Supporting Information: Spike time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity

Ila R. Fiete, Walter Senn, Claude Wang, and Richard H.R. Hahnloser

Contents

List of Figures

Calculation: Distribution of cycle lengths in the set of all permutation matrices.

An $N \times N$ permutation matrix applied to the vector $(1, 2, ..., N)$ performs a reordering of the elements $(1, 2, ..., N)$. The set of all permutation matrices of size $N \times N$ is equivalent to the set of all possible random rearrangements of $(1, 2, ..., N)$. If $N = 4$, the rearrangement (2341) contains a single cycle of length 4, while the rearrangement (21)(43) contains two cycles of length 2 each. We want to compute the probability of finding a cycle of length L in the full set of permutation matrices of size $N \times N$. The number of possible ways to select elements to make a cycle of length L out of N elements is

$$N \cdot (N-1) \cdots (N-L+1)/L \tag{1}$$

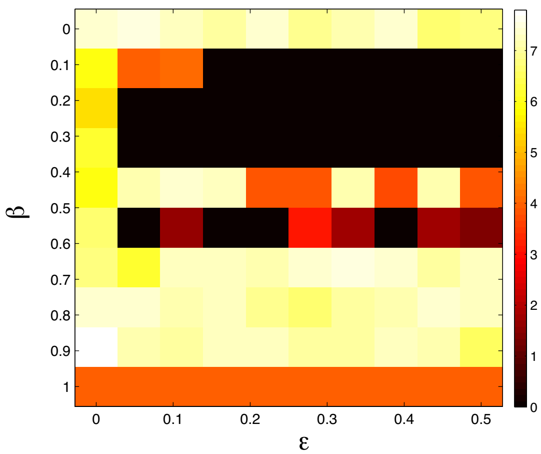
where there is a division by L because the cycle is unchanged under L cyclic rotations of the elements (the cycle $(2341) = (3412) = (4123) = (1234)$). In addition, there are $(N-L)!$ ways to arrange the remaining $N-L$ elements in the permutation. (Note that the rearrangements of the remaining elements may include the formation of additional chains of length L, if $L \leq N/2$: this possible multiplicity of shorter chains is included in our counting.) The total number of ways to arrange the set of N numbers, regardless of cycle length, is $N!$. Therefore, the probability of getting a chain of length L is proportional to

$$\frac{[N \cdot (N-1) \cdots (N-L+1)] \cdot (N-L)!}{N!L} = \frac{1}{L}$$

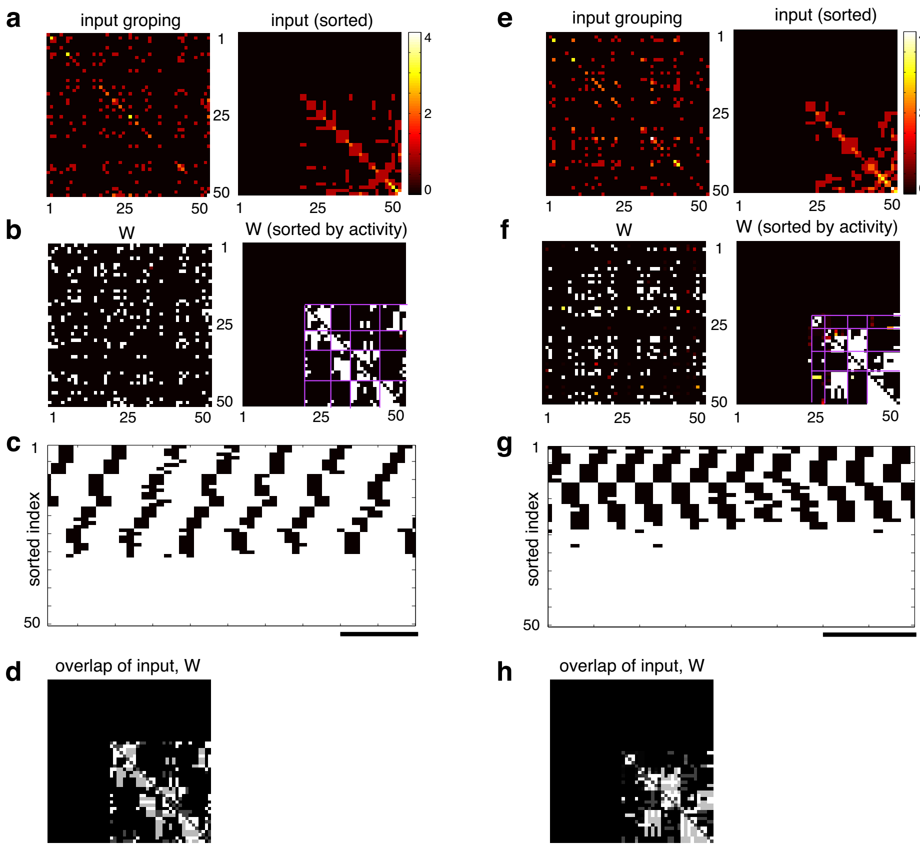
The probability of an L-cycle is $p(L) = c/L$, for $L \leq L_{max}$, with normalization $c^{-1} = \sum_{L=1}^{L_{max}} 1/L \approx \log(L_{max}) + \gamma_c + \mathcal{O}(1/L_{max})$, where $\gamma_c \approx .577$ is the Euler-Mascheroni constant. $L_{max} = N$. (In the case of m width chains, $L_{max} = N/m$.)

