aligned}
\tag{17}
$$

Note that we have put the in-degree conditions in the last row and the out-degree conditions in the last column, but there is a permutation symmetry and they could apply to any row-column combination. The hyperplane for individual neurons has $N_E(N_E - 3) + 1$ dimensions, if there are no self-connections. In the case of n_E excitatory populations, the hyperplane has $n_E(n_E - 2) + 1$ dimensions.

Slow manifold and diffusion to the global fixed point

Now we calculate the stationary state for the stochastic system and assess the geometry and stability of the underlying phase space. In the case of finite noise, $\sigma_d > 0$, Eq 11 has an attractive fixed point, instead of a hyperplane attractor. The fixed point is given by

$$
\bar{C}^*_{i,j} = \frac{{\rho'_{di}}^{+} {\rho'_{aj}}^{+}}{\rho({\rho_{di}}^{-} K_j^{\text{out}} + {\rho_{aj}}^{-} K_j^{in})} K_i^{in} K_j^{\text{out}}. \tag{18}
$$

Here, without loss of generality, we have assumed that the growth rate of dendritic elements is smaller or equal as compared to axonal elements. We obtain a normalized outer product of the indegree and outdegree vectors, respectively. In the case of identical dynamics for axonal and dendritic elements, and if all neurons have the same target rate $\nu_i = \nu$, we obtain the fixed point by plugging in the homogeneous solution

$$
\bar{C}^*_{i,j} = \frac{K^{\text{in}}}{N_E} = c. \tag{19}
$$

This fixed point depends only on the indegree of neurons. For a two-population system, the fixed point is $\mathbf{y} = \mathbf{L}\left(\frac{K^{\text{in}}}{N_E}\right)$, the most entropic configuration of the line segment attractor. This suggests a special importance for the relic of the line-attractor, which, as we show later, turns into a stochastic slow manifold through the presence of noise.

We now calculate the relaxation time for the three-population system to assess the persistence of a memory trace. To this end, we calculate the Jacobian $\mathcal{J} = \frac{\partial \mathbf{F}(\mathbf{y},t)}{\partial \mathbf{y}}\Big|_{\mathbf{y}=\mathbf{L}(x)}$ on the slow manifold, using a linearization of the vector field $\mathbf{F}$ (see Section Line attractor of the

deterministic system). For the stochastic system, the Jacobian $\mathcal{J}(\mathbf{C}^*)$ about the global fixed point is

$$\mathcal{J} = \begin{bmatrix} -\frac{1}{\tau_{\mathrm{Ca}}} & 0 & 0 & 0 & 0 & 0 & \frac{1}{\tau_{\mathrm{Ca}}} & 0 & 0 \\ 0 & -\frac{1}{\tau_{\mathrm{Ca}}} & 0 & 0 & 0 & 0 & 0 & \frac{1}{\tau_{\mathrm{Ca}}} & 0 \\ J_{31} & J_{32} & J_{33} & J_{34} & J_{35} & J_{36} & 0 & 0 & 0 \\ J_{41} & J_{42} & J_{43} & J_{44} & J_{45} & J_{46} & 0 & 0 & 0 \\ J_{51} & J_{52} & J_{53} & J_{54} & J_{55} & J_{56} & 0 & 0 & 0 \\ J_{61} & J_{62} & J_{63} & J_{64} & J_{65} & J_{66} & 0 & 0 & 0 \\ 0 & 0 & J_{73} & J_{74} & 0 & 0 & J_{77} & J_{78} & J_{79} \\ 0 & 0 & 0 & 0 & J_{85} & J_{86} & J_{87} & J_{88} & J_{89} \\ 0 & 0 & 0 & 0 & 0 & 0 & J_{97} & J_{98} & J_{99} \end{bmatrix}. \tag{20}$$

The Jacobian has a block structure, which corresponds to calcium activity (black rectangle in Eq 20), connectivity $\mathcal{J}_c$ (blue rectangle) and spike activity $\mathcal{J}_r$ (red rectangle). The connectivity block around fixed point $\mathbf{y} = \mathbf{L}(c)$ is given by

$$\mathcal{J}_c = c_0 \begin{bmatrix} -N_{E2}^2 & N_{E2}^2 & N_{E2}^2 & -N_{E2}^2 \\ N_{E1}N_{E2} & -N_{E1}N_{E2} & -N_{E1}N_{E2} & N_{E1}N_{E2} \\ N_{E1}N_{E2} & -N_{E1}N_{E2} & -N_{E1}N_{E2} & N_{E1}N_{E2} \\ -N_{E1}^2 & N_{E1}^2 & N_{E1}^2 & -N_{E1}^2 \end{bmatrix},$$

where $c_0 = \frac{\eta\sqrt{\frac{\nu}{\pi\tau_{\mathrm{Ca}}}}}{2cN_E^3}\left(\frac{1}{\beta_a}+\frac{1}{\beta_b}\right)$. Jacobian entries responsible for interaction between connectivity variables and calcium activity variables are $-J_{42} = -J_{52} = J_{61} = \frac{N_{E1}}{N_E^2\beta_b}$, $J_{32} = -J_{41} = -J_{51} = \frac{N_{E2}}{N_E^2\beta_b}$, $J_{31} = -\frac{n_{E1}+2N_{E2}}{N_E^2\beta_b}$ and $J_{62} = -\frac{2N_{E1}+N_{E2}}{N_E^2\beta_b}$. The spike activity terms in the Jacobian are

