Supplementary Information

The interplay of synaptic plasticity and scaling enables self-organized formation and allocation of multiple memory representations

Johannes Maria Auth, Timo Nachstedt, Christian Tetzlaff

September 25, 2018

Section A: Measures used in the network simulation

Average shortest path length

The *average shortest path length* (ASPL, Figure 2 D) is considered here as a measure to analyze the spatial distribution of activation within the memory area. A high ASPL between neurons indicates that these neurons are spatially broadly distributed across the memory area. By contrast, a low ASPL indicates that the neurons are clustered. In particular, as strongly activated neurons are supposed to become part of a memory representation, we focus on the distribution of highly activated neurons. For this, for each trial, we identified the 10% of neurons with the highest activity level (index set P) and calculated the shortest path length (SPL; [**?**]); using the *networkX* package for *Python*) between them and averaged over all those paths (denoted by $\langle\cdot\rangle$):

$$\mathrm{ASPL} = \langle \mathrm{SPL}_{\mathrm{i,j}} \rangle_{i,j \in P, \mathrm{i} \neq \mathrm{j}}. \tag{12}$$

Dynamic equilibria of synaptic weights

The *average outgoing recurrent synaptic weight* (Figure 2 E) is a measure of the interconnection within a neuronal sub-population in the memory area (index set Q). We therefore averaged the synaptic weight over all the connections among neurons within the sup-population:

$$\bar{w}^{\mathrm{rec}} = \langle w^{\mathrm{rec}}_{\mathrm{i,j}} \rangle_{i,j \in Q, \mathrm{i} \neq \mathrm{j}}. \tag{13}$$

The *average incoming feed-forward synaptic weight* (Figure 2 F,G) is the average synaptic weight of connections between a sub-population in the memory area (index set Q) and a specific stimulus pattern in the input area (index set H):

$$\bar{w}^{\mathrm{ff}} = \langle w^{\mathrm{ff}}_{\mathrm{i,k}} \rangle_{i \in Q, k \in H}. \tag{14}$$

Response disparity dependent on stimulus similarity

To analyse the response disparity, stimulus S1 is presented 10 times for 5 sec with 1 sec pause in between to form a single CA. After that, plasticity is shut off and we present variations of stimulus S1 with increasing *stimulus disparity* until the stimulus equals stimulus S2 (Figure 2 H). Stimulus disparity measures the relative amount of non-overlap between two stimulus patterns, in this case stimulus S1 and its variation (in the following called stimulus S1'). Both stimuli are of identical size $N^{\mathrm{S}} = 0.5 \cdot N^{\mathrm{I}}$, so that the stimulus disparity is calculated as follows:

$$\textit{stimulus disparity}\,(S1,\, S1') \,=\, 1 - \frac{1}{N^{\mathrm{S}}} \cdot \sum_{\mathrm{k}}^{N^{\mathrm{I}}} S_{\mathrm{k}}(S1) \cdot S_{\mathrm{k}}(S1'), \tag{15}$$

with binary stimulus patterns for a given stimulus $X \in S1, S1'$:

$$S_{\mathrm{k}}(X) = \begin{cases} 1, & \text{if } I_{\mathrm{k}}(X) \,=\, 130, \\ 0, & \text{if } I_{\mathrm{k}}(X) \,=\, 0. \end{cases} \tag{16}$$

Thus, a stimulus disparity equal zero describes two identical stimuli, whereas a disparity equal one indicates two non-overlapping stimulus patterns. The input area size of $N^{\mathrm{I}} = 36$ allows for 18 steps in

variation of $5.\bar{5}\%$ each. At the end of each presentation, we compare the resulting response in the memory area with the one at the end of the learning phase (i.e. the response to the original stimulus S1). The *response vector overlap* (RVO; Figure 2 H) describes the similarity between the response patterns in the memory area due to the presentation of stimuli S1 and S1':

$$RVO(S1,\, S1') \,=\, \sum_{\mathrm{i}}^{\mathrm{N^M}} R_{\mathrm{i}}(S1) \cdot R_{\mathrm{i}}(S1') \tag{17}$$

with binary response of neuron i to a given stimulus $X \in S1, S1'$:

$$R_{\mathrm{i}}(X) = \begin{cases} 1, & \text{if } F_{\mathrm{i}}(X) \geqslant 0.5 \cdot \alpha, \\ 0, & \text{else.} \end{cases} \tag{18}$$