Matlab code: Formation of chains of width $m = 1$ in a simple binary neuron network.

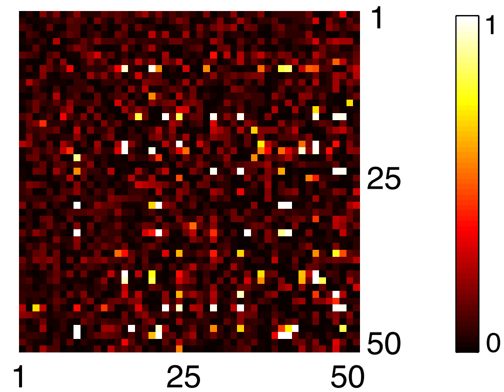
This code can be downloaded from the Neuron website at http://www.cell.com/neuron/. It can be run in Matlab without external inputs or additional files after saving it as a file with extension '.m'.

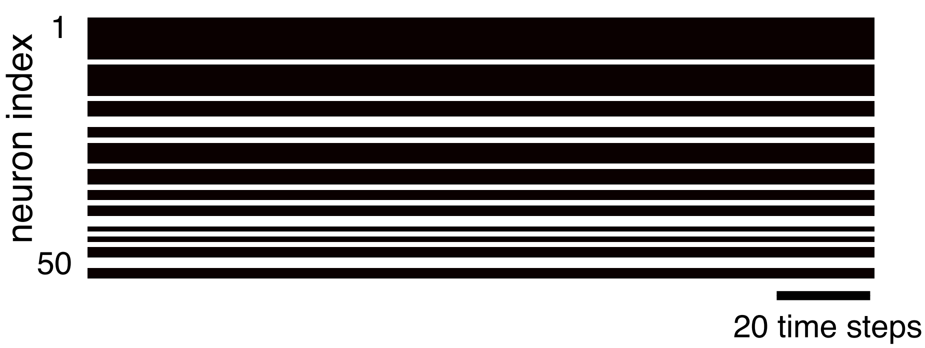


Supporting Figure 1: Related to Figure 2. **Insensitivity of chain formation to parameters.** The figure illustrates parameter regimes where the discrete-time binary neuron network forms synaptic chains from arbitrary random but weak initial weights. Two parameters are varied: the strength of global inhibition (β), and the strength of heterosynaptic LTD (ϵ), which is triggered when the summed-weight bound is hit. Plotted is the log of the number of entries in the matrix $\frac{1}{\max WW^T} WW^T$ that deviate beyond a threshold amount from the identity matrix. If the formed weight matrix W is a permutation matrix, then WW^T is the identity matrix, and the deviation measure used here will give the log of 0 (-∞). We have set the color scale so that black represents log(0). It is clear that the network forms into permutation matrices for a range of ϵ and β. In fact, beyond the necessity of some heterosynaptic LTD to promote competition, the strength of heterosynaptic LTDP does not require much tuning. The strength of global inhibition can vary 3-fold. All other network details and parameters are identical to those used for Figure 2 in the main paper.