$$\mathcal{J}_r = \frac{1}{\tau_{\mathrm{rate}}} \begin{bmatrix} -1+j_{EE}\bar{C}_{E_1E_1}N_{E1} & j_{EE}\bar{C}_{E_1E_2}N_{E2} & j_{EI}\epsilon N_I \\ j_{EE}\bar{C}_{E_2E_1}N_{E1} & -1+j_{EE}\bar{C}_{E_2E_2}N_{E2} & j_{EI}\epsilon N_I \\ j_{IE}\epsilon N_{E1} & J_{IE}\epsilon N_{E1} & -1+j_{II}\epsilon N_I \end{bmatrix}$$

Here we define the effective interaction of different neuron types neurons as excitatory-to-excitatory $j_{EE} = J\tau\frac{\partial f(\mu_E,\sigma_E)}{\partial\mu_E} + \frac{J^2\tau}{\sigma_E}\frac{\partial f(\mu_E,\sigma_E)}{\partial\sigma_E}$, inhibitory-to-excitatory $j_{EI} = -gJ\tau\frac{\partial f(\mu_E,\sigma_E)}{\partial\mu_E} + \frac{J^2\tau}{\sigma_E}\frac{\partial f(\mu_E,\sigma_E)}{\partial\sigma_E}$, excitatory-to-inhibitory $j_{IE} = J\tau\frac{\partial f(\mu_I,\sigma_I)}{\partial\mu_I} + \frac{J^2\tau}{\sigma_I}\frac{\partial f(\mu_I,\sigma_I)}{\partial\sigma_I}$ and inhibitory-to-inhibitory neurons as $j_{II} = -gJ\tau\frac{\partial f(\mu_I,\sigma_I)}{\partial\mu_I} + \frac{g^2J^2\tau}{\sigma_I}\frac{\partial f(\mu_I,\sigma_I)}{\partial\sigma_I}$. The Jacobian elements which are responsible for influence of connectivity change to rates are $J_{73} = \frac{1}{\tau_{\mathrm{rate}}}j_{EE}N_{E1}r_{E1}$, $J_{74} = \frac{1}{\tau_{\mathrm{rate}}}j_{EE}N_{E2}r_{E2}$, $J_{85} = \frac{1}{\tau_{\mathrm{rate}}}j_{EE}N_{E1}r_{E1}$ and $J_{86} = \frac{1}{\tau_{\mathrm{rate}}}j_{EE}N_{E2}r_2$. These terms are the largest in the Jacobian, but since connectivity changes only through a change of calcium, they will not play an essential role for long-term stability.

We exploit the block structure of the Jacobian $\mathcal{J}$ to separate between connectivity variables, the fast firing rate and calcium variables. We can solve the connectivity eigenproblem

analytically,

$$\boldsymbol{\lambda}_0^0 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{\lambda}_0^1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -\frac{N_{E1}}{N_{E2}} \\ -\frac{N_{E1}}{N_{E2}} \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{\lambda}_0^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -\frac{N_{E1}}{N_{E2}} \\ 1 \\ -\frac{N_{E1}}{N_{E2}} \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \boldsymbol{\lambda}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -\frac{N_{E1}}{N_{E2}} \\ -\frac{N_{E1}}{N_{E2}} \\ \left(\frac{N_{E1}}{N_{E2}}\right)^2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{21}$$

with eigenvalues $\lambda_0 = 0$ and $\lambda_1 = -\frac{R_{\sigma_a}(0)+R_{\sigma_b}(0)}{cN_E}$.

The persistence of a memory trace in the system with noise is essentially determined by the diffusion to the global fixed point. In fact, the line attractor of the noiseless system turns into a slow manifold of the noisy system. The eigenvalue λ_0 is degenerate, and the three-dimensional invariant space defines the central manifold of the connectivity subsystem. The eigenvalue λ_1 is also an eigenvalue of the full Jacobian $\mathcal{J}$ and it represents the slowest time scale of the system in the direction of the slow manifold. It is easy to check that $\partial_x \mathbf{L}(x) = \lambda_1$. This means that the noiseless line attractor is exactly corresponding to the slow manifold of the system. The other eigenvectors have components orthogonal to the slow manifold. Those are dominated by fast variables, and they relax quickly. As a result, the relaxation dynamics is dominated by the eigenvalue λ_1, and the relaxation time of the system under consideration is

$$\tau_{\text{diffusion}} = \sqrt{\frac{4\pi\tau_{\text{Ca}}}{\eta^2 \nu}} \frac{N_E c}{\frac{1}{\beta_d} + \frac{1}{\beta_a}}.$$

For the default parameters used here, its value is around 5000 s, fully in accordance with numerical simulations. The system relaxes along the direction of the slow manifold. Therefore, although rates of creation and deletion are identical on the slow manifold, the system takes a fast trajectory off the slow manifold upon perturbation, and it relaxes slowly along the slow manifold towards the fixed point.