Explicit network simulation results

S 1. Explicit results depicted in Figure 2 A,B: average shortest path length (ASPL) and average synaptic weights.

	test0	**test1**	**test2**
$\bar{w}^{\mathrm{ff}}_{11}$	107 (36)	290 (11)	302 (11)
$\bar{w}^{\mathrm{ff}}_{12}$	96 (22)	145 (19)	25 (15)
$\bar{w}^{\mathrm{ff}}_{21}$	95 (21)	3.03 (0.12)	14.95 (0.13)
$\bar{w}^{\mathrm{ff}}_{22}$	108 (36)	103 (36)	305.004 (0.002)
$\bar{w}^{\mathrm{rec}}_{1}$	19.37 (0)	69 (10)	81 (22)
$\bar{w}^{\mathrm{rec}}_{2}$	19.37 (0)	16.6 (5.2)	63 (20)
$\bar{w}^{\mathrm{rec}}_{\mathrm{RR}}$	19.37 (0)	24.7 (6.2)	30 (11)
ASPL S1	3.592 (0.018)	1.809 (0.014)	1.808 (0.016)
ASPL S2	3.592 (0.017)	3.546 (0.036)	1.837 (0.035)

Section B: Nullclines and Equilibria

The equilibrium values $\bar{w}_i^{\mathrm{ff},*}$ and $\bar{w}_i^{\mathrm{rec},*}$ of the feed-forward and recurrent weights can be obtained as a function of the equilibrium values of the population activities $\bar{F}_i^*$ from Equation 10 and Equation 11:

$$\bar{w}_{ik}^{\mathrm{ff},*}(\bar{F}_i^*) = \sqrt{\frac{\kappa^{\mathrm{ff}} \bar{F}_i^* \bar{I}_k}{\bar{F}_i^* - F^{\mathrm{T}}}}. \tag{19}$$

$$\bar{w}_i^{\mathrm{rec},*}(\bar{F}_i^*) = \sqrt{\frac{\kappa^{\mathrm{rec}} (\bar{F}_i^*)^2}{\bar{F}_i^* - F^{\mathrm{T}}}}, \tag{20}$$

The equilibrium value $\bar{u}^*_{\mathrm{inh}}$ of the membrane potential of the inhibitory population can be formulated as a function of $\bar{F}_1^*$ and $\bar{F}_2^*$ based on Equation 8:

$$\bar{u}^*_{\mathrm{inh}}(\bar{F}_1^*, \bar{F}_2^*) = R_{\mathrm{inh}} \tau_{\mathrm{inh}} (w_{\mathrm{inh},1} N_1 \bar{F}_1^* + w_{\mathrm{inh},2} N_2 \bar{F}_2^*). \tag{21}$$

By inserting Equations 20, 19 and 21 into Equation 7 and using $\bar{F}_i^* = F(\bar{u}_i^*)$, we obtain a system of the two population nullclines that only depends on the equilibrium values $\bar{u}_1^*$ and $\bar{u}_2^*$ ($i \in \{1, 2\}$):

$$0 = -\frac{\bar{u}_i^*}{\tau} + R\left(\bar{n}_i^{\mathrm{rec}}\bar{w}_i^{\mathrm{rec},*}(\bar{F}_i^*)\bar{F}_i^* + w_{i,\mathrm{inh}}\bar{F}_{\mathrm{inh}}^*(\bar{F}_1^*, \bar{F}_2^*) + \sum_k \bar{n}^{\mathrm{ff}}\bar{w}_{ik}^{\mathrm{ff},*}(\bar{F}_i^*)\bar{I}_k\right).$$

