



# Emergence of Stable Synaptic Clusters on Dendrites Through Synaptic Rewiring

Thomas Limbacher and Robert Legenstein\*

*Institute of Theoretical Computer Science, Graz University of Technology, Graz, Austria*

The connectivity structure of neuronal networks in cortex is highly dynamic. This ongoing cortical rewiring is assumed to serve important functions for learning and memory. We analyze in this article a model for the self-organization of synaptic inputs onto dendritic branches of pyramidal cells. The model combines a generic stochastic rewiring principle with a simple synaptic plasticity rule that depends on local dendritic activity. In computer simulations, we find that this synaptic rewiring model leads to synaptic clustering, that is, temporally correlated inputs become locally clustered on dendritic branches. This empirical finding is backed up by a theoretical analysis which shows that rewiring in our model favors network configurations with synaptic clustering. We propose that synaptic clustering plays an important role in the organization of computation and memory in cortical circuits: we find that synaptic clustering through the proposed rewiring mechanism can serve as a mechanism to protect memories from subsequent modifications on a medium time scale. Rewiring of synaptic connections onto specific dendritic branches may thus counteract the general problem of catastrophic forgetting in neural networks.

## OPEN ACCESS

### Edited by:

Paul Miller,

Brandeis University, United States

### Reviewed by:

Yi Zhong,

Tsinghua University, China

Garibaldi Pineda Garcia,

University of Sussex, United Kingdom

### \*Correspondence:

Robert Legenstein

robert.legenstein@igi.tugraz.at

**Received:** 03 March 2020

**Accepted:** 22 May 2020

**Published:** 06 August 2020

### Citation:

Limbacher T and Legenstein R (2020)

Emergence of Stable Synaptic

Clusters on Dendrites Through

Synaptic Rewiring.

Front. Comput. Neurosci. 14:57.

doi: 10.3389/fncom.2020.00057

**Keywords:** dendrites, synaptic plasticity, structural plasticity, rewiring, synaptic clustering, catastrophic forgetting, neuroscience, spiking neural networks

## 1. INTRODUCTION

Long-term imaging studies of the living brain have revealed that the cortical connectivity structure is dynamic, with dendritic spines being added and deleted on the time scale of hours to days (Holtmaat et al., 2005; Stettler et al., 2006; Kasai et al., 2010; Loewenstein et al., 2011; Rumpel and Triesch, 2016). It has been proposed that this ongoing cortical rewiring serves important functions for learning and memory (Chklovskii et al., 2004; DeBello, 2008). According to this view, synaptic rewiring defines the connectivity structure of cortical circuits and interacts with synaptic plasticity of established synaptic connections.

Many theoretical studies of cortical rewiring have explored the implications of rewiring on the network level using point neuron models (see section 3). However, rewiring can also shape the connectivity structure on the sub-cellular level, defining the dendritic targets of synaptic connections onto pyramidal cells (PCs). As dendrites of cortical PCs exhibit various types of non-linear regenerative events, so-called dendritic spikes (Larkum et al., 2009), the specific placement of synaptic connections at the dendritic tree of a PC can strongly impact its computational function (Mel, 1994; Kastellakis et al., 2015; Bono et al., 2017). In particular, with regard to the organization of synapses on dendritic branches, a popular hypothesis with strong experimental support is the synaptic clustering hypothesis, which states that functionally related synapses tend to cluster on

dendritic branches (Govindarajan et al., 2006; Kastellakis et al., 2015). A theoretical justification for rewiring of synapses on non-linear dendrites was provided by Poirazi and Mel (2001). There, it was shown that a supervised structural plasticity rule can optimize memory performance of a simple non-spiking model of a pyramidal cell with non-linear dendrites. Later, Legenstein and Maass (2011) showed that branch-specific synaptic plasticity—without synaptic rewiring—can self-organize non-linear computations in neurons with non-linear branches.

In this article, we study synaptic rewiring based on simple synaptic plasticity rules in a spiking neuron model for PCs with dendritic non-linearities. In computer simulations of the model, we find that rewiring leads to a clustering of temporally correlated inputs onto dendritic branches, thus supporting the synaptic clustering hypothesis. Using this model, we argue that synaptic clustering through rewiring could play an important functional role in learning processes. It was shown by Cichon and Gan (2015) that different motor learning tasks induce dendritic spikes in different dendritic branches of PCs in mouse motor cortex. A possible explanation for this finding is that synaptic inputs are clustered onto dendritic branches in a task-specific manner such that a dendritic branch receives predominantly synaptic inputs that are activated at a specific task. This interpretation is consistent with the synaptic clustering hypothesis. When this segregation was pharmacologically disrupted, the animals could still learn the current task, but performance on an earlier learned task degraded, indicating a role of clustering for retaining older memories. One may thus hypothesize that synaptic clustering provides a mechanism to shelter previous memories from being overwritten by novel plasticity events. Testing this hypothesis, we found in another series of computer simulations that synaptic clustering through our simple rewiring dynamics is able to shelter previous memories from subsequent modifications.

To model synaptic rewiring, we take advantage of the synaptic sampling theory (Kappel et al., 2015), which provides a general framework for the interaction of synaptic plasticity and rewiring. Synaptic plasticity in this framework has a strong stochastic component (Yasumatsu et al., 2008; Dvorkin and Ziv, 2016). This stochasticity naturally gives rise to a stochastic addition and deletion of synaptic connections, which can be shown to implement a stochastic search over connectivity structures. In our application of this framework, we combine rewiring with simple plasticity rules where plasticity events are triggered by local dendritic spikes (Golding et al., 2002; Lisman and Spruston, 2005; Gordon et al., 2006). While the branch-specific plasticity mechanism inevitably leads to the independent self-organization of each dendritic branch, we show that the addition of a spike-timing-dependent plasticity mechanism, as proposed by Legenstein and Maass (2011), can orchestrate this process across branches, leading to an increased capacity to store input patterns.

To probe conditions under which clustering emerges in our model, we tested the model in various situations. While synaptic clustering can occur under various input statistics, we observed no clustering when individual synaptic inputs were activated in an unreliable stochastic manner. Since these statistics are presumably more characteristic for early sensory areas, this may explain why some studies reported a lack of

synaptic clustering (Jia et al., 2010). Finally, we theoretically analyzed the consequences of rewiring in our model. We found that the proposed dynamics approximates a stochastic search over connectivity structures that favors functionally clustered synaptic configurations.

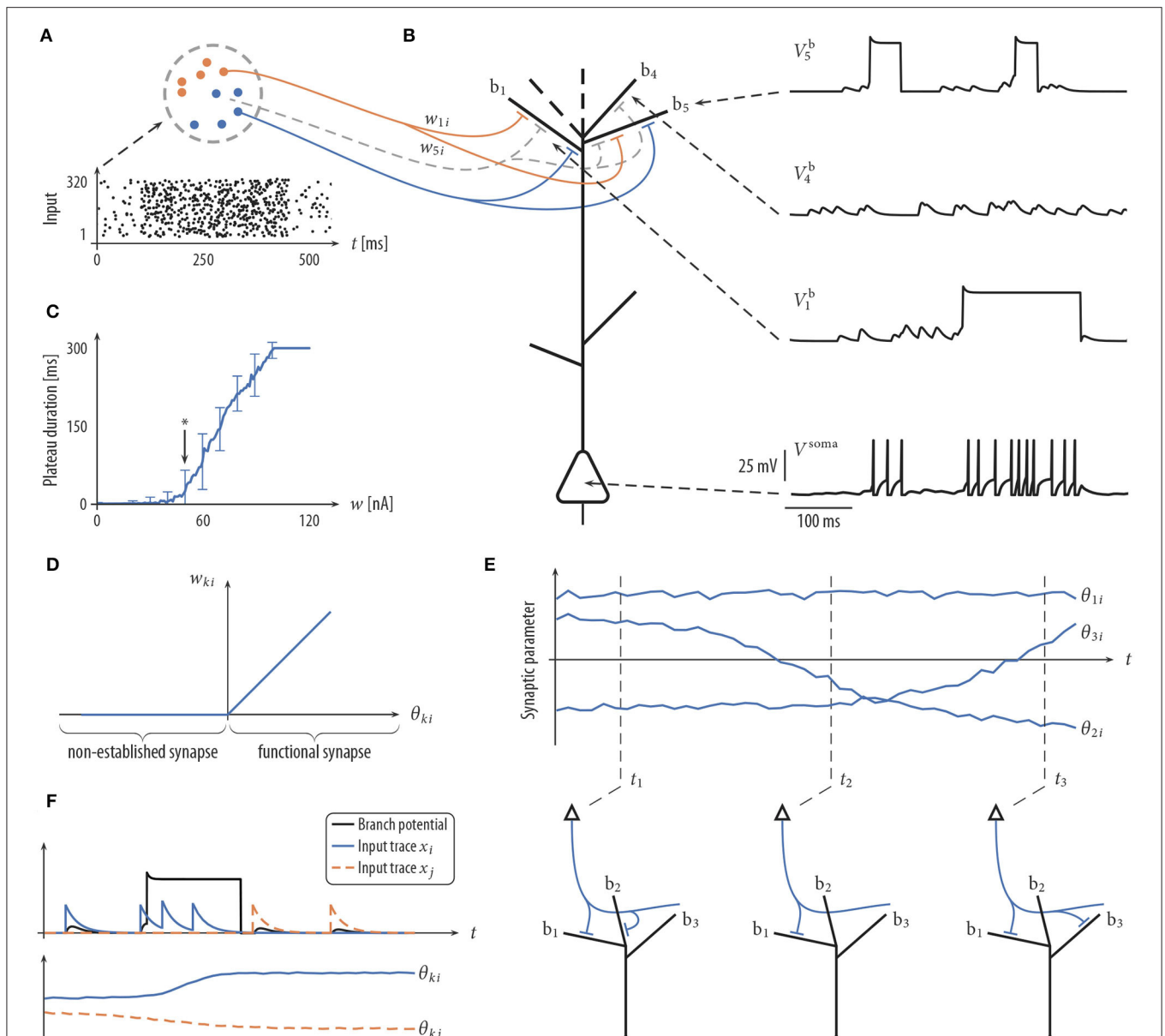
## 2. RESULTS

### 2.1. Synaptic Rewiring on Dendritic Branches

Dendritic structures of pyramidal cells (PCs) can be divided into integrative compartments (Losonczy and Magee, 2006; Major et al., 2008; Larkum et al., 2009; Branco et al., 2010). The compartmental model for PCs adopted in this article is illustrated in **Figures 1A,B**. The axon of a presynaptic neuron  $i$  can contact various dendritic branches  $k$  of the postsynaptic PC, establishing a synaptic connection  $ki$  with synaptic efficacy  $w_{ki}$ . Presynaptic spikes at such a synapse  $ki$  give rise to alpha-shaped currents of amplitude  $w_{ki}$  at the branch. Each dendritic branch acts as a leaky integrator that temporally and spatially sums synaptic input currents from its incoming synapses. Note that we do not consider any specific spatial structure of the branches. The structures shown in the figures are provided for illustration purposes only.

Besides this passive response to synaptic inputs, dendrites of PCs show a variety of non-linear responses. In particular, we model local N-methyl-D-aspartate (NMDA) spikes (Antic et al., 2010) that include a brief sodium spikelet followed by a plateau potential that lasts up to a few hundreds of milliseconds, see **Figure 1B** [amplitude values of the plateau potentials are taken from Antic et al. (2010), see Figure 1 there]. NMDA spikes are elicited stochastically at dendritic branches based on the local membrane potential (branch potential). The duration of the spike grows linearly with stimulus intensity (**Figure 1C**), whereas the amplitude exhibits an all-or-none highly non-linear behavior (Antic et al., 2010). Currents from dendritic compartments to the soma are driven by the difference between the branch potential and the membrane potential of the soma. The somatic compartment is again a leaky integrator that sums currents from all dendrites (with a certain degree of amplitude attenuation, see Equation 17). Somatic action potentials are elicited stochastically with an instantaneous firing rate that depends exponentially on the somatic membrane potential. See Jolivet et al. (2006) for a fit of this model of action potential generation to layer V pyramidal neurons in rat somatosensory cortex. A more detailed description of the neuron model is given in section 4.

The connectivity structure in the cortex is not static, rather synaptic connections are rewired on a time scale of hours to days (Holtmaat et al., 2005; Stettler et al., 2006; Minerbi et al., 2009; Yang et al., 2009; Ziv and Ahissar, 2009; Kasai et al., 2010; Loewenstein et al., 2011; Rumpel and Triesch, 2016; Chambers and Rumpel, 2017). To investigate the consequences of rewiring processes for information processing in PCs, we adopted the synaptic sampling framework introduced by Kappel et al. (2015, 2018). Each compartment has a set of potential synaptic connections that could be realized by some parameter setting.



**FIGURE 1 |** Schema of neuron model and plasticity/rewiring. **(A)** Input neurons (colored dots) are divided into different input assemblies (2 shown; color indicates assembly assignment). Bottom: spike raster of input neurons. Dots represent spike times. **(B)** Schematic drawing of neuron model (spatial structure of branches for illustrative purposes only) with dendritic membrane potentials  $V_i^b$  (top right) and somatic membrane potential  $V^{\text{soma}}$  (bottom right). Branch  $b_5$  emits two dendritic spikes with a duration of about 50 ms. Branch  $b_1$  spikes once with a longer plateau phase. No synaptic cluster was established on branch  $b_4$ . Therefore, this branch did not elicit dendritic spikes. Somatic spikes are indicated for illustrative purpose by vertical lines. **(C)** The duration of a plateau potential grows linearly as a function of the input intensity. The synaptic efficacy  $w$  of a single synaptic input was varied between 0 and 120 nA evenly spaced; a single strong synapse was used to mimic strong synchronous input. Shown is the mean and standard deviation of the resulting plateau durations over 100 independent trials. Arrow (at 50 nA) roughly marks the onset of this linear growth. Synaptic input exceeding this threshold has a fair chance of triggering a dendritic spike. **(D)** Mapping between the synaptic parameter  $\theta_{ki}$  and the synaptic efficacy  $w_{ki}$  of synapse  $i$  onto branch  $k$ . Negative values of  $\theta_{ki}$ , corresponding to non-established synapses, are mapped to zero in the  $w_{ki}$ -space. **(E)** Evolution of three synaptic parameters as a function of time  $t$  (top) and the corresponding wiring diagram at three points in time (bottom). At time  $t_1$  the values of parameters  $\theta_{1i}$  and  $\theta_{2i}$  are positive (indicating a functional synapse) while the value of parameter  $\theta_{3i}$  is negative (indicating a non-established synapse). Parameter  $\theta_{2i}$  crosses zero shortly before  $t_2$  (becoming non-established). **(F)** The plasticity process. Branch potential (black; sub-threshold potentials and a dendritic spike), the somatic spike traces of input neurons  $i$  and  $j$  (top), and the evolution of synaptic parameters  $\theta_{ki}$  and  $\theta_{kj}$  (bottom). Input neuron  $i$  is active shortly before and during the dendritic spike. The synaptic parameter  $\theta_{ki}$  of this connection is therefore increasing.  $\theta_{kj}$  decreases, since neuron  $j$  is not active during the plateau potential.