A quantitative measure for the volatility of network structure used in experiments is the turnover ratio (TOR) of dendritic spines [38]. It is defined as $\text{TOR} = \frac{\Delta N_{\text{new}} + \Delta N_{\text{deleted}}}{2N_{\text{spines}}}$, where the changes ΔN_{new} and $\Delta N_{\text{deleted}}$ are typically measured per day. For the case when the network plasticity is selectively driven by diffusion, this quantity corresponds to the eigenvalue λ_1, corresponding to a value around 18% per day for standard parameters.

Linear stability analysis

Numerical exploration of the eigenvalues of the Jacobian $\mathcal{J}$ (Eq 20) show that for our system there is linear stability for a wide parameter range. However this system can produce damped

oscillations, which may cause problems. In the case discussed here, dendritic and axonal elements have identical dynamics, $\beta_a = \beta_d$, and damped oscillations influence the connectivity block in the same way. Therefore it was sufficient to consider a reduced system of only one plastic excitatory population and one static inhibitory population to analyze the influence of the calcium time constant and element growth. This gives rise to the Jacobian

$$\mathcal{J}_1 = \begin{bmatrix} -\frac{1}{\tau_{\mathrm{Ca}}} & 0 & \frac{1}{\tau_{\mathrm{Ca}}} & 0 \\ -\frac{1}{N_E \beta_d} & 0 & 0 & 0 \\ 0 & \frac{1}{\tau_{\mathrm{rate}}} j_{EE} c N_E & -\frac{1}{\tau_{\mathrm{rate}}} + \frac{1}{\tau_{\mathrm{rate}}} j_{EE} c N_E & \frac{1}{\tau_{\mathrm{rate}}} j_{EI} \epsilon N_I \\ 0 & 0 & \frac{1}{\tau_{\mathrm{rate}}} J_{IE} \epsilon N_E & -\frac{1}{\tau_{\mathrm{rate}}} + \frac{1}{\tau_{\mathrm{rate}}} j_{II} \epsilon N_I \end{bmatrix}.$$

The eigenvalues of $\mathcal{J}_1$ have been reduced to radicals using *Mathematica* 12.0, and the complex conjugated eigenvalues $\lambda_{3/4}$ responsible for oscillations are plotted in Fig 6.

Although there is linear stability around the global fixed points for wide parameter range, the same is not the true for all points along the line attractor. The connectivity eigenvalue λ_1 is constant on the slow manifold, but if one population presents large recurrent connectivity, spike activity jumps to a persistent high activity state. We can track this instability in the Jacobian $\mathcal{J}(x)$ along the line attractor $\mathbf{L}(x)$. As the recurrent connectivity of the first ensemble $E1$ is increased, one of the eigenvalues becomes positive, and its corresponding eigenvector makes the biggest contributions in the direction of spike activity r_{E1} and r_{E2}. Here we use this fact and use the reduced Jacobian

$$\mathcal{J}_r = \begin{bmatrix} \frac{\partial F_7(\mathbf{y},t)}{\partial r_{E1}} & \frac{\partial F_7(\mathbf{y},t)}{\partial r_{E2}} \\ \frac{\partial F_8(\mathbf{y},t)}{\partial r_{E1}} & \frac{\partial F_8(\mathbf{y},t)}{\partial r_{E2}} \end{bmatrix} = \frac{1}{\tau_{\mathrm{rate}}} \begin{bmatrix} -1 + j_{EE} \bar{C}_{E_1E_1} N_{E1} & j_{EE} \bar{C}_{E_1E_2} N_{E2} \\ j_{EE} \bar{C}_{E_2E_1} N_{E1} & -1 + j_{EE} \bar{C}_{E_2E_2} N_{E2} \end{bmatrix}$$

to find the approximate position of this transition on the line attractor. The spiking activity becomes unstable, when the real part of the second eigenvalue λ_2^r becomes positive, and at this point the determinant of the Jacobian changes its sign (since both eigenvalues are real). We can use this criterion to determine when the system loses linear stability, leading to the critical value c_{crit} of connectivity $\bar{C}_{E_1E_1}$

$$c_{\mathrm{crit}} = \epsilon \left(1 + \frac{N_{E2}}{N_{E1} j_E}\right), \tag{22}$$

where j_E is a dimension-less measure of effective excitatory excitability $j_E = \epsilon N_E j_{EE}$. The value c_{crit} is the upper bound of full stability, but the system loses stability even below this value, as discussed in our results.

Network simulations

All simulations have been performed using the neural network simulator NEST 2.16.0 [67].

Grown networks. All numerical stimulation experiments start from grown networks. To this end, we initialize a network with random connections to and from inhibitory neurons and no excitatory-to-excitatory connections, whatsoever. The latter are then grown under the control of homeostatic structural plasticity. The controlled variable is the firing rate of excitatory neurons, the target rate is set to $\nu = 8$ Hz for all neurons. During this initial growth period, excitatory neurons receive external Poisson input of rate $\nu_{\mathrm{ext}} = 15$ kHz. After a long-enough

growth time t_{growth}, the network structure has reached its equilibrium and all neurons fire at their target rate, apart from small fluctuations.