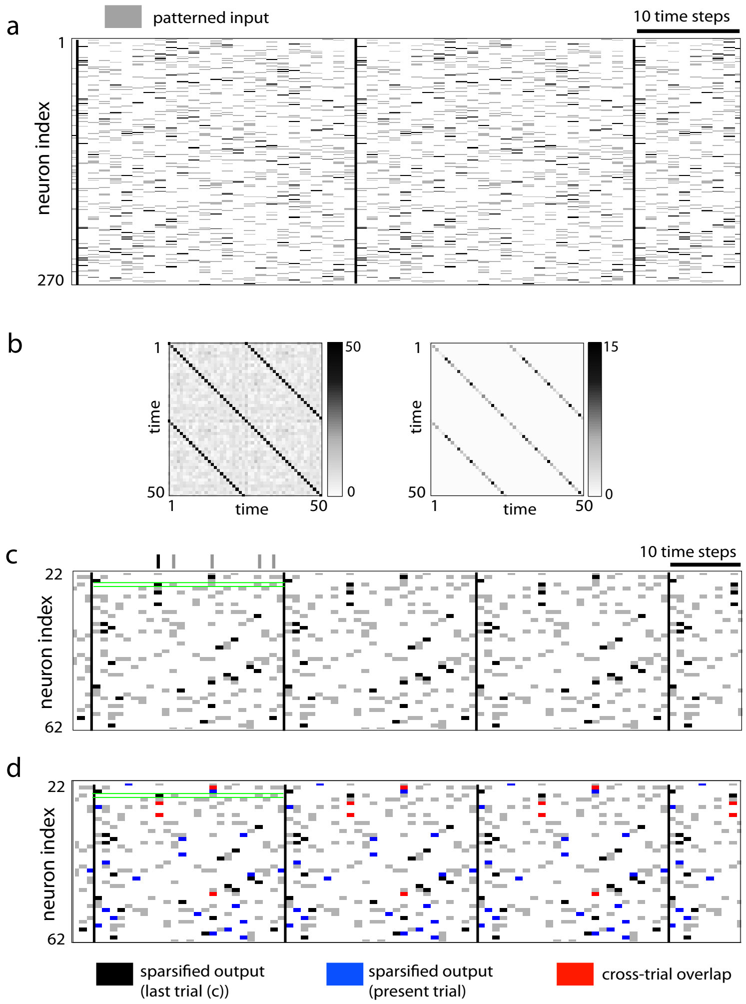


Supporting Figure 2: Related to Figure 5. **Wide chain formation with random overlapping input groups.** A, E) *Left:* 10 input groups of 5 randomly selected neurons each in a network of $N = 50$ neurons. If neurons i, j are never in the same group, the ij entry is 0. If the i, j neurons both belong to a single group, the ij entry is 1. If they both belong to the same two groups, the entry is 2. Different groups sometimes contain the same neuron(s) (entries larger than 1). *Right:* The neuron indices are re-sorted according to similarity (neurons with similar group membership or row pattern are put near each other). If the groups were disjoint in their membership, this procedure would produce non-overlapping blocks along the diagonal. Note that many neurons are never driven by the input and remain unused. B, F) *Left:* Weight matrices resulting from application of the summed-weight limit learning rule with temporally random input drive. *Right:* The same weight matrices, sorted according to the order of neural activation during playback of the formed activity sequences (shown in C, G). The underlying structure of the weights is suddenly apparent. The matrices have a roughly block permutation structure. Diagonal blocks represent output groups, or groups of neurons that are recurrently connected. Off-diagonal blocks represent the chain-like connectivity between groups. In (B), there are 4 clear blocks; the first connects to the second, which connects to the third, which connects to the fourth, which connects back to the first. In (F), there is one isolated block (first block) and 3 blocks in a chain: The second connects to the fourth, the fourth to the third, and the third to the second. The formed groups are largely disjoint even though the inputs did not stimulate disjoint blocks. These matrices, resulting from completely random input groupings, differ from the results obtained with disjoint input groups (Figure 5) in a few ways: the blocks are of varying sizes, and have a messier morphology. C, G) The activity in the network during playback of the learned sequences, after neurons are sorted to be near others with similar activity onset times. Scale bar = 10 time-steps. It is clear that the repeatable sequential activity patterns in the networks are unary. (Occasionally, neurons fail to respond or respond more than once in a sequence, due to the stochastic response of the system, but the repeatable response is unary.) Sequence length is ≈ 10 time-steps in C) and ≈ 4 in G). D, H) Groups in the learned weight matrix are different from input groups, but there is an elevated probability that neurons within a group receive the same external inputs. Plotted is the weight matrix summed with the input group matrix, both sorted in the same way. White areas indicate overlap of the input and formed groups; dark and light gray represent only input groups or formed groups, respectively. The difference between (A–D) and (E–H) is the value of m (defined as $W_{\max} = m w_{\max}$). In (A–D), $m = 7$. In (E–H), $m = 9$. All else is fixed. These simulations include probabilistic neuron and synaptic responses (see Methods).