However, at each point in time, only a subset of these connections is realized by functional synapses. More precisely, we maintain one parameter  $\theta_{ki}$  for each potential synapse  $ki$  from input neuron  $i$  to branch  $k$  of the neuron. This parameter encodes (a) through its sign whether the synapse is functional and (b) the synaptic efficacy  $w_{ki}$  of the synapse if it is functional. More precisely, the synaptic weight  $w_{ki}$  is  $c_\theta \theta_{ki}$  for a functional synapse ( $\theta_{ki} > 0$ ) and 0 otherwise ( $\theta_{ki} \leq 0$ ). Hence, the mapping from the parameter  $\theta_{ki}$  to the synaptic weight is given by  $w_{ki} = c_\theta \max\{0, \theta_{ki}\}$  where  $c_\theta = 1$  nA is the slope of this mapping for  $\theta_{ki} > 0$  (**Figure 1D**).

Recent experimental results show that synaptic modifications contain a strong autonomous component (Dvorkin and Ziv, 2016). In accordance with previous modeling approaches (Loewenstein et al., 2011), this component is modeled in the synaptic sampling framework as a stochastic process. In general, the stochastic synaptic dynamics is given in this model as a stochastic differential equation (SDE) of the form

$$d\theta_{ki}(t) = \eta \left( f_{ki}^S(t) + f_{ki}^L(t) \right) dt + \sqrt{2\eta T} d\mathcal{W}_{ki}, \quad (1)$$

where  $\eta > 0$  is a small learning rate (for parameter values, see **Table 2**). This SDE can be read as follows: the change of the synaptic parameter  $\theta_{ki}$  is the sum of three terms. The last term models the synapse-autonomous component that does not depend on any synaptic, presynaptic, or postsynaptic variables. Instead, it is given by the increments of a standard Wiener process  $\mathcal{W}_{ki}$ , thus implementing random walk behavior. The temperature  $T$  is a constant that scales the strength of this stochastic component. The term in the brackets defines the deterministic part of the dynamics. For conceptual clarity, it is divided here into two components. The first term  $f_{ki}^S$  describes changes that are not directly relevant for the functional goal of the plasticity process but rather used to enforce structural constraints on the connectivity (for example, bound the number of connections that can be established onto one branch). The functional goal of the plasticity process is captured by the second term  $f_{ki}^L$ , see below for an example. The deterministic terms will be chosen such that they vanish for  $\theta_{ki}(t) \leq 0$ . Hence, for non-established synapses, only the stochastic term remains. The random walk may at some point cross zero again, leading to the creation of a synaptic connection which may be strengthened or weakened depending on its functional relevance (**Figure 1E**). The model thus integrates synaptic dynamics with rewiring in one unifying framework. Note that the creation of new synapses solely depends on the Wiener process. New synapses can also directly be sampled from the set of non-established potential synapses, which however loses some of the nice mathematical properties of the model, see Bellec et al. (2018) for details.

Since in cortical PCs, the number of connections to a branch is bounded, we use for the structural component  $f_{ki}^S$  of the plasticity dynamics (1) a term that enforces a soft upper bound on the number of synaptic connections onto each branch. This term depresses weak synapses of a branch if the number of connections to it is close to or exceeds a predefined constant  $N_{\text{syn}}$  (set to 20 in our simulations), and it vanishes if the number of connections is well below  $N_{\text{syn}}$ .

According to the Hebbian theory of synaptic plasticity, synaptic connections are created and/or strengthened when the presynaptic neuron contributes to activity of the postsynaptic neuron (Hebb, 1949). More recent experiments, however, showed that postsynaptic somatic spiking is neither necessary nor sufficient to induce long-term potentiation (LTP) in pyramidal cells in various brain structures, both, at distal basal dendrites and at the apical tuft (Golding et al., 2002; Lisman and Spruston, 2005; Gordon et al., 2006; Gambino et al., 2014; Cichon and Gan, 2015; Brandalise et al., 2016). These studies indicate that local dendritic NMDA-dependent regenerative processes are necessary as the postsynaptic signal to trigger long-term potentiation (LTP). An important implication of this finding is that the learning signal is at least partially local to dendritic branches, leading to a synaptic organization that depends on local dendritic integration. To model such LTP within the stochastic rewiring framework, we include a term  $x_i(t)\Gamma_k(t)$  in the functional term  $f_{ki}^L$  of the plasticity dynamics (1). Here,  $x_i(t)$  is a trace of the spike train from presynaptic neuron  $i$ , and  $\Gamma_k(t) = 1$  indicates the presence of an NMDA plateau potential at time  $t$  ( $\Gamma_k(t) = 0$  otherwise). Hence, coincidence of presynaptic activity with a dendritic plateau potential will induce LTP (see **Figure 1F**, blue trace). While the dependence for LTP on dendritic spikes is well-established in the literature, less is known about long-term depression (LTD) in this regard. Interestingly, while Golding et al. (2002) observed LTP when pairing high-frequency stimulation of synaptic inputs with dendritic spikes in a theta-burst pairing protocol, Holthoff et al. (2004) reported long-term depression (LTD) with a single presynaptic stimulation leading to a dendritic spike, indicating that low-frequency pairing may lead to LTD, consistent with the classical findings that weak synaptic stimulation leads to LTD (Dudek and Bear, 1995) under the assumption that dendritic spikes are elicited in the postsynaptic neuron occasionally. We therefore included a term in our plasticity model that weakens a synapse if its activity is below a certain level during a plateau potential. Together, the functional component  $f_{ki}^L$  of the plasticity dynamics (1) was given by

$$f_{ki}^L(t) = \begin{cases} c_L \Gamma_k(t) x_i(t) & \text{if } \theta_{ki}(t) > 0 \text{ (functional connection)} \\ -\gamma(1 - x_i(t)), & \\ 0, & \text{if } \theta_{ki}(t) \leq 0 \text{ (non-established connection),} \end{cases} \quad (2)$$

with  $\gamma > 0$  being a constant that determines the threshold activity that switches from LTD to LTP. Note that changes occur only during the presence of a plateau potential [ $\Gamma_k(t) = 1$ ]. The rule can be described qualitatively as follows: if there is a dendritic spike at branch  $k$ , potentiate active synapses and depress inactive synapses to this branch  $k$ . Weak synapses will ultimately retract, and non-established potential synapses will be established at random times according to the stochastic process. We show in the section “Analysis of stochastic rewiring dynamics” that this dynamic corresponds to a distribution  $p_{\mathcal{L}}$  that defines functionally preferred parameter settings that lead to synaptic clustering.

## 2.2. Synaptic Clustering Through Rewiring

We first investigated the patterns of synaptic connections that emerge when synaptic rewiring is driven by the simple dendritic



spike-dependent plasticity mechanism (2). We simulated a neuron with 12 independent dendritic compartments that received input from 320 input neurons. Input neurons were divided into 8 disjoint assemblies of 40 neurons each, producing Poisson spike trains at a background rate of 1 Hz. Every 500 ms, one assembly was chosen randomly and activated for 300 ms. These neurons elevated their firing rate to 35 Hz, leading to a specific input firing pattern (Poisson spike trains were randomly generated at every pattern presentation), see **Figure 2A** (top).

Input neurons were initially connected to dendritic branches randomly such that 20 synapses were established on each branch (**Figure 2A** middle). Synaptic parameters of these connections were independently drawn from a uniform distribution over the interval [4, 8]. A soft upper bound of 20 synapses per branch was enforced through the structural component  $f_{ki}^S$  of the synaptic sampling update (1). At this initial configuration, brief dendritic plateau potentials appeared occasionally and the neuronal firing rate was low ( $\approx 6.19$  Hz). Dendritic plateau potentials were elicited preferentially during presynaptic assembly activations, but they were not specific to any of the input assemblies. We observed a gradual assembly-specific clustering of synapses onto individual branches during the simulation. After 17 min of assembly activations, each branch usually hosted a synaptic cluster from one of the input assemblies, receiving approximately 20 synapses from that assembly with the weight close to the maximum, while other assemblies were either not connected or had weights close to zero, see **Figure 2B**. Many synapses that were functional after rewiring were not connected to the branches initially but were established during the rewiring process (**Figure 2C**). In order to quantify the behavior of the model, we performed 25 simulations with different randomly chosen initial conditions and different random assembly activations. We say that assembly synapses cluster on a branch if the branch receives at least 10 synapses from one assembly with a total weight of at least 50 nA. The value of 50 nA was chosen since synaptic input exceeding this threshold has a fair chance of triggering a dendritic spike (see arrow in **Figure 1C**). In our 25 independent simulations, from the 8 input assemblies, between 5 and 8 were represented on dendritic branches with a mean of 6.36 and a standard deviation (SD) of 0.84. In all simulations, most of the dendritic branches had a synaptic cluster from one assembly, and some assemblies were clustered on several branches.

Since inputs to each branch were preferentially originating from a specific input assembly after rewiring, branch inputs were temporally correlated. This collective activity was sufficient to trigger long-lasting dendritic plateau potentials and somatic spikes (**Figure 2B**; voltage traces at active assemblies  $A_1$  and  $A_2$ ). In total, three branches received synaptic clusters from assembly  $A_2$  and two branches from assembly  $A_1$ . This was reflected in the firing rate of the neuron during the activation of these assemblies. The firing rate of the neuron was highest ( $\approx 63.5$  Hz) during the pattern presentation of assembly  $A_2$  and lowest ( $\approx 54.2$  Hz) during the activation of assembly  $A_1$ . To test the response of the neuron to assemblies that were not clustered, we activated assembly  $A_5$  that was active during the rewiring process but did not evolve a synaptic cluster on any branch. Branches did not

elicit spikes during the pattern presentation of this assembly and the neuronal firing rate was low ( $\approx 0.97$  Hz; **Figure 2B**; voltage traces at active assembly  $A_5$ ).

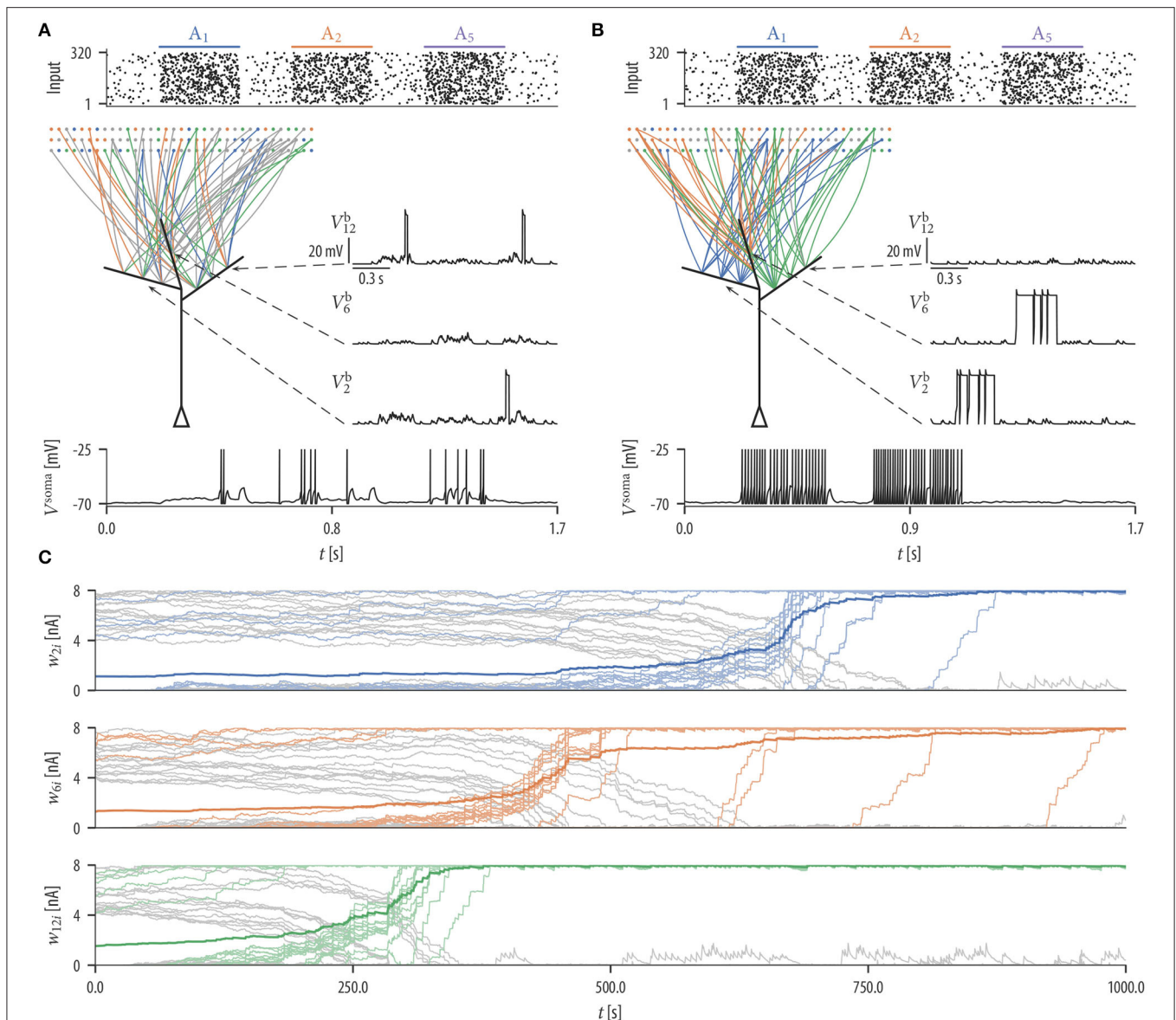
We hypothesized that dendritic plateau potentials are crucial for assembly-specific clustering of synapses in our model. To verify this we conducted simulations of an altered model with linear dendritic integration (i.e., the same model but without dendritic spikes). We did not observe clustering of functionally related inputs onto dendrites in this altered model on a large range of relevant parameters (see section 1.1 in the **Supplementary Material** for details), indicating that dendritic plateau potentials are indeed necessary for assembly-specific clustering of synapses onto individual dendrites.

In summary, these results show that synaptic rewiring dynamics can give rise to (a) clustering of functionally related inputs onto dendritic compartments and (b) segregation of different assemblies onto different branches. This leads to segregated assembly-specific activation of dendritic branches. Note that this plasticity-driven self-organization of network connectivity is independent of postsynaptic somatic spikes.