Conditioning paradigm. Non-overlapping neuronal ensembles comprising 10% of all excitatory neurons each are selected and labeled US, C1 and C2. Non-plastic connections of weight $J_E = 0.1$ mV are created from all neurons labeled US to a readout neuron, which has the same properties as all the other neurons in the network. Stimulation is specific for a certain group of neurons. A stimulation cycle consists of an increased external input rate of $1.4\nu_{\text{ext}}$ for a period of 2 s, followed by a relaxation period of 48 s, during which the external input is set back to ν_{ext}.

The whole protocol consists of 5 different episodes: growth, baseline, encoding, decay and retrieval. During "baseline" each of the 3 groups is stimulated alone in the order US, C1, C2. The "encoding" episode consists of 6 stimulation cycles. In 3 out of the 6 cycles, C2 is stimulated alone. In the other 3 cycles, neurons from both US and C1 are stimulated at the same time. The "decay" episode lasts 100 s, during which no stimulation beyond ν_{ext} is applied. During "retrieval", there are 2 stimulation cycles, C1 alone is followed by C2 alone. Connectivity is recorded every 1 s. In this protocol, plasticity is always on, and all measurements are performed in the plastic network. Growth time is $t_{\text{growth}} = 100$ s, the remaining parameters are $\beta_d = \beta_a = 0.4$, $\tau_{\text{Ca}} = 1$ s, $\Delta T_s = 10$ ms.

Repeated stimulation. Starting from a grown network, a random neuronal ensemble comprising 10% of all excitatory neurons in the network is repeatedly stimulated for 8 cycles. Here, a stimulation cycle consists of a stimulation period of 150 s, during which the external input to the ensemble neurons is increased to $1.05\ \nu_{\text{ext}}$. It is followed by a relaxation period of 150 s, during which the external input is set back to ν_{ext}. During stimulation, the connectivity is recorded every 15 s. After encoding the engram, plasticity is turned off, and all measurements are now performed in a non-plastic network. Growth time is $t_{\text{growth}} = 500$ s, the remaining parameters are $\beta_d = \beta_a = 2$, $\tau_{\text{Ca}} = 10$ s, $\Delta T_s = 100$ ms.

Readout neuron. Two non-overlapping ensembles comprising 10% of all excitatory neurons each are randomly selected and labeled A_1 and A_2. Starting from a grown network, A_1 is stimulated twice. During each stimulation cycle, the external input to stimulated neurons is increased to $1.1\ \nu_{\text{ext}}$ for a time period of 150 s. This is followed by a relaxation period of 150 s, during which the external input is set back to ν_{ext}. After a pause of 100 s, A_2 is stimulated once using otherwise the same protocol. Growth time is $t_{\text{growth}} = 500$ s, the remaining parameters are $\beta_d = \beta_a = 2$, $\tau_{\text{Ca}} = 10$ s, $\Delta T_s = 100$ ms.

After the encoding of engrams, plasticity is turned off. A readout neuron is added to the network, which has the same properties as all other neurons in the network. Non-plastic connections of weight $J_E = 0.1$ mV are created from a random sample comprising 9% of all excitatory and 9% of all inhibitory neurons in the network. Two new non-overlapping ensembles comprising 10% of all excitatory neurons each are selected as random patterns A_3 and A_4. Neurons in the network are stimulated in the order A_3, A_2, A_1, A_4. During each stimulation cycle, the external input to neurons in the corresponding group is increased to $1.1\nu_{\text{ext}}$ for a period of 1 s duration. This is followed by a non-stimulation period of 4 s, during which external input rate is set back to ν_{ext}.

Formation and decay of engrams. Starting from a grown network, a subgroup comprising 10% of all excitatory neurons is randomly selected and stimulated for 150 s, followed by a prolonged relaxation period of 5500 s. During stimulation, the external input to stimulated neurons is increased to $1.1\nu_{\text{ext}}$. For all simulations, $\beta_d = \beta_a = 2$, $\Delta T_s = 100$ ms.

The simulations to demonstrate how τ_{decay} changes with τ_{Ca} were performed with a target rate $\nu = 8$ Hz and $\tau_{\text{Ca}} = 2, 4, 8, 16, 32$ s. The simulations showing how τ_{decay} changes with ν were performed with $\tau_{\text{Ca}} = 10$ s and $\nu = 2, 4, 8, 16, 32$ Hz. In the simulations performed with

ν = 2, 4 Hz, τ_{growth} = 5000 s, neurons are stimulated for 1500 s, and the relaxation period is 26000 s. The parameter τ_{decay} is estimated from simulated time series by performing a least-squares fit of an exponential function to the connectivity values during the decay period, and extracting the fitted time constant.

Linear stability. Starting from a grown network, a subgroup comprising 10% of all excitatory neurons is randomly selected and stimulated for a time period t_{stim}. During stimulation, the external input to stimulated neurons is increased to $1.1\nu_{\text{ext}}$. Stimulation is followed by a period of duration t_{relax}, during which the external input is set back to ν_{ext} and the connectivity relaxes back to a new equilibrium. The parameters used are for non-oscillatory regime: $t_{\text{growth}} = t_{\text{stim}} = t_{\text{relax}}$ = 800 s, $\beta_d = \beta_a$ = 2, τ_{Ca} = 5 s, ΔT_s = 1 ms; for the weakly oscillatory regime: $t_{\text{growth}} = t_{\text{stim}} = t_{\text{relax}}$ = 400 s, $\beta_d = \beta_a$ = 0.2, τ_{Ca} = 20 s, ΔT_s = 0.2 ms; and for the strongly oscillatory regime: $t_{\text{growth}} = t_{\text{stim}} = t_{\text{relax}}$ = 200 s, $\beta_d = \beta_a$ = 0.03, τ_{Ca} = 10 s, ΔT_s = 0.2 ms.