We solve this system numerically to receive the equilibrium values $\bar{u}_1^*$ and $\bar{u}_2^*$ and, in consequence, by means of equations 20, 19 and 21, also $\bar{w}_1^{\mathrm{rec},*}$, $\bar{w}_2^{\mathrm{rec},*}$, $\bar{w}_{1\mathrm{A}}^{\mathrm{ff},*}$, $\bar{w}_{1\mathrm{B}}^{\mathrm{ff},*}$, $w_{2\mathrm{A}}^{\mathrm{ff},*}$, $\bar{w}_{2\mathrm{B}}^{\mathrm{ff},*}$ and $\bar{u}_{\mathrm{inh}}^*$.

Section C: Stability

The stability of an equilibrium is determined by the sign of the eigenvalue with the largest real part of the system's Jacobi matrix evaluated at the equilibrium. The nonzero terms of the Jacobi matrix are ($i \in \{1, 2\}$, $k \in \{\mathrm{S1}, \mathrm{S2}\}$):

$$\begin{aligned}
\frac{\partial \dot{\bar{u}}_i}{\partial \bar{u}_i} &= -\frac{1}{\tau} + R\bar{n}_i^{\mathrm{rec}}\bar{w}_i^{\mathrm{lat}}\frac{\partial \bar{F}_i}{\partial \bar{u}_i}, & \frac{\partial \dot{\bar{u}}_i}{\partial \bar{u}_{\mathrm{inh}}} &= Rw_{i,\mathrm{inh}}\frac{\partial \bar{F}_{\mathrm{inh}}}{\partial \bar{u}_{\mathrm{inh}}},\\
\frac{\partial \dot{\bar{u}}_i}{\partial \bar{w}_i^{\mathrm{rec}}} &= R\bar{n}_i^{\mathrm{rec}}\bar{F}_i, & \frac{\partial \dot{\bar{u}}_i}{\partial \bar{w}_{ik}^{\mathrm{ff}}} &= R\bar{n}_i^{\mathrm{ff}}\bar{I}_k,\\
\frac{\partial \dot{\bar{u}}_{\mathrm{inh}}}{\partial \bar{u}_i} &= R_{\mathrm{inh}}w_{\mathrm{inh},i}N_i\frac{\partial \bar{F}_i}{\partial \bar{u}_i}, & \frac{\partial \dot{\bar{u}}_{\mathrm{inh}}}{\partial \bar{u}_{\mathrm{inh}}} &= -\frac{1}{\tau_{\mathrm{inh}}},\\
\frac{\partial \dot{\bar{w}}_i^{\mathrm{rec}}}{\partial \bar{u}_i} &= \mu^{\mathrm{rec}}\frac{\partial \bar{F}_i}{\partial \bar{u}_i}\left(2\bar{F}_i - \frac{(\bar{w}_i^{\mathrm{rec}})^2}{\kappa^{\mathrm{rec}}}\right), & \frac{\partial \dot{\bar{w}}_i^{\mathrm{rec}}}{\partial \bar{w}_i^{\mathrm{rec}}} &= \frac{2\mu^{\mathrm{rec}}}{\kappa^{\mathrm{rec}}}(F^{\mathrm{T}} - \bar{F}_i)\bar{w}_i^{\mathrm{rec}},\\
\frac{\partial \dot{\bar{w}}_{ik}^{\mathrm{ff}}}{\partial \bar{u}_i} &= \mu^{\mathrm{ff}}\frac{\partial \bar{F}_i}{\partial \bar{u}_i}\left(\bar{I}_k - \frac{(\bar{w}_{ik}^{\mathrm{ff}})^2}{\kappa^{\mathrm{ff}}}\right), & \frac{\partial \dot{\bar{w}}_{ik}^{\mathrm{ff}}}{\partial \bar{w}_{ik}^{\mathrm{ff}}} &= \frac{2\mu^{\mathrm{ff}}}{\kappa^{\mathrm{ff}}}(F^{\mathrm{T}} - \bar{F}_i)\bar{w}_{ik}^{\mathrm{ff}}
\end{aligned}$$