### 2.3. STDP Increases the Capacity of Neurons to Store Assembly Patterns

With the plasticity dynamics considered above, since plasticity depends solely on synaptic inputs and local dendritic potentials, all branches are adapted independently without taking the activity of other branches into account. This implies that synaptic patterns at different branches can become correlated in the sense that projections from one assembly cluster on two or more branches of the neuron. Such adaptation may in general waste dendritic resources such that all branches become occupied by a subset of input assemblies while other assemblies are neglected. This effect would become even more pronounced if assemblies were not activated in an intermixed manner but sequentially. In this case, early assemblies would occupy all branches, leaving no representational space for later ones.

Previous work has shown that dendritic synaptic patterns can be decorrelated by a simple additional spike-timing-dependent plasticity (STDP) mechanism where each somatic spike induces a small amount of depression in all synapses with a recent presynaptic spike (Legenstein and Maass, 2011). This rule is consistent with the experimental finding that in distal dendrites, a pre-before-post spike pair leads to depression (Kampa et al., 2007). The STDP update indirectly introduces competition between dendritic branches such that branches compete for becoming activated for each input pattern. Briefly, consider an input pattern for which one branch (or a few branches) has evolved a strong synaptic cluster. Whenever this input is presented, the branch becomes active, it produces dendritic plateau potentials which in turn will lead to somatic spikes. For branches that did not yet evolve a strong cluster for that input pattern, these somatic spikes then depress via STDP active synapses and thus prevent the emergence of a cluster for this pattern. On the other hand, the branch with the strong cluster and plateau potentials is



**FIGURE 2 |** Synaptic rewiring leads to synaptic clustering at dendritic branches. **(A)** Spike rasters of input neurons (top), initial wiring diagram of three selected branches (middle left), dendritic membrane potentials  $V_i^b$  of these branches (middle right), and the somatic membrane potential  $V^{\text{soma}}$  (bottom) for the successive activation of assemblies  $A_1$ ,  $A_2$ , and  $A_5$ . Color of graph edges indicate to which of the assemblies ( $A_1$ ,  $A_2$ , and  $A_3$ ; pattern presentation of assembly  $A_3$  is not shown) a connection was established. Connections in gray are connections to one of the other assemblies. At this initial configuration, the input neurons were connected to branches randomly. Brief dendritic spikes appeared occasionally (middle right) and the neuronal firing rate was low (bottom). **(B)** After 17 min of rewiring dynamics, inputs to each branch were originating predominantly from a specific input assembly  $A_i$  (middle left). Hence, temporally correlated branch input triggered long lasting dendritic plateau potentials (middle right). In this simulation, none of the branches received synaptic clusters from assembly  $A_5$ . As a consequence, the neuron elevated its firing rate during pattern presentations of input assemblies  $A_1$  and  $A_2$ , but rarely spiked during the pattern presentation of assembly  $A_5$  (bottom). **(C)** Evolution of the synaptic weights  $w_{ki}$  of branch  $b_2$  (top), branch  $b_6$  (middle), and branch  $b_{12}$  (bottom). Shown is the mean synaptic weight from three selected assembly ( $A_1$ ,  $A_2$ , and  $A_3$ ) which clustered on the branch (saturated color), weights of individual synapses from that assembly (desaturated color), and weights of synapses from other assemblies (gray). Colors correspond to colors in **(A,B)**. Synaptic connections are reorganized during the rewiring process, i.e., synapses with correlated activity are established on the branch while synapses of other assemblies retract over time. Occasional small gray bumps close to 0 in the second half of the simulation originate from synapses from other assemblies that are reconnected, but they quickly retract due to their irrelevance for activating the branch.

protected in the sense that potentiation from the dendritic-spike-dependent plasticity rule overcompensates the depression. This branch (or a few branches with a strong cluster) has thus won the competition for becoming activated by this input pattern.

To test the effect of this mechanism on synaptic rewiring, we performed 25 simulations as described above but with STDP added to the parameter dynamics (see section 4 for a definition of the STDP update). We found that the average number of represented assemblies increased from  $6.36 \pm 0.84$  to  $7.40 \pm 0.57$

(mean  $\pm$  SD). This is a significant increase of the number of assemblies that are stored on the dendrites [ $t_{(24)} = 5.01$ ,  $p < 0.0001$ , unpaired  $t$ -test]. A schematic drawing of the simulated neuron and the wiring diagram after 17 min of rewiring dynamics of four selected branches of one of the conducted trials is shown in **Figure 3A**. Each assembly in this simulation evolved a synaptic cluster onto exactly one dendritic branch. Four branches did not receive a synaptic cluster from any of the assemblies (inset **Figure 3A**). Since these branches did not specialize to a specific input assembly, they could, in principle, adapt later to respond to novel input assemblies. To test whether these results are sensitive to the specific values of parameters in our rewiring model, we performed a sensitivity analysis. This analysis showed that the number of synaptic clusters is quite robust to parameter variations (see section 1.2 in the **Supplementary Material**). We also tested effects of the pattern duration and the delay between patterns. We found that our rewiring mechanism is, up to some point, quite robust to variations in the pattern duration and the delay between patterns (see section 1.3 in the **Supplementary Material**).

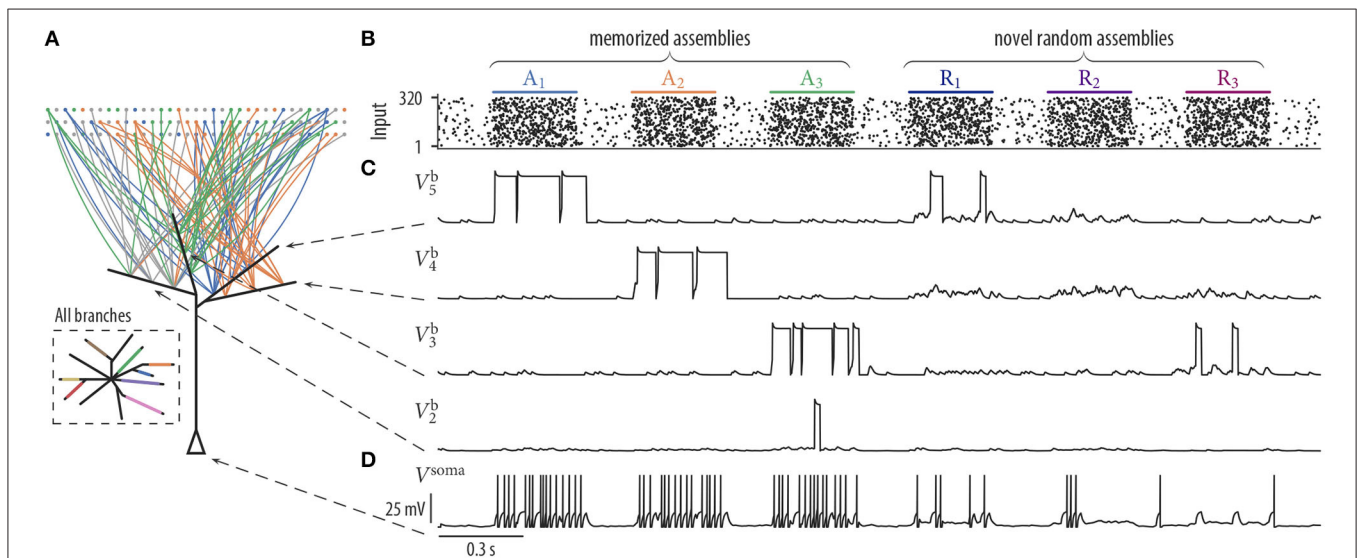
We next asked whether the neuron was able to differentiate after the rewiring process between stored and novel input patterns. We presented three firing patterns of memorized assemblies  $A_1$ ,  $A_2$ ,  $A_3$  and three firing patterns of novel random assemblies  $R_1$ ,  $R_2$ ,  $R_3$  to the neuron (**Figure 3B**). Each novel assembly was created by randomly choosing 40 distinct neurons

from all 320 input neurons. Hence, assemblies  $R_1$  to  $R_3$  were not necessarily disjoint and they shared a significant amount of neurons with assemblies  $A_1$  to  $A_8$ . The neuron responded with a high firing rate during the activation of memorized assemblies but spiked only occasionally for patterns of novel random assemblies (**Figure 3D**). Branches that did not receive a synaptic cluster from any of the assemblies still maintained rather strong individual synapses originating from multiple assemblies, but since their collective activity was not correlated; it was not sufficient to trigger long lasting dendritic plateau potentials in these branches (**Figure 3C**, bottom trace  $V_2^b$ ).

In summary, STDP increases the capacity of the neuron in terms of the number of assemblies that are stored on dendrites. If there are more branches than assemblies in the input, some branches remain assembly-unspecific and are therefore available to store novel assemblies. Unless otherwise stated, all of the following simulations presented below are conducted with the STDP update.

## 2.4. Rewiring Protects Stored Information

One effect of the synaptic clustering observed in our simulations was that activation of different input assemblies led to dendritic spikes in different dendritic branches, see **Figures 2A,B, 3B**. Similarly, Cichon and Gan (2015) observed task-specific segregation of dendritic activity in tuft branches of layer V pyramidal neurons in mouse motor cortex. More specifically,



**FIGURE 3** | Synaptic rewiring with spike-timing-dependent depression and neuron response to stored and novel input patterns. **(A)** Schematic neuron drawing wiring diagram of four selected branches after 17 min of rewiring dynamics (as in **Figure 2B**). Branches  $b_3$ ,  $b_4$ , and  $b_5$  received synaptic inputs predominantly from assemblies  $A_3$ ,  $A_2$ , and  $A_1$ , respectively. Branch  $b_2$  remained without a synaptic cluster from any assembly. Inset: Schematic drawing of all branches of the neuron. Branch color indicates assembly origin of a synaptic cluster. All assemblies were represented on the branches and each assembly established only a single cluster. Four branches remained without a synaptic cluster from any assembly. **(B–D)** Pattern-specific neuron responses after rewiring. **(B)** Input neurons spike raster. Assemblies  $A_1, \dots, A_3$  were also activated during the previous rewiring period. The last three firing patterns were generated by novel random assemblies  $R_i$ . **(C)** Dendritic membrane potentials  $V_i^b$  of four branches during the presentation of three test patterns of assemblies  $A_i$  and  $R_i$ . Dendritic plateau potentials were elicited preferentially during the activation of memorized assemblies  $A_i$  and appeared only occasionally at novel random assemblies  $R_i$ . **(D)** Somatic membrane potential  $V^{soma}$  during the presentation of input patterns of memorized assemblies ( $A_1$ ,  $A_2$ , and  $A_3$ ) and for patterns of novel random assemblies ( $R_1$ ,  $R_2$ , and  $R_3$ ). The neuron responded with a high firing rate ( $\approx 27.6$  Hz) to patterns of memorized assemblies and with a low firing rate ( $\approx 11.3$  Hz) during the activation of novel random assemblies.

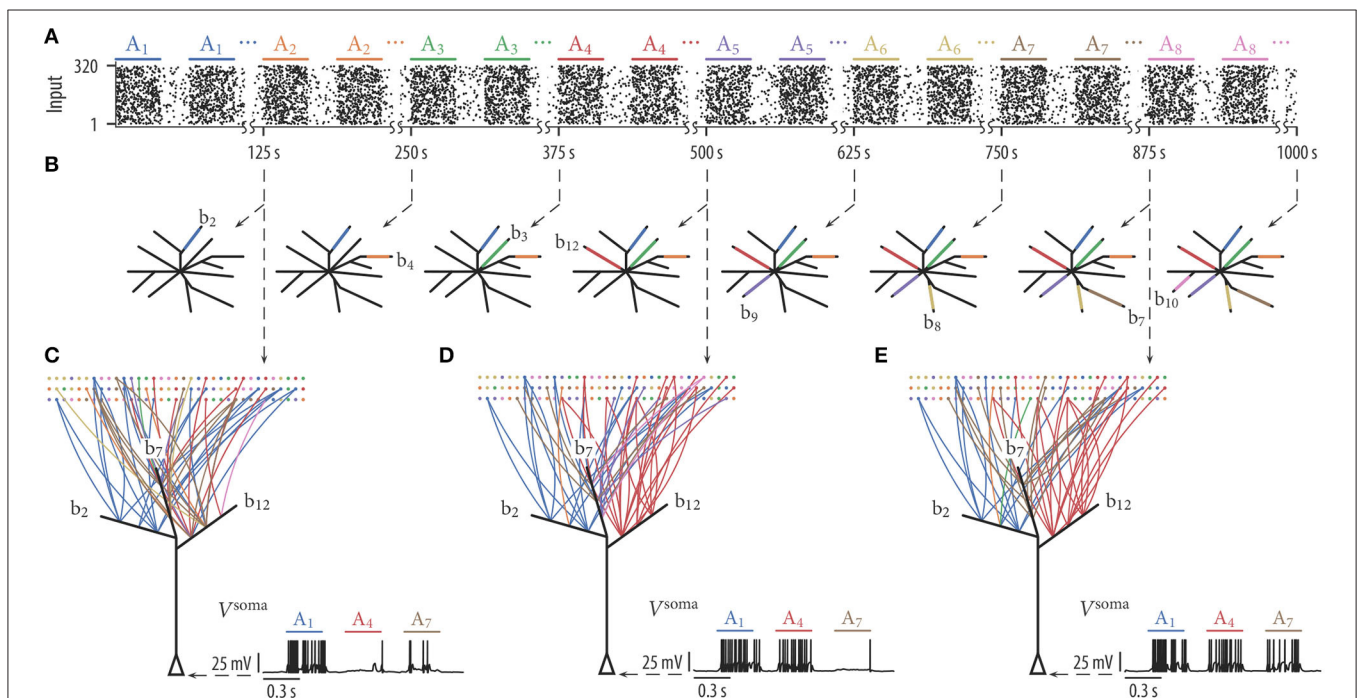


they observed that different motor learning tasks induce dendritic spikes on different branches. The processing of different behaviors is therefore separated onto different dendritic branches. Importantly, when this segregation was pharmacologically disrupted, the animals could still learn the current task, but performance on an earlier learned task degraded, indicating a role of this clustering for retaining older memories. The authors thus argued that this segregation may provide a mechanism to shelter previous memories from being overwritten by novel plasticity events.

To test whether simple rewiring rules can indeed support memory protection, we performed simulations as described above. Input assembly activations and background noise were generated as before, however, assemblies were activated in a sequential manner. In other words, we first activated exclusively assembly  $A_1$  250 times, then assembly  $A_2$ , then assembly  $A_3$ , etc. (with a delay between each pattern presentation as in the previous simulations; **Figure 4A**). **Figure 4B** shows schematic drawings of the 12 simulated branches at various times of the rewiring process. The color of a branch in **Figure 4B** indicates which of the 8 sequentially presented assemblies had evolved a synaptic cluster at the branch at the specified time. Clearly, branches were recruited sequentially to store input patterns. Branch  $b_2$  received a synaptic cluster from the first assembly, whereas none of the other branches was evolving large weights for this assembly. Beginning at 125 s, assembly  $A_2$  was activated while assembly  $A_1$

was silent. Shortly afterwards, branch  $b_4$  developed large synaptic weights for this input pattern while the synapse cluster at branch  $b_2$  remained stable. This stabilization is due to the fact that (a) synapses from assembly  $A_2$  to branch  $b_2$  were weak, and dendritic plateau potentials were consequently induced rarely, and (b), presynaptic activity from assembly  $A_1$  was weak, and depression due to inverse STDP was thus very small. As further assemblies became activated every 125 s, new branches were recruited to store these patterns while synaptic weights on old branches kept stable. In this simulation, all assemblies were represented on the branches. Four branches remained without a synapse cluster from any of the assemblies. We have repeated the experiment for 25 independent trials. The mean number of represented assemblies over these trials was  $6.92 \pm 0.89$  (mean  $\pm$  SD).