Non-linear stability. Starting from a grown network, a subgroup comprising 10% of all excitatory neurons is randomly selected and stimulated for 200 s. During stimulation, the external input rate to stimulated neurons is increased to 1.25 ν_{ext}. Stimulation is followed by a relaxation time in which external input rate is set back to ν_{ext}. Growth time t_{growth} = 500 s, $\beta_d = \beta_a$ = 2, τ_{Ca} = 10 s, ΔT_s = 0.1 ms. During and after stimulation, connectivity is recorded every 5 s.

Overlap measure. The similarity between network activity during stimulation, spontaneous and evoked responses is measured by the corresponding overlaps [68] defined as

$$m^{\mu} = [Na(1-a)]^{-1} \sum_i (\xi_i^{\mu} - a) s_i(t). \tag{23}$$

The pattern ξ^{μ} is a vector of dimension N. Each entry ξ_i^{μ} is a binary variable indicating whether or not neuron i is stimulated by pattern ξ^{μ}, and it has a mean value of $a = \langle \xi_i^{\mu} \rangle_i$. The activity vector $s_i(t)$ is also composed of binary variables, which indicate whether or not neuron i is active in a given time bin. In all figures, the bin size used for calculating overlaps is 10 ms.

Population firing rate. In Fig 3B, the population response of excitatory neurons is estimated for different connectivity values using a mean-field rate model [62]. We consider a model with three populations, two of which are excitatory (E_1 comprises 10% and E_2 comprises 90% of all N_E = 10000 excitatory neurons), and one is inhibitory with N_I = 2500 neurons. All connectivities involving inhibitory neurons are fixed and set to ϵ = 0.1. For the excitatory-to-excitatory connections, we systematically vary the connectivity within the E_1 population ($\bar{C}_{E_1E_1}$), and calculate the other values to achieve a constant excitatory in-degree of ϵN_E. All other parameters used are unchanged. For the different values of $\bar{C}_{E_1E_1}$ we calculate the population rate of excitatory neurons for E_1 receiving a larger external input of 1.05 ν_{ext} as compared to the E_2 and I populations (ν_{ext} = 15 kHz).

Pattern completion. To address pattern completion, we employ the non-plastic network after engram encoding (see Section Repeated stimulation). Different fractions of the neurons belonging to the engram are stimulated for 10 s, and we calculate the overlap averaged over the stimulation time $\langle m^{E_1} \rangle$. For each fraction of stimulated neurons, 50 different simulations are run, during which a different subsample of the engram neurons is stimulated.

Acknowledgments

We thank Sandra Diaz-Pier and Mikaël Naveau from the Research Center Jülich for support on new features of NEST, and Uwe Grauer from the Bernstein Center Freiburg as well as Bernd Wiebelt and Michael Janczyk from the Freiburg University Computing Center for their assistance with HPC applications.

Author Contributions

Conceptualization: Júlia V. Gallinaro, Nebojša Gašparović, Stefan Rotter.

Formal analysis: Júlia V. Gallinaro, Nebojša Gašparović, Stefan Rotter.

Investigation: Júlia V. Gallinaro, Nebojša Gašparović, Stefan Rotter.

Project administration: Stefan Rotter.

Software: Júlia V. Gallinaro, Nebojša Gašparović.

Supervision: Stefan Rotter.

Visualization: Júlia V. Gallinaro, Nebojša Gašparović, Stefan Rotter.

Writing – original draft: Júlia V. Gallinaro, Nebojša Gašparović, Stefan Rotter.

Writing – review & editing: Júlia V. Gallinaro, Nebojša Gašparović, Stefan Rotter.