with

$$\frac{\partial \bar{F}_i}{\partial \bar{u}_i} = \beta\bar{F}_i\left(1 - \frac{\bar{F}_i}{\alpha}\right) \qquad \text{and} \qquad \frac{\partial \bar{F}_{\mathrm{inh}}}{\partial \bar{u}_{\mathrm{inh}}} = \beta\bar{F}_{\mathrm{inh}}\left(1 - \frac{\bar{F}_{\mathrm{inh}}}{\alpha}\right).$$

The eigenvalues of the resulting matrix are determined numerically.

Section D: Feed-Forward Synaptic Weight Change

For constant pre- and post-synaptic activities (Figure 5 D), Equation 10 can be solved analytically by separation of variables. The resulting time-course $w_i^{\mathrm{ff}}(t)$ depends on the given parameters and initial conditions:

$$\bar{w}_i^{\mathrm{ff}}(t) = \begin{cases}
\bar{w}_i^{\mathrm{ff},*}\coth\left(\sqrt{\frac{\bar{F}_i\bar{I}_i(\bar{F}_i - F_{\mathrm{T}})}{\kappa^{\mathrm{ff}}}}(t - t_0)\mu^{\mathrm{ff}} + \operatorname{arcoth}\left(\frac{\bar{w}_i^{\mathrm{ff}}(t_0)}{w_i^{\mathrm{ff},*}}\right)\right) & \text{for } \bar{w}_i^{\mathrm{ff}}(t_0) > w_i^{\mathrm{ff},*} \wedge \bar{F}_i^{\mathrm{ff}} > F_{\mathrm{T}} \wedge \bar{I}_i > 0,\\
\bar{w}_i^{\mathrm{ff},*}\tanh\left(\sqrt{\frac{\bar{F}_i\bar{I}_i(\bar{F}_i - F_{\mathrm{T}})}{\kappa^{\mathrm{ff}}}}(t - t_0)\mu^{\mathrm{ff}} + \operatorname{artanh}\left(\frac{\bar{w}_i^{\mathrm{ff}}(t_0)}{w_i^{\mathrm{ff},*}}\right)\right) & \text{for } \bar{w}_i^{\mathrm{ff}}(t_0) < w_i^{\mathrm{ff},*} \wedge \bar{F}_i^{\mathrm{ff}} > F_{\mathrm{T}} \wedge \bar{I}_i > 0,\\
\bar{w}_i^{\mathrm{ff},\dagger}\tan\left(\sqrt{\frac{\bar{F}_i\bar{I}_i(\bar{F}_i - F_{\mathrm{T}})}{\kappa^{\mathrm{ff}}}}(t - t_0)\mu^{\mathrm{ff}} + \arctan\left(\frac{\bar{w}_i^{\mathrm{ff}}(t_0)}{w_i^{\mathrm{ff},*}}\right)\right) & \text{for } \bar{w}_i^{\mathrm{ff}}(t_0) < w_i^{\mathrm{ff},*} \wedge \bar{F}_i^{\mathrm{ff}} < F_{\mathrm{T}} \wedge \bar{I}_i > 0,\\
\left(\frac{1}{\bar{w}_i^{\mathrm{ff}}(t_0)} - \frac{F^{\mathrm{T}} - \bar{F}_i^{\mathrm{ff}}}{\kappa^{\mathrm{ff}}}(t - t_o)\mu^{\mathrm{ff}}\right)^{-1} & \text{for } \bar{F}_i^{\mathrm{ff}} = 0 \vee \bar{I}_i = 0
\end{cases}$$

with

$$\bar{w}_i^{\mathrm{ff},\dagger}(\bar{F}_i^*) = \sqrt{\frac{\kappa^{\mathrm{ff}} \bar{F}_i^* \bar{I}}{F^{\mathrm{T}} - \bar{F}_i^*}}.$$