We hypothesized that competition between branches through STDP is essential for this capability. Due to the competition, only one or a few branches evolve a synaptic cluster from a specific assembly while all other branches remain neutral to assemblies. This is especially important when assemblies are presented in a sequential manner. Without competition, most of the branches would adapt to respond to the first active assembly. Most branches would thus be occupied and not available for novel assemblies. To quantify the effect of dendritic competition, we repeated the experiment without STDP. In this case, the mean number of represented assemblies reduced to  $4.04 \pm 0.72$  (mean  $\pm$  SD over 25 independent trials). As expected, most of the



**FIGURE 4** | Dendritic-spike-dependent rewiring protects stored information by segregating functionally unrelated inputs onto different dendritic branches. When input assemblies were activated sequentially, novel assemblies clustered on free branches while previously established clusters were retained. **(A)** Input assemblies  $A_i$  activated in a sequential manner. We first activated exclusively assembly  $A_1$ , then assembly  $A_2$ , then  $A_3$ , etc. **(B)** Schematic drawing of the 12 simulated branches at 8 different time points (branch color indicates assembly origin of synaptic cluster at the branch). Branches were recruited sequentially. **(C–E)** Network connectivity graph and somatic membrane potential  $V_{\text{soma}}$  of three selected branches ( $b_2$ ,  $b_7$ , and  $b_{12}$ ) at three different time points. Graph edge color indicates to which of the input assemblies a connection is established. The somatic potential shows the response of the neuron to one test pattern of assembly  $A_1$ , assembly  $A_4$ , and assembly  $A_7$ . The neuron successively reorganized its synaptic connections such that it responds to any of the patterns that have been presented previously.



branches evolved synaptic clusters from the first shown assembly  $A_1$  early on.

Due to the random fluctuations of weights, there is always a chance that synapses are overwritten. This behavior can be interpreted as gradual forgetting. If a synaptic cluster is inactive for an extended time, then the synaptic connections gradually degrade. If, however, that assembly is again activated, the synapses to that assembly will again stabilize provided that the weights are large enough to trigger a dendritic spike. To investigate the impact of the random fluctuations of weights in the weakening of older memories over time, we have performed simulations with various noise strengths and analyzed the weight decrease (see section 1.4 in the **Supplementary Material** for more detail). In section 1.5 of the **Supplementary Material**, we show a simple way to reduce gradual forgetting in our model and in section 1.6 of the **Supplementary Material** we propose a simple consolidation mechanism. Both of these methods integrate nicely with our rewiring framework. Synaptic connections of old memories are well-protected against changes induced by the noise term in this extended model (see sections 1.5, 1.6 in the **Supplementary Material**).

## 2.5. Synaptic Clustering Depends on Input Statistics

In the above simulations, input assemblies were activated in temporal isolation and each assembly activation was clean in the sense that all assembly neurons had an elevated firing rate and assemblies were disjoint. Under these conditions, our rewiring model led to robust synaptic clustering. We next determined the sensitivity of this result on these input statistics.

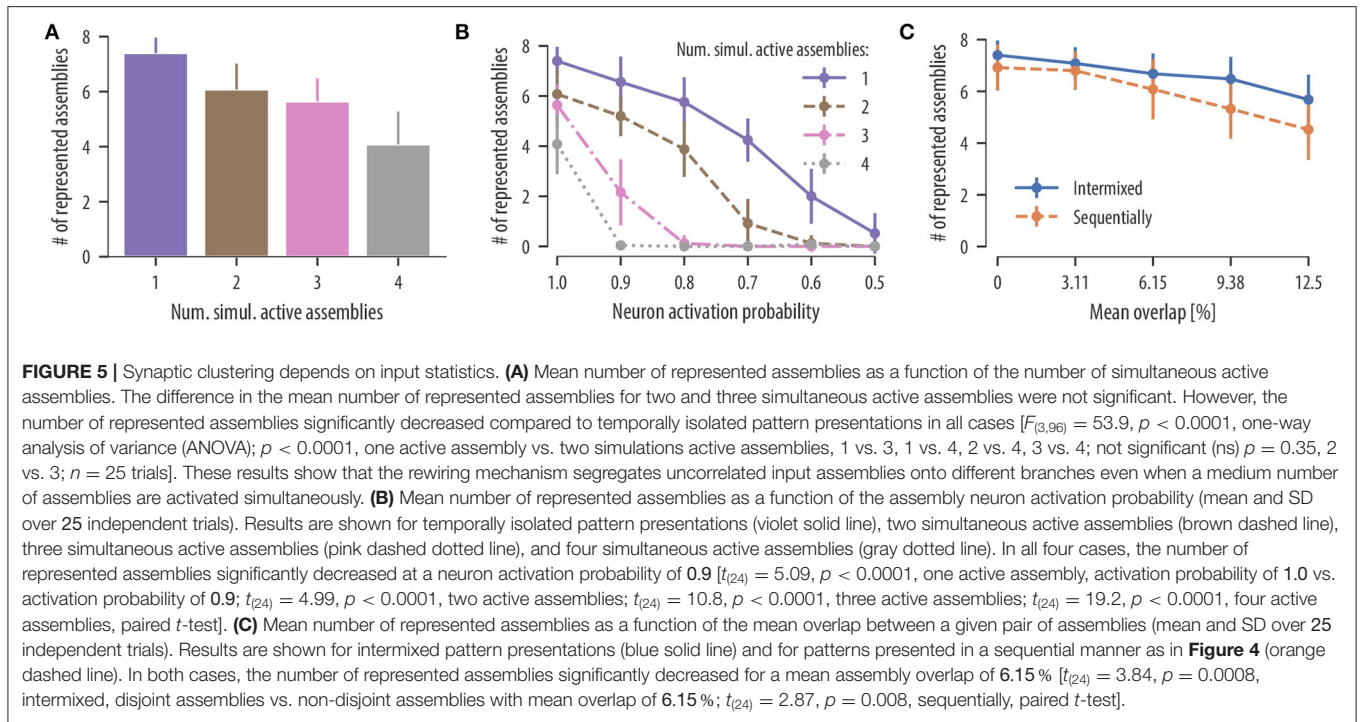
*Influence of co-active assemblies:* We first asked whether input assemblies can be clustered onto different dendritic compartments even when single assemblies are never activated in temporal isolation. As above, we simulated a neuron with 12 independent branches and a population of 320 input neurons divided into 8 disjoint neuronal assemblies of 40 neurons each. Input patterns and background noise were generated as before. At times of assembly activation, a medium number of assemblies was chosen uniformly at random from the set of all input assemblies and were simultaneously activated for 300 ms. For up to four simultaneous active assemblies, the rewiring dynamics separated connections from different assemblies onto different branches with some decrease of clustering. In 25 independent simulation runs, the mean number of represented assemblies on the neuron was  $6.08 \pm 0.94$  for two simultaneously activated assemblies,  $5.64 \pm 0.84$  for three simultaneously activated assemblies, and  $4.08 \pm 1.20$  for four simultaneously activated assemblies (**Figure 5A**). Our rewiring mechanisms consequently segregates uncorrelated inputs onto different branches even when assembly patterns are not temporally isolated.

*Influence of stochastic assembly activation:* While evidence for synaptic clustering has been reported in various cortical areas (Chen et al., 2011; Kleindienst et al., 2011; Fu et al., 2012), some studies found no evidence for it in early sensory areas (Jia et al., 2010). We hypothesized that this discrepancy may be contributed

to different input statistics in different cortical areas. In particular, we conjectured that inputs to neurons in early sensory areas may be more stochastic (or unreliable) than in higher cortical areas. To study this issue quantitatively, we investigated how clustering in our model depends on the reliability of assembly activations. To this end, we repeated the experiment with simultaneously activated assemblies but with a reduced number of active neurons in each assembly activation. When an assembly was activated for 300 ms, a fraction  $p$  of assembly neurons was chosen randomly which increased their firing rate to 35 Hz as before, while the rest remained at the background rate of 1 Hz. In successive simulations, we reduced the fraction  $p$  of active neurons in each assembly from 1.0 (all neurons were active at each pattern presentation) down to 0.5 (half of the neurons were active at each pattern presentation). The mean number of represented assemblies as a function of the neuron activation probability in each assembly is shown in **Figure 5B** (25 independent trials). The number of assembly-synapse clusters on dendritic branches decreased with the neuron activation probability. A moderate stochasticity is tolerated by the rewiring dynamics for temporally isolated patterns and two simultaneously activated assemblies. No stochasticity is tolerated, however, when more than two assemblies are activated simultaneously at times of pattern presentation (**Figure 5B**). In all simulations, a number of branches did not receive any synaptic cluster. This was expected as the neuron had 12 branches while only 8 assembly patterns were presented as input. These branches hosted rather stable synaptic connections originating from various input assemblies. In conclusion, this result predicts a continuum of synaptic clustering in PCs where the extent of synaptic clustering depends on the reliability of presynaptic assembly activations. This finding may provide an explanation for diverse results on synaptic clustering in cortex.

*Influence of assembly overlap:* In the simulations so far, we considered disjoint input assemblies. In contrast, it was proposed that associative memories learned within short time intervals will be represented by overlapping populations of neurons (Silva et al., 2009; Rogerson et al., 2014). This claim was recently verified by experimental studies of medial temporal lobe (Ison et al., 2015), the hippocampus (Cai et al., 2016), and the lateral amygdala (Rashid et al., 2016). Their results show that linking of memories relies on the overlap of distributed neuronal assemblies.

To test whether synaptic clustering emerges also for input assemblies with various amounts of neuron overlap, we considered a setup where some of the neurons of each assembly were chosen from a shared pool of neurons. As a result, some neurons of the shared pool participated in two or more input assemblies (see section 4 for details). By increasing the fraction of shared neurons in the assemblies, we were able to increase the assembly overlap. When repeating this procedure in 25 independent trials, the overlap between a given pair of assemblies was  $3.11 \pm 2.48\%$  (mean  $\pm$  SD),  $6.15 \pm 3.53\%$ ,  $9.38 \pm 4.19\%$ , and  $12.5 \pm 5.11\%$  for a fraction of shared neurons per input assembly of 25, 50, 75, and 100%, respectively. After 17 min of assembly activations, we determined the number of represented assemblies on the neuron. The results for 25



independent trials are summarized in **Figure 5C** (solid line). Our results show that synapses are organized on dendrites in a clustered manner; this is true even for assemblies that share a significant portion of neurons, where the number of represented assemblies decreases slightly with increasing assembly overlap.

In section 2.4, we showed that our rewiring rule can support memory protection by recruiting branches sequentially as new assemblies were activated. To test whether this ability of the model depends on disjoint input assemblies, we repeated the experiment described in section 2.4 but with overlapping assemblies and various assembly overlaps. Overlapping assemblies were created as described above. After 17 min of assembly activations, the mean number of represented assemblies was  $6.80 \pm 0.75$ ,  $6.08 \pm 1.16$ ,  $5.32 \pm 1.16$ , and  $4.52 \pm 1.17$  for mean assembly overlaps of  $3.11 \pm 2.48\%$ ,  $6.15 \pm 3.53\%$ ,  $9.38 \pm 4.19\%$ , and  $12.5 \pm 5.11\%$ , respectively (mean and SD over 25 independent trials). For all considered overlaps, synapses were still segregated in an assembly-specific manner. However, the number of represented assemblies significantly decreased for mean assembly overlaps of 6.15% (**Figure 5C**, dashed line). In summary, our model predicts that associated memories with overlapping assembly representations are segregated in an assembly-specific manner onto different dendritic branches and that this segregation can indeed support memory protection.

## 2.6. Analysis of Stochastic Rewiring Dynamics

The stochastic rewiring dynamics considered in this work is described by the general SDE (1), reiterated here for convenience:

$$d\theta_{ki}(t) = \eta \left( f_{ki}^S(t) + f_{ki}^L(t) \right) dt + \sqrt{2\eta T} dW_{ki}, \quad (3)$$

where the first term  $f_{ki}^S$  describes structural constraints on the connectivity, the second term  $f_{ki}^L$  captures the functional goal of the plasticity process, and the third term models stochastic contributions.

It has been shown by Kappel et al. (2015) that the long-term behavior of this dynamics can be described concisely if the deterministic parameter changes are defined through two probability distributions:

$$f_{ki}^S(t) = \left. \frac{\partial}{\partial \theta_{ki}} \log p_S(\boldsymbol{\theta}) \right|_{\boldsymbol{\theta}(t)} \quad \text{and} \quad f_{ki}^L(t) = \left. \frac{\partial}{\partial \theta_{ki}} \log p_L(\boldsymbol{\theta}) \right|_{\boldsymbol{\theta}(t)}, \quad (4)$$

where  $p_S$  is a probability distribution that defines the structural constraints (the *structural prior* on parameters), and  $p_L$  is a distribution that defines functionally preferred parameter settings (the *functional likelihood* of parameters). Preferred parameter settings have high probability in these distributions, others have low probability. Note that without the noise term, Equation (1) together with Equation (4) constitute continuous-time gradient ascent dynamics on the objective  $\log p_S(\boldsymbol{\theta}) + \log p_L(\boldsymbol{\theta}) = \log(p_S(\boldsymbol{\theta})p_L(\boldsymbol{\theta}))$ . With the inclusion of the noise term, Equation (1) defines Langevin sampling dynamics (Welling and Teh, 2011), that is, parameter configurations are visited in the long run according to the distribution

$$p^*(\boldsymbol{\theta}) \propto (p_S(\boldsymbol{\theta})p_L(\boldsymbol{\theta}))^{1/T}. \quad (5)$$

In other words, the probability density to observe some specific  $\boldsymbol{\theta}$  is given—after an initial transient period (often called burn-in time) where parameters move from the potentially low-probability initial setting to a high-probability region—by  $p^*(\boldsymbol{\theta})$ . We say that the network samples network configurations

from  $p^*(\theta)$ . Since this  $p^*(\theta)$  contains the product of the two distributions  $p_S$  and  $p_L$ , high-probability parameter settings are only possible if they fulfill both the structural and the functional demands as defined by these distributions. Since  $\theta$  defines both the synaptic strengths and the connectivity structure (a synapse  $ki$  is connected if  $\theta_{ki} > 0$ ), this stochastic process automatically samples network connectivity structures.