References

1. Takeuchi T, Duszkiewicz AJ, Morris RGM. The synaptic plasticity and memory hypothesis: encoding, storage and persistence. Philosophical Transactions of the Royal Society B: Biological Sciences. 2014; 369 (1633). https://doi.org/10.1098/rstb.2013.0288 PMID: 24298167
2. Litwin-Kumar A, Doiron B. Formation and maintenance of neuronal assemblies through synaptic plasticity. Nature Communications. 2014; 5:5319. https://doi.org/10.1038/ncomms6319 PMID: 25395015
3. Zenke F, Agnes EJ, Gerstner W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nature Communications. 2015; 6. https://doi.org/10.1038/ncomms7922 PMID: 25897632
4. Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. Science. 2011; 334(6062):1569–73. https://doi.org/10.1126/science.1211095 PMID: 22075724
5. Ocker GK, Doiron B. Training and spontaneous reinforcement of neuronal assemblies by spike timing plasticity. Cerebral Cortex. 2019; 29(3):937–951. https://doi.org/10.1093/cercor/bhy001 PMID: 29415191
6. Yger P, Gilson M. Models of Metaplasticity: A Review of Concepts. Frontiers in Computational Neuroscience. 2015; 9. https://doi.org/10.3389/fncom.2015.00138 PMID: 26617512
7. Litwin-Kumar A, Doiron B. Slow dynamics and high variability in balanced cortical networks with clustered connections. Nature Neuroscience. 2012; 15(11):1498–1505. https://doi.org/10.1038/nn.3220 PMID: 23001062
8. Holtmaat A, Svoboda K. Experience-dependent structural synaptic plasticity in the mammalian brain. Nature Reviews Neuroscience. 2009; 10(9):647–58. https://doi.org/10.1038/nrn2699 PMID: 19693029
9. Caroni P, Donato F, Muller D. Structural plasticity upon learning: regulation and functions. Nature Reviews Neuroscience 2012 13:7. 2012; 13(7):478–490. https://doi.org/10.1038/nrn3258 PMID: 22714019
10. Loewenstein Y, Yanover U, Rumpel S. Predicting the dynamics of network connectivity in the neocortex. Journal of Neuroscience. 2015; 35(36):12535–12544. https://doi.org/10.1523/JNEUROSCI.2917-14.2015 PMID: 26354919
11. Mongillo G, Rumpel S, Loewenstein Y. Intrinsic volatility of synaptic connections—a challenge to the synaptic trace theory of memory. Current Opinion in Neurobiology. 2017; 46:7–13. https://doi.org/10.1016/j.conb.2017.06.006
12. Fauth MJ, van Rossum MC. Self-organized reactivation maintains and reinforces memories despite synaptic turnover. eLife. 2019; 8:e43717. https://doi.org/10.7554/eLife.43717 PMID: 31074745
13. Fauth M, Wörgötter F, Tetzlaff C. Formation and Maintenance of Robust Long-Term Information Storage in the Presence of Synaptic Turnover. PLOS Computational Biology. 2015; 11(12):e1004684. https://doi.org/10.1371/journal.pcbi.1004684 PMID: 26713858
14. Turrigiano G. Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function. Cold Spring Harbor Perspectives in Biology. 2012; 4(1):a005736. https://doi.org/10.1101/cshperspect.a005736 PMID: 22086977

15. Hengen KB, Torrado Pacheco A, McGregor JN, Van Hooser SD, Turrigiano GG. Neuronal Firing Rate Homeostasis Is Inhibited by Sleep and Promoted by Wake. Cell. 2016; 165(1):180–191. https://doi.org/10.1016/j.cell.2016.01.046 PMID: 26997481

16. Torrado Pacheco A, Tilden EI, Grutzner SM, Lane BJ, Wu Y, Hengen KB, et al. Rapid and active stabilization of visual cortical firing rates across light–dark transitions. Proceedings of the National Academy of Sciences of the United States of America. 2019; p. 201906595. https://doi.org/10.1073/pnas.1906595116 PMID: 31366632

17. Zenke F, Gerstner W. Hebbian plasticity requires compensatory processes on multiple timescales. Philosophical Transactions of the Royal Society B. 2017; 372(1715):20160259. https://doi.org/10.1098/rstb.2016.0259 PMID: 28093557

18. Keck T, Toyoizumi T, Chen L, Doiron B, Feldman DE, Fox K, et al. Integrating Hebbian and homeostatic plasticity: the current state of the field and future research directions. Philosophical Transactions of the Royal Society B. 2017; 372(1715):20160158. https://doi.org/10.1098/rstb.2016.0158 PMID: 28093552

19. Gallinaro JV, Rotter S. Associative properties of structural plasticity based on firing rate homeostasis in recurrent neuronal networks. Scientific Reports. 2018; 8(1):3754. https://doi.org/10.1038/s41598-018-22077-3 PMID: 29491474

20. Lu H, Gallinaro JV, Rotter S. Network remodeling induced by transcranial brain stimulation: A computational model of tDCS-triggered cell assembly formation. Network Neuroscience. 2019; 3(4):924–943. https://doi.org/10.1162/netn_a_00097 PMID: 31637332

21. Butz M, Lehmann K, Dammasch IE, Teuchert-Noodt G. A theoretical network model to analyse neurogenesis and synaptogenesis in the dentate gyrus. Neural Networks. 2006; 19(10):1490–1505. https://doi.org/10.1016/j.neunet.2006.07.007 PMID: 17014989

22. Butz M, Teuchert-Noodt G, Grafen K, van Ooyen A. Inverse relationship between adult hippocampal cell proliferation and synaptic rewiring in the dentate gyrus. Hippocampus. 2008; 18(9):879–898. https://doi.org/10.1002/hipo.20445 PMID: 18481284

23. Butz M, Steenbuck ID, van Ooyen A. Homeostatic structural plasticity can account for topology changes following deafferentation and focal stroke. Frontiers in Neuroanatomy. 2014; 8:115. https://doi.org/10.3389/fnana.2014.00115 PMID: 25360087

24. Butz M, Van Ooyen A, Wörgötter F. A model for cortical rewiring following deafferentation and focal stroke. Frontiers in Computational Neuroscience. 2009; 3(August):10. https://doi.org/10.3389/neuro.10.010.2009 PMID: 19680468

25. Butz M, van Ooyen A. A simple rule for dendritic spine and axonal bouton formation can account for cortical reorganization after focal retinal lesions. PLOS computational biology. 2013; 9(10):e1003259. https://doi.org/10.1371/journal.pcbi.1003259 PMID: 24130472