Section E: Comparison of Bifurcation Curve with Network Simulation

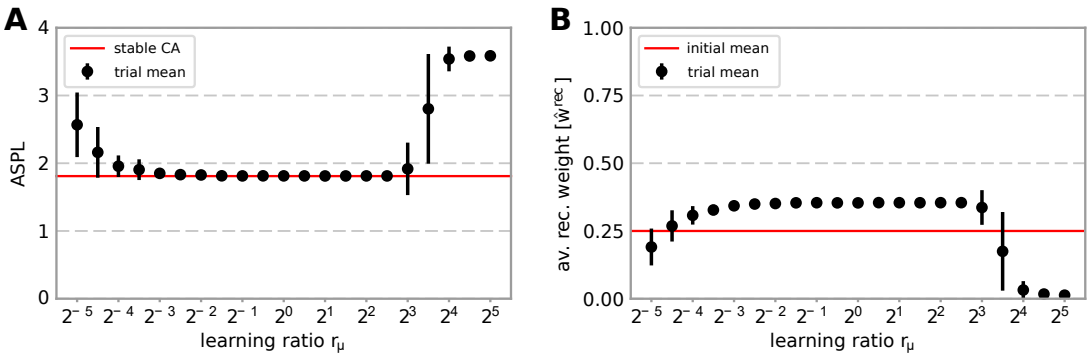
When comparing the equilibrium structure of the population model dependent on the input amplitude (bifurcation parameter) with the equilibria reached in network simulations (Figure 5 C), the network simulations are initialized close to the different expected stable configurations. For every input amplitude I, we perform two simulations with different initial conditions:

- $w_{ij}^{\mathrm{rec}} = 0.25\hat{w}^{\mathrm{rec}}$ for all realized recurrent synapses and $w_{ij}^{\mathrm{ff}} = \hat{w}^{\mathrm{ff}}$ for all realized feed-forward synapses.
- $w_{ij}^{\mathrm{rec}} = \hat{w}^{\mathrm{rec}}$ for synapses in between 121 neurons in a circle-shaped population, $w_{ij}^{\mathrm{rec}} = 0.25\hat{w}^{\mathrm{rec}}$ for all other realized recurrent synapses and $w_{ij}^{\mathrm{ff}} = \hat{w}^{\mathrm{ff}}$ for all realized feed-forward synapses.

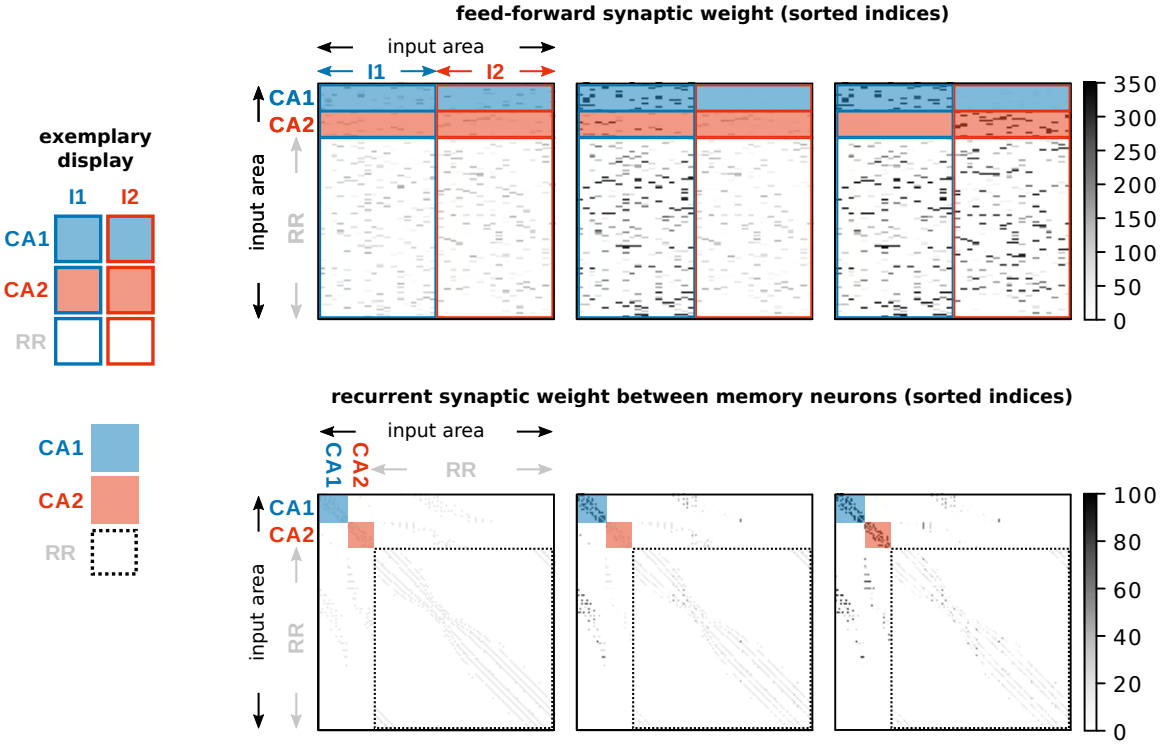
In each case, the network is simulated for $50,000\,\mathrm{s}$. Every simulation is repeated 50 times with different random connectivities. To avoid simulation artefacts related to absolute silence of input channels, we assume a small background activity of 0.1α for inactive inputs. In the final state, we either consider all neurons with activity higher than 0.5α or, if there are none, 120 neurons centred around the activity centre of the network as population 1. Population 2 is defined as the circular group of 120 neurons with the highest distance (respecting the periodic boundary conditions) to population 1. Within these two population, we evaluate the mean recurrent weight.