The structural prior corresponding to the structural plasticity term  $f_{ki}^S$  is discussed in detail in section 4. The functional term  $f_{ki}^L(t)$  is given by

$$f_{ki}^L(t) = \begin{cases} c_L \Gamma_k(t)(x_i(t) - \gamma(1 - x_i(t))), & \text{if } \theta_{ki}(t) > 0 \\ & \text{(functional connection)} \\ 0, & \text{if } \theta_{ki}(t) \leq 0 \\ & \text{(non-established connection).} \end{cases} \quad (6)$$

Synapses that contribute or are co-active with dendritic plateau potentials are potentiated, while others are depressed and eventually retract, making room for new connections.

To interpret this functional term from the perspective of the sampling distribution of the stochastic dynamics, we want to exhibit a corresponding functional likelihood  $p_L$ . We show in section 4 that this distribution factors into branch-specific distributions

$$p_L(\theta) = \prod_k p_L(\theta_k), \quad (7)$$

so each branch optimizes its likelihood independently from other branches. For the functional likelihood, we consider the situation where the neuron is exposed for an extended time  $t_{\max}$  to input spike trains  $x_i$  according to some distribution and it stochastically responds with branch spikes. We can interpret the functional dynamics using the following functional likelihood (see section 4)

$$p_L(\theta_k) = \frac{1}{Z_L} \exp \left( \frac{\hat{c}_L}{t_{\max}} \int_0^{t_{\max}} \Gamma_k(s) \left( \sum_i w_{ki} x_i(s) - \gamma \sum_i w_{ki} (1 - x_i(s)) \right) ds \right)_{X, \Gamma_k}, \quad (8)$$

where  $Z_L$  is a normalizing factor (we consider here parameters bounded in the range  $[\theta_{\min}, \theta_{\max}]$  to keep this factor bounded) and  $\hat{c}_L > 0$  is a constant. The angular brackets  $\langle \cdot \rangle_{X, \Gamma}$  denote an average over stochastic realizations of inputs  $x_i$  and branch spikes  $\Gamma_k$ . Because of the function  $\Gamma_k$  in the integral, only times at which a plateau potential is present at the branch contribute to the integral. In combination with the first term in the brackets, the probability of the functional likelihood is high if efficacies of those synapses are large that are active slightly before or during a plateau potential. In combination with the second term in the brackets, the probability of the functional likelihood is high if efficacies of those synapses are small that are inactive slightly before or during a plateau potential. The constant  $\gamma > 0$  trades off the importance of large active vs. small inactive synapses.

Hence, the rewiring dynamics prefers parameter settings where inputs that jointly recruit dendritic regenerative events (i.e., temporally correlated inputs) have strong synaptic connections to the branch. Under the assumption that such inputs are also functionally related (corresponding for example to a specific motor behavior), the rewiring dynamics contributes to functional synaptic clustering (Kastellakis et al., 2015). While cortical connectivity is sparse even at the local scale, it has been estimated that potential (i.e., possibly realizable) connectivity is huge—a spatial scale of a few hundred microns—and every pair of PCs could be potentially connected (Chklovskii et al., 2004; Kalisman et al., 2005). Our analysis shows that in this large space of potential connectivity structures, synaptic rewiring based on simple dendritic-spike-dependent plasticity rules could implement a stochastic search algorithm that favors functionally clustered synaptic configurations, as proposed by Chklovskii et al. (2004).

### 3. DISCUSSION

We have proposed in this article a model for synaptic rewiring in neurons with dendritic non-linearities. We have observed that simple plasticity rules that depend on postsynaptic dendritic spikes lead to robust synaptic clustering. The use of the synaptic sampling framework allowed us to analyze the connectivity dynamics as a sampling process from a distribution of network configurations. Our analysis revealed that synaptic clusters are preferred configurations in this sampling process. Cichon and Gan (2015) found that dendritic spike events for different motor tasks were segregated onto different tuft branches of layer V PCs in mouse motor cortex. They also found that this segregation was necessary to avoid a detrimental effect of training on a previously learned task. Our model can explain these findings. First, we found that branch activity is segregated onto different dendritic branches if it is elicited by different input assemblies. Second, we found that this segregation allows memory retention in the sense that synaptic clusters are retained after memorization of new assembly input patterns.

#### 3.1. Synaptic Plasticity and Rewiring

Throughout this article, we assumed that strong presynaptic activity coincident with a postsynaptic dendritic spike leads to LTP whereas a postsynaptic dendritic spike paired with weak presynaptic activity leads to LTD. While these plasticity dynamics are qualitatively consistent with experimental findings both for LTP (Golding et al., 2002; Lisman and Spruston, 2005; Gordon et al., 2006; Gambino et al., 2014; Cichon and Gan, 2015; Brandalise et al., 2016) and LTD (Holthoff et al., 2004), the exact plasticity rules and their relation to dendritic regenerative events are still to be determined. For example, experimental results of a study by Cichon and Gan (2015) indicated that at distal dendrites of layer V PCs, presynaptic activity during dendritic spikes leads to LTP, which is consistent with our model. However, different from previous findings, they found that LTD was induced by presynaptic activity that preceded the dendritic spike. To test whether our model is sensitive to such alterations in the plasticity rules, we reran the simulations for **Figure 3** with such depression

(see section 4 for details). We found that the model behaved comparably ( $7.28 \pm 0.60$  represented assemblies, mean  $\pm$  SD over 25 independent trials).

We found that the number of input patterns that can be stored by a neuron is increased if dendritic-spike-dependent plasticity is extended by an STDP mechanism where pre-before-post spike pairs induce depression. This phenomenon has already been described and analyzed in a modeling study by Legenstein and Maass (2011). In general, the back-propagating action potential provides a means for communicating global information to the dendritic branches, that is, it can indicate to all branches that the neuron is active for the current input pattern. With STDP, a pattern that does already activate enough branches to elicit strong firing of the neuron will depress synapses at inactive branches. This competition reduces the number of activated branches per pattern and thus increases the number of patterns that can be stored in a single neuron. While we showed that clustering is still observed without this mechanism (Figure 2), it is beneficial in a sequential learning paradigm, as it keeps branches free to store patterns presented later in the sequence (Figure 4).

### 3.2. Protection of Memories

Our results indicate that synaptic clustering may protect older memories (Figure 4). It can thus avoid catastrophic forgetting, a well-known deficiency of standard incremental learning algorithms for neural networks (McCloskey and Cohen, 1989; French, 1999; Kirkpatrick et al., 2017). The mechanism is simple: a branch with a synaptic cluster from a specific input pattern is silent to most other patterns. Since LTP and LTD is conditioned on the local dendritic spike, memories on a branch are protected for most other input patterns. At the same time, other branches can develop specificity to novel patterns as plasticity there is decoupled from protected branches. However, this does not shelter memories completely. There are still mechanisms in our model that are detrimental to old memories. In particular, the noise term in Equation (1) will inevitably corrode memories. We did indeed observe weakening of older memories over time if they were not repeated, but on a time scale much longer than the one of pattern presentation (sections 1.4, 1.5 in the **Supplementary Material**). This suggests that synaptic clustering may provide a mechanism to avoid catastrophic forgetting temporarily. Additional consolidation mechanisms seem necessary to protect memories over long time scales, and we have investigated one way how consolidation can be integrated in our rewiring framework (section 1.6 in the **Supplementary Material**). An alternative role of dendritic structures in extending memory lifetime was proposed in Bono and Clopath (2017). We have demonstrated in this article that simple dendritic-spike-dependent rewiring can protect memories in a pattern memorization setup. In future work, one could apply these ideas to larger deep-learning networks in the context of continual learning tasks, possibly using simplified neuron and plasticity models with similar properties.

### 3.3. Related Work

Many theoretical studies have linked rewiring processes to cortical connection statistics (Deger et al., 2012; Fauth et al.,

2015), computation and memory (Knoblauch et al., 2014; Zheng and Triesch, 2014; Deger et al., 2016; Miner and Triesch, 2016; Spiess et al., 2016), and homeostasis (Gallinaro and Rotter, 2018). These studies were, however, based on point neuron models, where the phenomenons considered in the current study cannot be observed. It has been proposed by Rhodes (2008) that neurons with non-linear dendrites could serve as a recoding stage in order to orthogonalize input representations, that is, to reduce the overlap between different input patterns. It is discussed there that a dendritic-spike-dependent rewiring could be used to provide efficient wiring diagrams for this operation. Our model exhibits one possible implementation of such a rewiring strategy. In a theoretical study, the impact of active dendrites on the memory capacity of neural networks was investigated (Poirazi and Mel, 2001). They calculated that non-linear dendrites can vastly increase the number of patterns that can be stored by a simple non-spiking model of a pyramidal cell. They also considered a supervised structural plasticity rule, the “clusteron” (Mel, 1992), that was designed to cluster synaptic inputs while optimizing the memory performance of the neuron. Their simulations showed that this rule leads to memory capacities as predicted by their theoretical considerations. Our model shows that such clustering emerges also in a more detailed spiking neuron model with a simpler online dendritic-spike-dependent rewiring principle that does not need supervision. We exhibited a potential functional relevance of this clustering in terms of protection of memory but did not consider memory capacity *per se*. It would be interesting to study similar capacity questions in the proposed model. Memory capacity of networks of neurons with non-linear dendrites was also studied by Wu and Mel (2009). They also used a simple non-spiking neuron model which basically consisted of a “dendritic” layer with a threshold non-linearity and a somatic layer which summed the outputs of the binary dendritic elements. This article did not consider rewiring. Nevertheless, a simple plasticity rule was used that depended on the dendritic activation, which resembles some similarity with dendritic-spike-dependent plasticity considered here. Interestingly, they found that memory capacity could be increased if not only the dendritic activation but also the number of activated synapses at the dendrite was used to gate LTP. In our rewiring framework, such a rule could be applied to weak synaptic connections, but not to non-established connections as their dynamics are assumed independent of pre- and postsynaptic activity.

### 3.4. Relation to Experimental Literature and Experimentally Testable Predictions

The proposed model gives rise to a number of experimentally testable predictions. We have shown that temporally correlated input leads to the clustering of corresponding synaptic connections on single dendritic branches. This suggests that neuronal assemblies should connect to PCs preferentially in a branch-specific clustered manner. To verify this prediction, methods for labeling memory engrams (Tonegawa et al., 2015) could either be combined with tracing techniques or with engram reactivation and target cell  $\text{Ca}^{2+}$  imaging at dendritic branches. Another prediction of our model is that synaptic clustering



depends on input statistics. This finding may explain that no evidence for clustering has been observed in early sensory cortical areas (Jia et al., 2010). One could test this prediction more directly through glutamate uncaging at tuft dendrites. Our model predicts that concerted correlated uncaging at single branches should lead to a clustered strengthening of synapses at the branch. Conversely, a more unreliable uncaging where each uncaging site is activated only with a certain probability would not have this effect. Various experimental studies have indicated that associated memories are stored in overlapping assemblies (Ison et al., 2015; Cai et al., 2016; Rashid et al., 2016). Our model predicts that such associations should also be visible in the clustering structure, that is, neurons that respond to each of two associated memories would exhibit synaptic co-clustering of these memories within dendrites. This prediction is consistent with the findings of a modeling study by Kastellakis et al. (2016).

### 3.5. Conclusions

We have shown through computer simulations and theoretical analysis that synaptic clustering emerges from simple rewiring dynamics based on dendritic-spike-dependent synaptic plasticity. This result provides support for the synaptic clustering hypothesis. We further found that this clustering in turn may serve as a mechanism to protect memories from subsequent modifications on a medium time scale. More generally, our results emphasize the importance of dendritic structures in learning and memory and demonstrate the effectiveness of synaptic rewiring to find sparse network structures in the high-dimensional search space over potential connectivities.

## 4. MATERIALS AND METHODS

Here, we provide details to the models and simulations.

### 4.1. Neuron Model

We assume a spiking neuron model with  $N_b$  independent dendritic compartments. Here and in the following, we will indicate variables and parameters of dendritic branches with a  $b$  superscript or subscript. Presynaptic spikes are transformed at the synapse  $ki$ —the synapse from presynaptic neuron  $i$  to branch  $k$  of the postsynaptic neuron—into alpha-shaped postsynaptic currents, which are weighted by the corresponding synaptic efficacies  $w_{ki}$ . The total current at branch  $k$  is given by

$$I_k^b(t) = \sum_i w_{ki} \sum_f \alpha(t - t_{pre,i}^{(f)}) \quad (9)$$

where  $t_{pre,i}^{(f)}$  is the  $f^{\text{th}}$  spike-time of presynaptic neuron  $i$ . The two sums run over all presynaptic neurons  $i$  and all firing times  $t_{pre,i}^{(f)} < t$  of neuron  $i$ . Each individual alpha-current is given by

$$\alpha(s) = \frac{s}{\tau_{syn}} \exp\left(1 - \frac{s}{\tau_{syn}}\right) H(s), \quad (10)$$

where  $H$  denotes (here and in the following) the Heaviside step function (for parameter values, see **Table 1**). The membrane

**TABLE 1 |** Neuron parameters for computer simulations.

Branch parameters		
$V_{th}^b$	Stochastic threshold	−55 mV
$E_L^b$	Leak reversal potential	−70 mV
$R_b$	Membrane resistance	40 M $\Omega$
$C_b$	Membrane capacity	250 pF
$\tau_s$	Time constant of sodium spikelet	4 ms
$V_{ds}$	Amplitude of plateau potential	−30 mV
$V_s$	Initial amplitude of sodium spikelet (rel. to $V_{ds}$ )	5 mV
$C_{ds}$	Scale factor for plateau duration	0.04 s <sup>2</sup> /mV
$\rho_0^b$	Instantaneous firing rate at threshold	2.5 Hz
$\beta^b$	Sensitivity of spike emission	0.5 mV <sup>−1</sup>
$\Delta_{min}^{ds}$	Minimum duration of a plateau potential	20 ms
$\Delta_{max}^{ds}$	Maximum duration of a plateau potential	300 ms
Soma parameters		
$V_{th}^{soma}$	Stochastic threshold	−55 mV
$E_L^{soma}$	Leak reversal potential	−70 mV
$R_m$	Membrane resistance	40 M $\Omega$
$C_m$	Membrane capacity	250 pF
$R_l$	Longitudinal resistance along dendrites and apical trunk	2 $\Omega$
$\Delta_{abs}^{soma}$	Absolute refractory period	5 ms
$\rho_0^{soma}$	Instantaneous firing rate at threshold	2.5 Hz
$\beta^{soma}$	Sensitivity of spike emission	0.5 mV <sup>−1</sup>
Synapse parameters		
$\tau_{syn}^b$	Synaptic time constant of branch	2 ms
$\theta_{min}$	Lower bound of synaptic parameter	−2
$\theta_{max}$	Upper bound of synaptic parameter	8

potential of each branch ( $V_k^b$ ; branch potential) follows the dynamics of a leaky integrator of the current with resting potential  $E_L^b$ . In particular, the dynamics of potential  $V_k^b$  of branch  $k$  is given by

$$C_b \frac{dV_k^b(t)}{dt} = -\frac{1}{R_b} (V_k^b(t) - E_L^b) + I_k^b(t), \quad (11)$$

with membrane resistance  $R_b$  and membrane capacity  $C_b$  (the membrane time constant is given by  $\tau_b = R_b C_b$ ).