26. Tetzlaff C, Okujeni S, Egert U, Wörgötter F, Butz M. Self-Organized Criticality in Developing Neuronal Networks. PLOS Computational Biology. 2010; 6(12):e1001013. https://doi.org/10.1371/journal.pcbi.1001013 PMID: 21152008

27. Butz M, Steenbuck ID, van Ooyen A. Homeostatic structural plasticity increases the efficiency of small-world networks. Frontiers in Synaptic Neuroscience. 2014; 6(April):7. https://doi.org/10.3389/fnsyn.2014.00007 PMID: 24744727

28. Josselyn SA, Köhler S, Frankland PW. Finding the engram. Nature Reviews Neuroscience. 2015; 16(9):521–534. https://doi.org/10.1038/nrn4000 PMID: 26289572

29. Josselyn SA, Frankland PW. Memory Allocation: Mechanisms and Function. Annual Review of Neuroscience. 2018; 41(1):389–413. https://doi.org/10.1146/annurev-neuro-080317-061956 PMID: 29709212

30. Kriener B, Tetzlaff T, Aertsen A, Diesmann M, Rotter S. Correlations and population dynamics in cortical networks. Neural computation. 2008; 20(9):2185–2226. https://doi.org/10.1162/neco.2008.02-07-474 PMID: 18439141

31. Turrigiano GG, Nelson SB. Homeostatic plasticity in the developing nervous system; 2004; 20(9):2185–2226. https://doi.org/10.1162/neco.2008.02-07-474

32. Buzsáki G, Mizuseki K. The log-dynamic brain: how skewed distributions affect network operations. Nature Reviews Neuroscience. 2014; 15(4):264–278. https://doi.org/10.1038/nrn3687 PMID: 24569488

33. Rubin DC, Wenzel AE. One hundred years of forgetting: A quantitative description of retention. Psychological Review. 1996; 103(4):734–760. https://doi.org/10.1037/0033-295X.103.4.734

34. Grewe BF, Helmchen F. High-speed two-photon calcium imaging of neuronal population activity using acousto-optic deflectors. Cold Spring Harbor Protocols. 2014; 2014(6):618–629. https://doi.org/10.1101/pdb.prot081778 PMID: 24890212

35. Grewe BF, Langer D, Kasper H, Kampa BM, Helmchen F. High-speed in vivo calcium imaging reveals neuronal network activity with near-millisecond precision. Nature Methods. 2010; 7(5):399–405. https://doi.org/10.1038/nmeth.1453 PMID: 20400966

36. Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, et al. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nature Neuroscience. 2011; 14(8):1045–1052. https://doi.org/10.1038/nn.2876 PMID: 21765421

37. Ziv Y, Burns LD, Cocker ED, Hamel EO, Ghosh KK, Kitch LJ, et al. Long-term dynamics of CA1 hippocampal place codes. Nature Neuroscience. 2013; 16(3):264–266. https://doi.org/10.1038/nn.3329 PMID: 23396101

38. Trachtenberg JT, Chen BE, Knott GW, Feng G, Sanes JR, Welker E, et al. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature. 2002; 420(6917):788–794. https://doi.org/10.1038/nature01273 PMID: 12490942

39. Nader K, Schafe GE, Le Doux JE. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature. 2000; 406(6797):722–726. https://doi.org/10.1038/35021052 PMID: 10963596

40. Yiu AP, Mercaldo V, Yan C, Richards B, Rashid AJ, Hsiang HLL, et al. Neurons Are Recruited to a Memory Trace Based on Relative Neuronal Excitability Immediately before Training. Neuron. 2014; 83(3):722–735. https://doi.org/10.1016/j.neuron.2014.07.017 PMID: 25102562

41. Zhou Y, Won J, Karlsson MG, Zhou M, Rogerson T, Balaji J, et al. CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nature Neuroscience. 2009; 12(11):1438–1443. https://doi.org/10.1038/nn.2405 PMID: 19783993

42. Holtmaat A, Wilbrecht L, Knott GW, Welker E, Svoboda K. Experience-dependent and cell-type-specific spine growth in the neocortex. Nature. 2006; 441(7096):979–983. https://doi.org/10.1038/nature04783 PMID: 16791195

43. Moser MB, Trommald M, Andersen P. An increase in dendritic spine density on hippocampal CA1 pyramidal cells following spatial learning in adult rats suggests the formation of new synapses. Proceedings of the National Academy of Sciences of the United States of America. 1994; 91(26):12673–5. https://doi.org/10.1073/pnas.91.26.12673 PMID: 7809099

44. Trommald M, Hulleberg G, Andersen P. Long-term potentiation is associated with new excitatory spine synapses on rat dentate granule cells. Learning & memory (Cold Spring Harbor, NY). 1996; 3(2-3):218–228. https://doi.org/10.1101/lm.3.2-3.218 PMID: 10456092

45. Bonhoeffer T, Engert F. Dendritic spine changes associated with hippocampal long-term synaptic plasticity. Nature. 1999; 399(6731):66–70. https://doi.org/10.1038/19978 PMID: 10331391