Large Input Amplitudes In the network simulation, the functional role of the inhibitory population is two-fold: On the one hand, inhibition mediates the competition between different populations. This role is also captured by the population model. On the other hand, it prevents an active cell assembly from growing without limit by inhibiting neighbouring neurons. This aspect is not reflected in the population model as in the latter the size of the populations is approximated as being fixed. Due to this discrepancy, the population model predicts equilibria also for very large input amplitudes while in the network simulation these input amplitudes lead to full activation of the complete network.

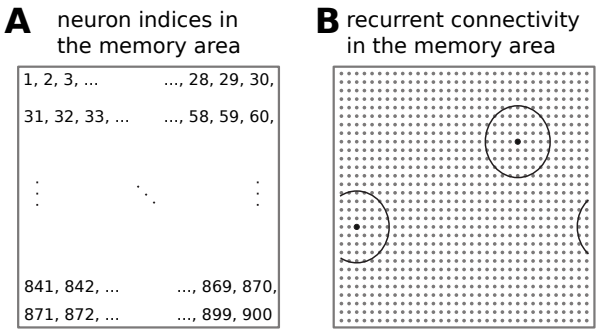
Section F: Supplementary Figures



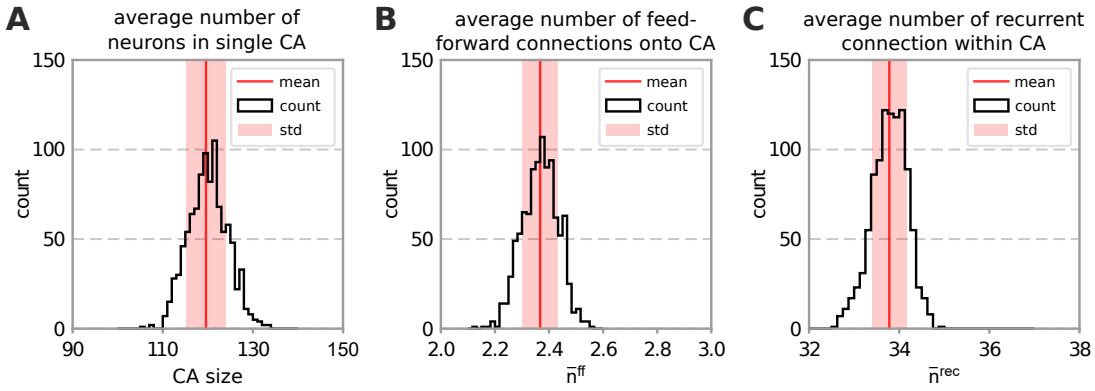
S 1. The robustness of CA formation according to parameter variations. The formation of a cell assembly is robust against changes in the velocity of synaptic adaptations of the feed-forward (μ^{ff}; Equation 5) and of the recurrent synapses (μ^{rec}; Equation 6). Given one learning phase, (A) ASPL (between the 10 mostly active neurons) as well as (B) average recurrent synaptic weights within the CA indicate that a CA is formed for a wide range of time scale ratios ($r_\mu = \frac{\mu^{\text{rec}}}{\mu^{\text{ff}}}$). Please note that a low ASPL (≈ 1.8; A, red line) and a significant increase of the average recurrent synaptic weight above the initial mean value (0.25; B, red line) indicate a proper formation of a CA. In the main text, synaptic changes of both types of synapses (feed-forward as well as recurrent) occur with the same time scale $\mu^{\text{ff}} = \mu^{\text{rec}} = \mu$.



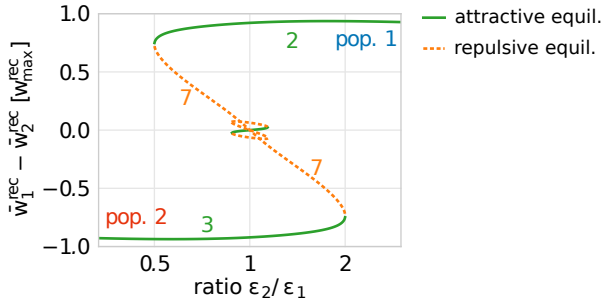
S 2. All synaptic weights of feed-forward and recurrent connections. Raw data of the weights of (top) feed-forward and (bottom) recurrent synapses (left) before learning, (middle) after the first learning phase, and (right) after the second learning phase for one system initiation. Indices of neurons in the memory area are sorted according to formed CAs (blue and red shading) while indices of neurons in the input area are sorted according to their input-affiliation (blue and red boxes).