Branch spikes are elicited by the branch in a stochastic manner. In particular, the voltage-dependent instantaneous firing rate  $f_k^b(t)$  is given by

$$f_k^b(t) = \begin{cases} \rho_0^b \exp(\beta^b (V_k^b(t) - V_{th}^b)), & \text{if } \frac{dV_k^b(t)}{dt} > 0 \\ 0, & \text{else,} \end{cases} \quad (12)$$

where  $V_{th}^b$  is the threshold for branch-spike initiation,  $\rho_0^b$  is the instantaneous firing rate at threshold, and  $\beta^b$  is the sensitivity of spike emission (for parameter values, see **Table 1**). We denote the times of branch-spike onsets at branch  $k$  as  $t_k^{(1)}, t_k^{(2)}, \dots$ . A branch spike includes a brief sodium spikelet followed by a plateau potential that lasts for up to a few hundreds of milliseconds, see **Figure 1B**. Assuming that branch  $k$  has initiated its last dendritic

spike at time  $\hat{t}_k(t)$ , the evolution of  $V_k^b(t)$  immediately after initiation is given by

$$V_k^b(t) = \eta_k(t - \hat{t}_k(t)) \quad \text{for } \hat{t}_k(t) < t \leq \hat{t}_k(t) + \Delta_k^{\text{ds}}(\hat{t}_k(t)), \quad (13)$$

where the function  $\eta_k$  describes the form of the branch spike (see below). The duration of the dendritic spike is proportional to the slope of  $V_k^b(t)$  at the last spike time of the branch, i.e.,

$$\Delta_k^{\text{ds}}(\hat{t}_k(t)) = c_{\text{ds}} \left. \frac{dV_k^b(t)}{dt} \right|_{t=\hat{t}_k(t)}, \quad (14)$$

where  $c_{\text{ds}}$  is a scaling parameter. In our simulations, we clipped the duration between a minimum of  $\Delta_{\text{min}}^{\text{ds}} = 20$  ms and a maximum of  $\Delta_{\text{max}}^{\text{ds}} = 300$  ms to avoid unnaturally brief and long plateau durations (for other parameter values, see **Table 1**). The functional relation between input strength and the duration of the plateau potential  $\Delta_k^{\text{ds}}(\hat{t}_k(t))$  is shown in **Figure 1C**. The shape of the dendritic spike is given by a brief sodium spikelet followed by a plateau (Antic et al., 2010):

$$\eta_k(s) = V_s \exp\left(-\frac{s}{\tau_s}\right) H(s) + V_{\text{ds}} \left[ H(s) - H\left(s - \Delta_k^{\text{ds}}(\hat{t}_k(s))\right) \right]. \quad (15)$$

The first term in Equation (15) models the sodium spikelet, and the second term models the plateau potential. The sodium spikelet decays exponentially with time constant  $\tau_s$  with initial amplitude  $V_s$  (added to the amplitude  $V_{\text{ds}}$  of the plateau potential). For  $t > \hat{t}_k(t) + \Delta_k^{\text{ds}}(\hat{t}_k(t))$  the membrane potential  $V_k^b(t)$  again follows the dynamics given by Equation (11). As an example, the evolution of the dendritic membrane potential for an arbitrary chosen input is depicted in **Figure 1B**.

During the time of the plateau potential, the dendritic branch is absolute refractory, that is, no further regenerative event can be initiated. The propagation of dendritic plateau potentials toward the soma produces a sustained depolarization of the cell body. In general, the currents from dendritic compartments to the soma are driven by the difference between the membrane potential at the branch and the membrane potential of the soma. The somatic compartment sums currents from all dendrites (with a certain degree of amplitude attenuation).

The membrane potential at the soma of the neuron is given by

$$C_m \frac{dV^{\text{soma}}(t)}{dt} = -\frac{1}{R_m} (V^{\text{soma}}(t) - E_L^{\text{soma}}) + I^{\text{soma}}(t), \quad (16)$$

where  $E_L^{\text{soma}}$  is the resting potential. The membrane time constant of the neuron is given by  $\tau_m = R_m C_m$  and  $I^{\text{soma}}(t)$  is

$$I^{\text{soma}}(t) = \frac{1}{R_l} \sum_k \max\{0, V_k^b(t) - V^{\text{soma}}(t)\}, \quad (17)$$

where  $R_l$  represents the total of the longitudinal resistance along the dendrites and apical trunk. Note that we restrict the currents to non-negative values since we only simulate forward currents, i.e., currents from the dendritic branches to the soma.

Somatic spikes are characterized by their spike times  $t^{(f)}$ . The instantaneous firing rate is given analogously to the branch spike mechanism by

$$f^{\text{soma}}(t) = \begin{cases} \rho_0^{\text{soma}} \exp(\beta^{\text{soma}}(V^{\text{soma}}(t) - V_{\text{th}}^{\text{soma}})), & \text{if } \frac{dV^{\text{soma}}(t)}{dt} > 0 \\ 0, & \text{else.} \end{cases} \quad (18)$$

Immediately after  $t^{(f)}$ , the somatic potential enters a refractory period of duration  $\Delta_{\text{abs}}^{\text{soma}}$  in which  $V^{\text{soma}}(t)$  is set to  $E_L^{\text{soma}}$ . For  $t > t^{(f)} + \Delta_{\text{abs}}^{\text{soma}}$  the dynamics is again given by Equation (16). To illustrate the dynamics of the membrane potential, a neuron which receives arbitrary input was simulated. The results of this simulation are shown in **Figure 1B**.

## 4.2. Plasticity and Rewiring

Here, we summarize the plasticity and rewiring dynamics in the network.

### 4.2.1. Synaptic Parameters and Efficacies

Each dendritic compartment has a set of potential synaptic connections that could be realized by some parameter setting. However, at each point in time, only a subset of these connections is realized by functional synapses. More precisely, we maintain one parameter  $\theta_{ki}$  for each synapse  $ki$  from input neuron  $i$  to branch  $k$  of the neuron. This parameter encodes (a) whether the synapse is functional and (b) the synaptic efficacy  $w_{ki}$  of the synapse (if it is functional). In particular,  $\theta_{ki} > 0$  indicates a functional synapse, while  $\theta_{ki} \leq 0$  indicates that the connection is not established. The synaptic weight  $w_{ki}$  for a functional synapse, and 0 otherwise. Hence, the mapping from the parameter  $\theta_{ki}$  to the synaptic weight is given by  $w_{ki} = c_\theta \max\{0, \theta_{ki}\}$  where  $c_\theta = 1$  nA is the slope of this mapping for  $\theta_{ki} > 0$  (for other parameter values, see **Table 2**).

### 4.2.2. Parameter Dynamics

The stochastic dynamics of parameters are given by

$$d\theta_{ki}(t) = \eta \left( f_{ki}^{\mathcal{S}}(t) + f_{ki}^{\mathcal{L}}(t) \right) dt + \sqrt{2\eta T} d\mathcal{W}_{ki}, \quad (19)$$

**TABLE 2 |** Plasticity parameters for computer simulations.

$\eta$	Learning rate	0.002
$T$	Temperature parameter	0.3
$c_{\mathcal{L}}$	Scale factor of functional plasticity term	1.5
$\gamma$	Scale factor for depression of functional plasticity term	0.2
$\lambda$	Scale factor for steepness of sigmoid in structural plasticity term	10
$c_w$	Scale factor for steepness of sigmoid in synapse count	0.55 nA <sup>-1</sup>
$c_\theta$	Slope of $\theta$ to $w$ mapping (for $\theta > 0$ )	1 nA
$c_{\text{STDP}}$	STDP, LTD factor	3.2
$\text{STDP}_{\text{th}}$	Threshold for STDP initiation	-67 mV
$\tau_x$	Time constant of presynaptic trace	20 ms

where  $\eta > 0$  is a small learning rate and  $\mathcal{W}_{ki}$  denote increments of a standard Wiener process, thus implementing random walk behavior. The temperature  $T > 0$  is a constant that scales the strength of this stochastic component. In the simulations, the parameters  $\theta_{ki}$  were clipped to the range  $[\theta_{\min}, \theta_{\max}]$ . The upper bound is necessary to avoid that the synaptic efficacies become arbitrary large. The lower bound is useful to avoid that parameters become very negative. Such parameters would be stuck for a long time in the negative domain and hence would be unlikely to be re-established, which would slow down the rewiring dynamics. The term in the brackets defines the deterministic part of the dynamics. It is divided here into two components. The first term  $f_{ki}^S$  describes changes that are used to enforce structural constraints on the connectivity. The functional goal of the plasticity process is captured by the second term  $f_{ki}^L$ . The deterministic terms are chosen such that they vanish for  $\theta_{ki}(t) \leq 0$ . Hence, for non-established synapses, only the stochastic term remains.

For the structural term, we define a soft count of synaptic connections on branch  $k$

$$N(\theta_k) = \sum_i 2 \left( \sigma(c_w w_{ki}) - \frac{1}{2} \right), \quad (20)$$

where  $\sigma$  denotes the logistic sigmoid function  $\sigma(x) = \frac{1}{1 + \exp(-x)}$ . For a non-established synapse ( $w_{ki} = 0$ ), the term in the brackets is 0 and this potential connection is not counted. For increasing efficacies larger than 0, this term approaches 1 with the scaling constant  $c_w$  defining the speed of this approach. Using this soft count, the structural term is given by

$$f_{ki}^S(t) = -2\lambda c_w c_\theta [1 - \sigma(\lambda(N_{\text{syn}} - N(\theta_k(t))))] \dot{\sigma}(c_w w_{ki}(t)) H(\theta_{ki}(t)), \quad (21)$$

where the constant  $N_{\text{syn}}$  defines a soft upper bound on the number of active synapses per branch (we used  $N_{\text{syn}} = 20$  in all simulations), the constant  $\lambda > 0$  defines how hard this soft bound is enforced, and  $\dot{\sigma}$  denotes the derivative of the logistic sigmoid.

The functional term combines a presynaptic activity trace with a postsynaptic term that indicates the presence of a dendritic spike. The exponential presynaptic activity trace with time constant  $\tau_x > 0$  is given by:

$$x_i(t) = \sum_{f \in \mathcal{F}_i(t)} \exp\left(-\frac{t - t_{\text{pre},i}^{(f)}}{\tau_x}\right) H\left(t - t_{\text{pre},i}^{(f)}\right), \quad (22)$$

where  $\mathcal{F}_i(t) = \{f : t_{\text{pre},i}^{(f)} \leq t\}$  is the set of all indices of input spikes with spike times before  $t$ . The presence of a postsynaptic dendritic spike is formalized as a function  $\Gamma_k$  which is 1 if a plateau potential is present at time  $t$  and 0 otherwise. More formally, using  $\hat{t}_k(t)$  to indicate the time of the last branch spike initiation before  $t$ , i.e.,  $\hat{t}_k(t) = \max_f \{t_k^{(f)} : t_k^{(f)} \leq t\}$ , we define

$$\Gamma_k(t) = H(t - \hat{t}_k(t)) - H\left(t - \hat{t}_k(t) - \Delta_k^{\text{ds}}(\hat{t}_k(t))\right), \quad (23)$$

where  $\Delta_k^{\text{ds}}(\hat{t}_k(t))$  is the length of this plateau potential as defined in Equation (14). The functional term is then given by

$$f_{ki}^L(t) = c_{\mathcal{L}} \Gamma_k(t) (x_i(t) - \gamma(1 - x_i(t))) H(\theta_{ki}(t)) \quad (24)$$

with constants  $c_{\mathcal{L}} > 0$  and  $\gamma > 0$ .

#### 4.2.3. STDP

Denoting the postsynaptic somatic spike times by  $t^{(1)}, t^{(2)}, \dots$ , we define the postsynaptic spike trains as sums of Dirac delta pulses, i.e.,  $S(t) = \sum_f \delta(t - t^{(f)})$ . Using this notation, the STDP update term is given by

$$f_{ki}^{\text{STDP}}(t) = \begin{cases} -c_{\text{STDP}} x_i(t) S(t), & \text{if } V_k^b(t) \geq \text{STDP}_{\text{th}} \\ 0, & \text{otherwise,} \end{cases} \quad (25)$$

where  $c_{\text{STDP}} > 0$  is a constant,  $x_i$  is an exponential trace of the presynaptic spike times with time constant  $\tau_x$  as defined in Equation (22), and where  $\text{STDP}_{\text{th}}$  is the threshold for STDP initiation (for parameter values, see **Table 2**). Note that this update models the induction of depression for a pre-before-post spike pair (an inverted STDP rule) as found for synapses in the distal dendrites (Kampa et al., 2007). The plasticity dynamics with STDP for the parameters  $\theta_{ki}$  is therefore

$$d\theta_{ki}(t) = \eta H(\theta_{ki}(t)) \left( f_{ki}^S(t) + f_{ki}^L(t) + f_{ki}^{\text{STDP}}(t) \right) dt + \sqrt{2\eta T} d\mathcal{W}_{ki}. \quad (26)$$

Note that we include the term  $H(\theta_{ki}(t))$  in Equation (26) in order to stress again that only functional synapses are subject to non-stochastic parameter changes.