46. Müller M, Gähwiler B, Rietschin L, Thompson S. Reversible loss of dendritic spines and altered excitability after chronic epilepsy in hippocampal slice cultures. Proceedings of the National Academy of Sciences of the United States of America. 1993; 90(1):257–261. https://doi.org/10.1073/pnas.90.1.257 PMID: 8093558

47. Drakew A, Müller M, Gähwiler B, Thompson S, Frotscher M. Spine loss in experimental epilepsy: quantitative light and electron microscopic analysis of intracellularly stained CA3 pyramidal cells in hippocampal slice cultures. Neuroscience. 1996; 70(1):31–45. https://doi.org/10.1016/0306-4522(95)00379-W PMID: 8848134

48. Fauth M, Tetzlaff C. Opposing Effects of Neuronal Activity on Structural Plasticity. Frontiers in Neuroanatomy. 2016; 10:75. https://doi.org/10.3389/fnana.2016.00075 PMID: 27445713

49. Rao-Ruiz P, Visser E, Mitrić M, Smit AB, Oever MCvd. A Synaptic Framework for the Persistence of Memory Engrams. Frontiers in Synaptic Neuroscience. 2021; 13:661476. https://doi.org/10.3389/fnsyn.2021.661476 PMID: 33841124

50. Zuo Y, Lin A, Chang P, Gan WB. Development of long-term dendritic spine stability in diverse regions of cerebral cortex. Neuron. 2005; 46(2):181–189. https://doi.org/10.1016/j.neuron.2005.04.001 PMID: 15848798

51. Holtmaat AJGD, Trachtenberg JT, Wilbrecht L, Shepherd GM, Zhang X, Knott GW, et al. Transient and persistent dendritic spines in the neocortex in vivo. Neuron. 2005; 45(2):279–291. https://doi.org/10.1016/j.neuron.2005.01.003 PMID: 15664179

52. Stokes MG. 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework; 2015.

53. Helias M, Rotter S, Gewaltig MO, Diesmann M. Structural plasticity controlled by calcium based correlation detection. Frontiers in Computational Neuroscience. 2008; 2(DEC):7. https://doi.org/10.3389/neuro.10.007.2008 PMID: 19129936

54. Gershman SJ, Monfils MH, Norman KA, Niv Y. The computational nature of memory modification. eLife. 2017; 6. https://doi.org/10.7554/eLife.23763

55. Gerstein GL, Bedenbaugh P, Aertsen AMHJ. Neuronal assemblies. IEEE transactions on bio-medical engineering. 1989; 36(1):4–14. https://doi.org/10.1109/10.16444 PMID: 2646211

56. Friston KJ. Functional and effective connectivity in neuroimaging: A synthesis. Human Brain Mapping. 1994; 2(1-2):56–78. https://doi.org/10.1002/hbm.460020107

57. Knoblauch A, Sommer FT. Structural Plasticity, Effectual Connectivity, and Memory in Cortex. Frontiers in Neuroanatomy. 2016; 10(JUNE):63. https://doi.org/10.3389/fnana.2016.00063 PMID: 27378861

58. Scharfman HE. The neurobiology of epilepsy. Current Neurology and Neuroscience Reports. 2007; 7 (4):348–354. https://doi.org/10.1007/s11910-007-0053-z

59. Buckmaster PS, Zhang GF, Yamawaki R. Axon Sprouting in a Model of Temporal Lobe Epilepsy Creates a Predominantly Excitatory Feedback Circuit. The Journal of Neuroscience. 2002; 22(15):6650–6658. https://doi.org/10.1523/JNEUROSCI.22-15-06650.2002 PMID: 12151544

60. Semon RW. The mneme. London: Allen & Unwin; 1921. Available from: https://archive.org/stream/cu31924100387210.

61. Diaz-Pier S, Naveau M, Butz-Ostendorf M, Morrison A. Automatic Generation of Connectivity for Large-Scale Neuronal Network Models through Structural Plasticity. Frontiers in Neuroanatomy. 2016; 10:57. https://doi.org/10.3389/fnana.2016.00057 PMID: 27303272

62. Brunel N. Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. Journal of Computational Neuroscience. 2000; 8(3):183–208. https://doi.org/10.1023/A:1008925309027 PMID: 10809012

63. Tuckwell HC. Introduction to theoretical neurobiology. Cambridge University Press; 1988.

64. Nawrot MP, Boucsein C, Rodriguez Molina V, Riehle A, Aertsen A, Rotter S. Measurement of variability dynamics in cortical spike trains. Journal of Neuroscience Methods. 2008; 169(2):374–390. https://doi.org/10.1016/j.jneumeth.2007.10.013 PMID: 18155774

65. Ledoux E, Brunel N. Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. Frontiers in Computational Neuroscience. 2011; 5:25. https://doi.org/10.3389/fncom.2011.00025 PMID: 21647353

66. Gerstner W. Population dynamics of spiking neurons: Fast transients, asynchronous states, and locking. Neural Computation. 2000; 12(1):43–89. https://doi.org/10.1162/089976600300015899 PMID: 10636933

67. Linssen C, Lepperød ME, Mitchell J, Pronold J, Eppler JM, Keup C, et al. NEST 2.16.0. 2018. https://doi.org/10.5281/ZENODO.1400175

68. Gerstner W, Kistler WM, Naud R, Paninski L. Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press; 2014.