S 3. Indexing of neurons in sub-plots and topology. Indexing of neurons and recurrent connectivity in the memory area. (A): The neurons are arranged on a 30x30-grid with indices running from left-top to right-bottom such that each dot in sub-plots of Figure 3 and Figure 4 indicate properties of one neuron. (B): Each neuron is connected to its neighbours, if the position of the neighbouring neuron is within a circle of radius 4 (measured in neuronal units; see two circles as examples). Periodic boundary conditions are introduced to avoid boundary and finite-size effects.



S 4. Parameter estimation for population model. Some parameters of the population model are estimated from the full network model. Each panel shows the histogram of a parameter from 1000 different network initializations. Average values are given as *mean*±*standard deviation*. (A): numbers of neurons within the first formed CA, with average $\bar{N}^{\mathrm{CA}} = 120 \pm 4$. (B): average number of feed-forward connections per CA-neuron from the active input population I1 to corresponding CA, with average $\bar{n}^{\mathrm{ff}} = 2.37 \pm 0.07$. (C): average number of recurrent connections each neuron within the first formed CA receives from other neurons in the same CA, with average $\bar{n}^{\mathrm{rec}} = 33.8 \pm 0.4$).



S 5. Excitability bifurcation diagram as obtained from the population model. Excitability bifurcation diagram as obtained from the population model. Note that a higher value of ϵ_i means a lower excitability of population i and vice-versa. For a given ratio $\frac{\epsilon_2}{\epsilon_1}$, the used values are $\epsilon_1 = \epsilon_0 \cdot \sqrt{\frac{\epsilon_2}{\epsilon_1}}^{-1}$ and $\epsilon_2 = \epsilon_0 \cdot \sqrt{\frac{\epsilon_2}{\epsilon_1}}$.