#### 4.2.4. Alternative Plasticity Rule

To test the dependence of the results on the specific form of the plasticity rule, we considered an alternative rule that models findings in a study by Cichon and Gan (2015). Here, presynaptic activity during a dendritic spike still leads to LTP, but, LTD is induced by presynaptic activity *preceding* a dendritic spike. The functional term in this plasticity rule combines two presynaptic activity traces [with fast (LTP) and slow (LTD) decay] with postsynaptic terms that indicate the onset and the presence of a dendritic spike. The fast decaying presynaptic activity trace  $x_{ki}^{\text{LTP}}(t)$  with time constant  $\tau_{x^{\text{LTP}}} = 20$  ms is incremented at each presynaptic spike of neuron  $i$  that appears within the time period of a dendritic spike of branch  $k$ . The trace is given by:

$$x_{ki}^{\text{LTP}}(t) = \sum_{f \in \mathcal{F}_{ki}^{\text{LTP}}(t)} \exp\left(-\frac{t - t_{\text{pre},i}^{(f)}}{\tau_{x^{\text{LTP}}}}\right) H\left(t - t_{\text{pre},i}^{(f)}\right), \quad (27)$$

where  $\mathcal{F}_{ki}^{\text{LTP}}(t) = \{f : t_{\text{pre},i}^{(f)} \leq t \wedge \Gamma_k(t_{\text{pre},i}^{(f)}) = 1\}$  is the set of all indices of input spikes of neuron  $i$  with spike times before  $t$  and during the presence of a dendritic spike in branch  $k$  (recall that the function  $\Gamma_k$  is 1 during a dendritic spike and 0 otherwise). The slow decaying presynaptic activity trace  $x_{ki}^{\text{LTD}}(t)$  is defined alike. However, this trace is only incremented at each presynaptic spike

if there is no postsynaptic dendritic spike. The trace  $x_{ki}^{\text{LTD}}(t)$  with time constant  $\tau_{x^{\text{LTD}}} = 500$  ms is defined as

$$x_{ki}^{\text{LTD}}(t) = \sum_{f \in \mathcal{F}_{ki}^{\text{LTD}}(t)} \exp\left(-\frac{t - t_{\text{pre},i}^{(f)}}{\tau_{x^{\text{LTD}}}}\right) H\left(t - t_{\text{pre},i}^{(f)}\right), \quad (28)$$

where  $\mathcal{F}_{ki}^{\text{LTD}}(t) = \{f: t_{\text{pre},i}^{(f)} \leq t \wedge \Gamma_k(t_{\text{pre},i}^{(f)}) = 0\}$  is the set of all indices of input spikes of neuron  $i$  with spike times before  $t$  and in the absence of a postsynaptic dendritic spike. Similar to somatic spike trains, we formalize the dendritic spike train as a sum of Dirac delta pulses at times of dendritic spike onset  $\Lambda_k(t) = \sum_f \delta(t - t_k^{(f)})$ , where  $t_k^{(f)}$  denotes the onset time of the  $f^{\text{th}}$  spike of dendrite  $k$ . The functional term is then given by

$$\hat{f}_{ki}^{\mathcal{L}}(t) = [c_{\mathcal{L}}^+ x_{ki}^{\text{LTP}}(t) \Gamma_k(t) - c_{\mathcal{L}}^- x_{ki}^{\text{LTD}}(t) \Lambda_k(t)] H(\theta_{ki}(t)), \quad (29)$$

where the constants  $c_{\mathcal{L}}^+ = 6$  and  $c_{\mathcal{L}}^- = 2$  scale the contribution of potentiation and depression. The structural term and the STDP term are not altered and are kept as defined above (Equation (21) and (25), respectively). The alternative plasticity dynamics for the parameters  $\theta_{ki}$  is therefore,

$$d\theta_{ki}(t) = \eta H(\theta_{ki}(t)) \left( f_{ki}^{\mathcal{S}}(t) + \hat{f}_{ki}^{\mathcal{L}}(t) + f_{ki}^{\text{STDP}}(t) \right) dt + \sqrt{2\eta T} dW_{ki}. \quad (30)$$

### 4.3. Experimental Design and Statistical Analysis

All simulations as well as statistical analyses were performed using Python v3.5.3 with SciPy v0.18.1 and NumPy v1.13.0. The model was simulated with a time step of 1 ms. We used the Euler method for the integration of the deterministic neuron dynamics, and we used the Euler-Maruyama method for the integration of the stochastic dynamics of the synaptic parameters. Every neuron state, that is, membrane potential, remaining refractory period, etc., as well as the synaptic parameters are represented by fixed size state vectors that are updated at each time step of the simulation.

D'Agostino-Pearson tests were used to assess normality. Statistical analyses were performed using different tests as appropriate, as stated in the main text and the figure legends. All of the data is presented as mean  $\pm$  SD. Significant levels for all significance tests were set at  $p \leq 0.05$ .

Simulation parameters are listed in **Tables 1, 2**. All simulations were based on the same parameter sets. Parameters used in the simulations with the alternative plasticity rule are given in section 4.3.3. To investigate the impact of the parameters on the simulation results we performed a one-at-a-time sensitivity analysis (see section 1.2 in the **Supplementary Material** for more detail).

#### 4.3.1. Details to Simulations for Figure 2

A total of 320 input neurons were divided into eight disjoint assemblies of 40 neurons each. Neurons within each input assembly were preferentially active together. More precisely, at times  $t_{\text{start}} + k(\Delta t_{\text{act}} + \Delta t_{\text{delay}})$ , with  $0 \leq k < N_p$  and  $N_p$

being the number of input patterns, one assembly was chosen uniformly at random from the set of all input assemblies and each neuron in the assembly emitted a 35 Hz Poisson spike train (randomly generated at every pattern presentation) for  $\Delta t_{\text{act}} = 300$  ms. In addition, to model background activity, each neuron from the whole input population was spiking at a rate of 1 Hz. The delay between two pattern presentations (the delay between successive active assemblies) was set to  $\Delta t_{\text{delay}} = 200$  ms. During this time and from  $t = 0$  to  $t_{\text{start}} = 200$  ms only the background noise was active. We assume that each input neuron could potentially establish a connection to each branch. Note that this is a reasonable assumption if the neurons are considered to be located within the same cortical column (Chklovskii et al., 2004; Kalisman et al., 2005). Input neurons were initially connected to branches randomly, such that exactly 20 synapses were established on each branch. Synaptic parameters of these connections were independently drawn from a uniform distribution on the interval [4, 8]. The soft maximum of synapses per branch ( $N_{\text{syn}}$ ) was also set to 20. A total of  $N_p = 2,000$  patterns as described above were presented to the neuron. Neuron and plasticity parameters were set according to **Tables 1, 2**, respectively.

#### 4.3.2. Details to Simulations for Figure 3

Patterns were generated and parameters were set as in simulations for **Figure 2**. In these simulations, however, we added the STDP update (Equation 25) to the parameter dynamics. All of the following simulations, that is, simulations with alternative plasticity rule, simulations for **Figures 4, 5**, were conducted with the STDP update.

#### 4.3.3. Details to Simulations With Alternative Plasticity Rule

We repeated the simulations for **Figure 3** but with an alternative functional term in the plasticity dynamics (according to Equation 29). We presented a total of  $N_p = 2,000$  patterns to the neuron and we used the following plasticity parameters:  $c_{\mathcal{L}}^+ = 6$ ,  $c_{\mathcal{L}}^- = 2$ ,  $\tau_{x^{\text{LTP}}} = 20$  ms,  $\tau_{x^{\text{LTD}}} = 500$  ms. All other parameters were set according to **Tables 1, 2**. Results of these simulations are given in the Discussion.

#### 4.3.4. Details to Simulations for Figure 4

Input assembly activations and background noise were generated as in the simulations described above (simulations for **Figures 2, 3**), but instead of choosing one assembly at random at times of pattern presentation, we activated the assemblies sequentially. In other words, we first activated exclusively assembly  $A_1$  250 times, then assembly  $A_2$ , then assembly  $A_3$ , etc. The delay between two pattern presentations was set to  $\Delta t_{\text{delay}} = 200$  ms as before. Parameters were set according to **Tables 1, 2**. At the end of all pattern presentations of an assembly, we evaluated for each branch which, if any, assemblies had synaptic clusters on that branch.

#### 4.3.5. Details to Simulations for Figure 5A

Patterns were generated similar as in simulations for **Figure 2**, but instead of choosing one assembly at random at times of



pattern presentations, up to 4 assemblies were chosen uniformly at random from the set of all input assemblies and were simultaneously activated. In addition, in successive simulations, we reduced the fraction  $p$  of active neurons in each assembly from 1.0 down to 0.5 (in steps of 0.1). The fraction  $p$  of assembly neurons was chosen randomly at each pattern presentation. These neurons increased their firing rate to 35 Hz as before, while the rest remained at the background rate of 1 Hz. In each of these simulations we presented a total of  $N_p = 2,000$  patterns to the neuron. Neuron and plasticity parameters were set according to **Tables 1, 2**, respectively.

#### 4.3.6. Details to Simulations for Figures 5B,C

These simulations were performed with non-disjoint input assemblies. Non-disjoint assemblies were created by randomly choosing for each assembly either 10, 20, 30, or all 40 neurons from a shared pool of input neurons consisting of 80, 160, 240, and 320 neurons, respectively. Since the number of neurons per assembly and the total number of input neurons were kept as before (40 assembly neurons and 320 input neurons), this procedure led to an overlap between assemblies, where the overlap between a given pair of assemblies was  $3.11 \pm 2.48\%$  (mean  $\pm$  SD, over 25 independent trials),  $6.15 \pm 3.53\%$ ,  $9.38 \pm 4.19\%$ , and  $12.5 \pm 5.11\%$  for a shared neuron pool of size of 80, 160, 240, and 320, respectively. Patterns of these assemblies were then created as given in details to simulations for **Figure 2**. A total of  $N_p = 2,000$  patterns were presented to the neuron. Parameters were set according to **Tables 1, 2**.

### 4.4. Details to Analysis of Stochastic Rewiring Dynamics

#### 4.4.1. Structural Prior

We first consider the structural plasticity term  $f_{ki}^S$ . Note that this term is meant to model the physical constraint that the number of synapses on a branch is bounded. We used

$$f_{ki}^S(t) = \begin{cases} -2\lambda c_w c_\theta [1 - \sigma(\lambda(N_{\text{syn}} - N(\theta_k(t))))] \dot{\sigma}(c_w w_{ki}(t)), & \text{if } \theta_{ki}(t) > 0 \quad (\text{functional connection}) \\ 0, & \text{if } \theta_{ki}(t) \leq 0 \quad (\text{non-established connection}), \end{cases} \quad (31)$$

where  $\sigma$  denotes the logistic sigmoid function,  $\dot{\sigma}$  denotes its derivative, and  $N(\theta_k)$  counts the number of active synapses at branch  $k$  in a smooth manner (see section 4). The constants  $\lambda$  and  $c_w$  will be discussed below. Equation (31) has a very simple interpretation. When the number of synapses at the branch is below  $N_{\text{syn}}$ , the term in the brackets is approximately 0, and the structural term has a tiny influence on the dynamics. However, when the number of synapses is above  $N_{\text{syn}}$ , this term is approximately 1, which will tend to decrease synaptic efficacies until the desirable synapse population size at the branch is reestablished. Due to the last term,  $\dot{\sigma}(c_w w_{ki}(t))$ , which is 0.25 for zero weights and then decreases to 0, weak synapses depress faster than stronger ones. This is reasonable since strong synapses are expected to play a more important role in the function of the neuron, and it is consistent with the experimental finding that large dendritic spines tend to be more stable than small ones. The constant  $c_w$  defines how fast this term decreases to 0.

We show below that this structural term in the stochastic dynamics corresponds to a structural prior  $p_S(\theta)$  that factorizes into one structural prior per branch

$$p_S(\theta) = \prod_k p_S(\theta_k), \quad (32)$$

with each distribution given by

$$p_S(\theta_k) = \frac{1}{Z_S} \sigma(\lambda(N_{\text{syn}} - N(\theta_k))), \quad (33)$$

where  $Z_S$  is a normalizing constant. We can interpret this distribution as follows. At each branch, the synapse count  $N(\theta_k)$  on the branch is compared to the maximum number of synapses  $N_{\text{syn}}$ . All configurations where the actual number of synapses is below  $N_{\text{syn}}$  have approximately the same probability, while the prior goes to zero quickly for counts above that number. The parameter  $\lambda > 0$  scales the steepness of the sigmoid and therefore how strictly this bound is enforced. In other words, using this structural prior, the stochastic plasticity dynamics tend to sample network configurations where the number of synapses per branch is at most  $N_{\text{syn}}$ .

We now show that the structural plasticity term  $f_{ki}^S$  follows from the structural prior given in Equations (32), (33). According to Equation (4), we have,

$$\begin{aligned} f_{ki}^S(t) &= \left. \frac{\partial}{\partial \theta_{ki}} \log p_S(\theta) \right|_{\theta(t)} \\ &= \left. \frac{\partial}{\partial \theta_{ki}} \sum_l \log \left( \frac{1}{Z_S} \sigma(\lambda(N_{\text{syn}} - N(\theta_l))) \right) \right|_{\theta(t)} \\ &= \left. \frac{\partial}{\partial \theta_{ki}} \log \left( \frac{1}{Z_S} \sigma(\lambda(N_{\text{syn}} - N(\theta_k))) \right) \right|_{\theta(t)} \\ &= \left. \frac{\partial}{\partial \theta_{ki}} \log(\sigma(\lambda(N_{\text{syn}} - N(\theta_k)))) \right|_{\theta(t)} \end{aligned}$$

$$\begin{aligned} &= -(1 - \sigma(\lambda(N_{\text{syn}} - N(\theta_k)))) \left. \frac{\partial}{\partial \theta_{ki}} N(\theta_k) \right|_{\theta(t)} \\ &= -2\lambda c_w c_\theta [1 - \sigma(\lambda(N_{\text{syn}} - N(\theta_k(t))))] \dot{\sigma}(c_w w_{ki}(t)) H(\theta_{ki}(t)). \end{aligned} \quad (34)$$

#### 4.4.2. Functional Likelihood

We next link the functional plasticity term  $f_{ki}^L$  to the functional likelihood Equations (7), (8). This link is again established using Equation (4). First, we note that

$$\begin{aligned} \frac{\partial}{\partial \theta_{ki}} \log p_{\mathcal{L}}(\theta) &= \left\langle \frac{\hat{c}_{\mathcal{L}}}{t_{\text{max}}} \int_0^{t_{\text{max}}} \frac{\partial}{\partial \theta_{ki}} \Gamma_k(s) \left( \sum_j w_{kj} x_j(s) - \gamma \sum_j w_{kj} (1 - x_j(s)) \right) ds \right\rangle_{X, \Gamma_k}. \end{aligned} \quad (35)$$

Under the simplifying assumption that  $\Gamma_k(t)$  does not depend on  $\theta_{ki}$  (which is approximately satisfied if each input has only a small influence on the branch potential), we obtain

$$\frac{\partial}{\partial \theta_{ki}} \log p_{\mathcal{L}}(\boldsymbol{\theta}) \approx \left\langle \frac{\hat{c}_{\mathcal{L}} c_{\theta}}{t_{\max}} \int_0^{t_{\max}} \Gamma_k(s) (x_i(s) - \gamma(1 - x_i(s))) H(\theta_{ki}) ds \right\rangle_{\chi, \Gamma_k} \quad (36)$$

Assuming slow parameter updates, we approximate this with a self-averaging online update

$$\frac{\partial}{\partial \theta_{ki}} \log p_{\mathcal{L}}(\boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}(t)} \approx c_{\mathcal{L}} \Gamma_k(t) (x_i(t) - \gamma(1 - x_i(t))) H(\theta_{ki}(t)). \quad (37)$$

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher. The Python code that was used to generate the results reported in this article is available at: [https://github.com/IGITUGraz/dendritic\\_rewiring](https://github.com/IGITUGraz/dendritic_rewiring).

## REFERENCES

- Antic, S. D., Zhou, W.-L., Moore, A. R., Short, S. M., and Ikonomu, K. D. (2010). The decade of the dendritic NMDA spike. *J. Neurosci. Res.* 88, 2991–3001. doi: 10.1002/jnr.22444
- Bellec, G., Kappel, D., Maass, W., and Legenstein, R. (2018). “Deep rewiring: training very sparse deep networks,” in *International Conference on Learning Representations (ICLR)* (New Orleans, LA).
- Bono, J., and Clopath, C. (2017). Modeling somatic and dendritic spike mediated plasticity at the single neuron and network level. *Nat. Commun.* 8:706. doi: 10.1038/s41467-017-00740-z
- Bono, J., Wilmes, K. A., and Clopath, C. (2017). Modelling plasticity in dendrites: from single cells to networks. *Curr. Opin. Neurobiol.* 46, 136–141. doi: 10.1016/j.conb.2017.08.013
- Branco, T., Clark, B. A., and Häusser, M. (2010). Dendritic discrimination of temporal input sequences in cortical neurons. *Science* 329, 1671–1675. doi: 10.1126/science.1189664
- Brandalise, F., Carta, S., Helmchen, F., Lisman, J., and Gerber, U. (2016). Dendritic NMDA spikes are necessary for timing-dependent associative LTP in CA3 pyramidal cells. *Nat. Commun.* 7:13480. doi: 10.1038/ncomms13480
- Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* 534:115. doi: 10.1038/nature17955
- Chambers, A. R., and Rumpel, S. (2017). A stable brain from unstable components: emerging concepts, implications for neural computation. *Neuroscience* 357, 172–184. doi: 10.1016/j.neuroscience.2017.06.005
- Chen, X., Leischner, U., Rochefort, N. L., Nelken, I., and Konnerth, A. (2011). Functional mapping of single spines in cortical neurons *in vivo*. *Nature* 475, 501–505. doi: 10.1038/nature10193
- Chklovskii, D. B., Mel, B., and Svoboda, K. (2004). Cortical rewiring and information storage. *Nature* 431:782. doi: 10.1038/nature03012
- Cichon, J., and Gan, W.-B. (2015). Branch-specific dendritic Ca<sup>2+</sup> spikes cause persistent synaptic plasticity. *Nature* 520, 180–185. doi: 10.1038/nature14251
- DeBello, W. M. (2008). Micro-rewiring as a substrate for learning. *Trends Neurosci.* 31, 577–584. doi: 10.1016/j.tins.2008.08.006
- Deger, M., Helias, M., Rotter, S., and Diesmann, M. (2012). Spike-timing dependence of structural plasticity explains cooperative synapse formation in the neocortex. *PLoS Comput. Biol.* 8:e1002689. doi: 10.1371/journal.pcbi.1002689

## AUTHOR CONTRIBUTIONS

RL and TL conceived the research idea, wrote the manuscript, analyzed the simulation data, and prepared the figures. TL wrote the code for the computational models and performed the simulations.

## FUNDING

This project has received funding from the Austrian Science Fund #I3251-N33, and the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement #785907 (HBP SGA2).

## ACKNOWLEDGMENTS

We thank Wolfgang Maass for helpful discussions and ideas.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2020.00057/full#supplementary-material>

- Deger, M., Seeholzer, A., and Gerstner, W. (2016). Multi-contact synapses for stable networks: a spike-timing dependent model of dendritic spine plasticity and turnover. *arXiv preprint arXiv:1609.05730*. doi: 10.1093/cercor/bhx339
- Dudek, S. M., and Bear, M. F. (1995). “Homonynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade,” in *How We Learn; How We Remember: Toward An Understanding Of Brain And Neural Systems: Selected Papers of Leon N Cooper* (Hackensack, NJ: World Scientific), 200–204. doi: 10.1142/9789812795885\_0013
- Dvorkin, R., and Ziv, N. E. (2016). Relative contributions of specific activity histories and spontaneous processes to size remodeling of glutamatergic synapses. *PLoS Biol.* 14:e1002572. doi: 10.1371/journal.pbio.1002572
- Fauth, M., Wörgötter, F., and Tetzlaff, C. (2015). The formation of multi-synaptic connections by the interaction of synaptic and structural plasticity and their functional consequences. *PLoS Comput. Biol.* 11(1):e1004031. doi: 10.1371/journal.pcbi.1004031
- French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends Cogn. Sci.* 3, 128–135. doi: 10.1016/S1364-6613(99)01294-2
- Fu, M., Yu, X., Lu, J., and Zuo, Y. (2012). Repetitive motor learning induces coordinated formation of clustered dendritic spines *in vivo*. *Nature* 483, 92–96. doi: 10.1038/nature10844
- Gallinaro, J. V., and Rotter, S. (2018). Associative properties of structural plasticity based on firing rate homeostasis in recurrent neuronal networks. *Sci. Rep.* 8:3754. doi: 10.1038/s41598-018-22077-3
- Gambino, F., Pagés, S., Kehayas, V., Baptista, D., Tatti, R., Carleton, A., et al. (2014). Sensory-evoked LTP driven by dendritic plateau potentials *in vivo*. *Nature* 515:116. doi: 10.1038/nature13664
- Golding, N. L., Staff, N. P., and Spruston, N. (2002). Dendritic spikes for cooperative long-term potentiation. *Nature* 418:326. doi: 10.1038/nature00854
- Gordon, U., Polsky, A., and Schiller, J. (2006). Plasticity compartments in basal dendrites of neocortical pyramidal neurons. *J. Neurosci.* 26, 12717–12726. doi: 10.1523/JNEUROSCI.3502-06.2006
- Govindarajan, A., Kelleher, R. J., and Tonegawa, S. (2006). A clustered plasticity model of long-term memory engrams. *Nat. Rev. Neurosci.* 7:575. doi: 10.1038/nrn1937
- Hebb, D. O. (1949). *The Organization of Behavior*. New York, NY: Wiley.
- Holthoff, K., Kovalchuk, Y., Yuste, R., and Konnerth, A. (2004). Single-shock LTD by local dendritic spikes in pyramidal neurons of mouse visual cortex. *J. Physiol.* 560, 27–36. doi: 10.1113/jphysiol.2004.072678

- Holtmaat, A. J., Trachtenberg, J. T., Wilbrecht, L., Shepherd, G. M., Zhang, X., Knott, G. W., et al. (2005). Transient and persistent dendritic spines in the neocortex *in vivo*. *Neuron* 45, 279–291. doi: 10.1016/j.neuron.2005.01.003
- Ison, M. J., Quiñones, R., and Fried, I. (2015). Rapid encoding of new memories by individual neurons in the human brain. *Neuron* 87, 220–230. doi: 10.1016/j.neuron.2015.06.016
- Jia, H., Rochefort, N. L., Chen, X., and Konnerth, A. (2010). Dendritic organization of sensory input to cortical neurons *in vivo*. *Nature* 464:1307. doi: 10.1038/nature08947
- Jolivet, R., Rauch, A., Lüscher, H., and Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *J. Comput. Neurosci.* 21, 35–49. doi: 10.1007/s10827-006-7074-5
- Kalisman, N., Silberberg, G., and Markram, H. (2005). The neocortical microcircuit as a tabula rasa. *Proc. Natl. Acad. Sci. U.S.A.* 102, 880–885. doi: 10.1073/pnas.0407088102
- Kampa, B. M., Letzkus, J. J., and Stuart, G. J. (2007). Dendritic mechanisms controlling spike-timing-dependent synaptic plasticity. *Trends Neurosci.* 30, 456–463. doi: 10.1016/j.tins.2007.06.010
- Kappel, D., Habenschuss, S., Legenstein, R., and Maass, W. (2015). Network plasticity as Bayesian inference. *PLoS Comput. Biol.* 11:e1004485. doi: 10.1371/journal.pcbi.1004485
- Kappel, D., Legenstein, R., Habenschuss, S., Hsieh, M., and Maass, W. (2018). A dynamic connectome supports the emergence of stable computational function of neural circuits through reward-based learning. *eNeuro* 5:ENEURO.0301-17.2018. doi: 10.1523/ENEURO.0301-17.2018
- Kasai, H., Fukuda, M., Watanabe, S., Hayashi-Takagi, A., and Noguchi, J. (2010). Structural dynamics of dendritic spines in memory and cognition. *Trends Neurosci.* 33, 121–129. doi: 10.1016/j.tins.2010.01.001
- Kastellakis, G., Cai, D. J., Mednick, S. C., Silva, A. J., and Poirazi, P. (2015). Synaptic clustering within dendrites: an emerging theory of memory formation. *Prog. Neurobiol.* 126, 19–35. doi: 10.1016/j.pneurobio.2014.12.002
- Kastellakis, G., Silva, A. J., and Poirazi, P. (2016). Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. *Cell Rep.* 17, 1491–1504. doi: 10.1016/j.celrep.2016.10.015
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. U.S.A.* 114, 3521–3526. doi: 10.1073/pnas.1611835114
- Kleindienst, T., Winnubst, J., Roth-Alpermann, C., Bonhoeffer, T., and Lohmann, C. (2011). Activity-dependent clustering of functional synaptic inputs on developing hippocampal dendrites. *Neuron* 72, 1012–1024. shows functional clustering. doi: 10.1016/j.neuron.2011.10.015
- Knoblauch, A., Körner, E., Körner, U., and Sommer, F. T. (2014). Structural synaptic plasticity has high memory capacity and can explain graded amnesia, catastrophic forgetting, and the spacing effect. *PLoS ONE* 9:e96485. doi: 10.1371/journal.pone.0096485
- Larkum, M. E., Nevian, T., Sandler, M., Polsky, A., and Schiller, J. (2009). Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science* 325, 756–760. doi: 10.1126/science.1171958
- Legenstein, R., and Maass, W. (2011). Branch-specific plasticity enables self-organization of nonlinear computation in single neurons. *J. Neurosci.* 31, 10787–10802. doi: 10.1523/JNEUROSCI.5684-10.2011
- Lisman, J., and Spruston, N. (2005). Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity. *Nat. Neurosci.* 8:839. doi: 10.1038/nn0705-839
- Loewenstein, Y., Kuras, A., and Rumpel, S. (2011). Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex *in vivo*. *J. Neurosci.* 31, 9481–9488. doi: 10.1523/JNEUROSCI.6130-10.2011
- Losonczy, A., and Magee, J. C. (2006). Integrative properties of radial oblique dendrites in hippocampal CA1 pyramidal neurons. *Neuron* 50, 291–307. doi: 10.1016/j.neuron.2006.03.016
- Major, G., Polsky, A., Denk, W., Schiller, J., and Tank, D. W. (2008). Spatiotemporally graded NMDA spike/plateau potentials in basal dendrites of neocortical pyramidal neurons. *J. Neurophys.* 99, 2584–2601. doi: 10.1152/jn.00011.2008
- McCloskey, M., and Cohen, N. J. (1989). “Catastrophic interference in connectionist networks: the sequential learning problem,” in *Psychology of Learning and Motivation*, Vol. 24, ed G. H. Bower (New York, NY: Elsevier), 109–165. doi: 10.1016/S0079-7421(08)60536-8
- Mel, B. W. (1992). “The clusteron: toward a simple abstraction for a complex neuron,” in *Advances in Neural Information Processing Systems*, eds J. E. Moody, S. J. Hanson, and R. P. Lippmann (San Mateo, CA: Morgan-Kaufmann), 35–42.
- Mel, B. W. (1994). Information processing in dendritic trees. *Neural Comput.* 6, 1031–1085. doi: 10.1162/neco.1994.6.6.1031
- Miner, D., and Triesch, J. (2016). Plasticity-driven self-organization under topological constraints accounts for non-random features of cortical synaptic wiring. *PLoS Comput. Biol.* 12:e1004759. doi: 10.1371/journal.pcbi.1004759
- Minerbi, A., Kahana, R., Goldfeld, L., Kaufman, M., Marom, S., and Ziv, N. E. (2009). Long-term relationships between synaptic tenacity, synaptic remodeling, and network activity. *PLoS Biol.* 7:e1000136. doi: 10.1371/journal.pbio.1000136
- Poirazi, P., and Mel, B. W. (2001). Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* 29, 779–796. doi: 10.1016/S0896-6273(01)00252-5
- Rashid, A. J., Yan, C., Mercaldo, V., Hsiang, H.-L. L., Park, S., Cole, C. J., et al. (2016). Competition between engrams influences fear memory formation and recall. *Science* 353, 383–387. doi: 10.1126/science.aaf0594
- Rhodes, P. A. (2008). Recoding patterns of sensory input: higher-order features and the function of nonlinear dendritic trees. *Neural Comput.* 20, 2000–2036. doi: 10.1162/neco.2008.04-07-511
- Rogerson, T., Cai, D. J., Frank, A., Sano, Y., Shobe, J., Lopez-Aranda, M. F., et al. (2014). Synaptic tagging during memory allocation. *Nat. Rev. Neurosci.* 15:157. doi: 10.1038/nrn3667
- Rumpel, S., and Triesch, J. (2016). The dynamic connectome. *eNeuroforum* 7, 48–53. doi: 10.1007/s13295-016-0026-2
- Silva, A. J., Zhou, Y., Rogerson, T., Shobe, J., and Balaji, J. (2009). Molecular and cellular approaches to memory allocation in neural circuits. *Science* 326, 391–395. doi: 10.1126/science.1174519
- Spies, R., George, R., Cook, M., and Diehl, P. U. (2016). Structural plasticity denoises responses and improves learning speed. *Front. Comput. Neurosci.* 10:93. doi: 10.3389/fncom.2016.00093
- Stettler, D. D., Yamahachi, H., Li, W., Denk, W., and Gilbert, C. D. (2006). Axons and synaptic boutons are highly dynamic in adult visual cortex. *Neuron* 49, 877–887. doi: 10.1016/j.neuron.2006.02.018
- Tonegawa, S., Pignatelli, M., Roy, D. S., and Ryan, T. J. (2015). Memory engram storage and retrieval. *Curr. Opin. Neurobiol.* 35, 101–109. doi: 10.1016/j.conb.2015.07.009
- Welling, M., and Teh, Y. W. (2011). “Bayesian learning via stochastic gradient Langevin dynamics,” in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (Madison, WI), 681–688.
- Wu, X. E., and Mel, B. W. (2009). Capacity-enhancing synaptic learning rules in a medial temporal lobe online learning model. *Neuron* 62, 31–41. doi: 10.1016/j.neuron.2009.02.021
- Yang, G., Pan, F., and Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature* 462, 920–924. doi: 10.1038/nature08577
- Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., and Kasai, H. (2008). Principles of long-term dynamics of dendritic spines. *J. Neurosci.* 28, 13592–13608. doi: 10.1523/JNEUROSCI.0603-08.2008
- Zheng, P., and Triesch, J. (2014). Robust development of synfire chains from multiple plasticity mechanisms. *Front. Comput. Neurosci.* 8:66. doi: 10.3389/fncom.2014.00066
- Ziv, N. E., and Ahissar, E. (2009). Neuroscience: new tricks and old spines. *Nature* 462, 859–861. doi: 10.1038/462859a

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Limbacher and Legenstein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.