Supporting Figure 3: Related to Figures 2 and 6. **STDP is not sufficient for synaptic chain formation.** Depicted here are results from the same simulation as in Figure 1 of the main manuscript (same inputs, architecture, learning rule, and all parameters), except that ϵ, the strength of heterosynaptic LTD, is set to 0. This means that learning is based only on STDP, with hard bounds of $w_{\max}$ on each synapse, but no bounds on the summed weights into or out of each neuron. A) The resultant weight matrix starting from the same initial conditions. Most rows and columns have no large entries; but the rows and columns associated with a few neuron indices each contain several non-zero entries. These neurons have become activity hubs, receiving many inputs and sending many outputs. B) The activity pattern in the network consists of continuous activation through time of some neurons (the hubs from (A)), and silence in the rest of the neurons. The total level of activation in the network and the density of large entries in the matrix depends on the level of global inhibition in the network. The network with STDP alone fails to assemble into a synaptic chain architecture, and does not support sequential activity.



Supporting Figure 4: Related to Figure 7. **The weight-growth limit rule with sequential input can produce longer sequences than determined by τ_{ada}.** Vertical black lines highlight the period of the activity loop (A, ,C, D). A) Patterned (sequential) input (with added noise) provided to the network (grey bars), and the self-organized sequence produced by the network in response (overlaid black bars). B) Temporal cross-correlations of the input sequence (left) and the network response (right). The input sequence is dense with large cross-correlations, while the network response is perfectly unary: multiple neurons are active at one time-step but each neuron is active exactly once. C) Magnified version of the inputs to and responses of a few neurons from (A). The set of active neurons at each time in the unary sequence (black) is a subset of the neurons activated at that time by the dense input sequence (grey). Black and grey bars over the boxed neural response: the multiple times that neuron 25 is driven to be active by the input sequence. Black: the self-generating sequence of activity in the network. Neuron 25 is active only once in the sequence, at one of the times it was driven by its training inputs. D) The network is trained again from scratch in a new learning trial (new random seed for noise generation) on the same input sequence (grey bars). As before, the network response (blue) fully overlaps the input patterns, and is unary. But the unary sequence on this trial is distinct from the preceding trail (C, black). The few cross-trial overlaps are shown in red. Numbers: The average number of neurons active per time is 25. $N = 270$, $\alpha = 5$; $\beta = 3.33$; $W_{\max} = 1.25$; $pn = 30$; $\eta = .05$; $\epsilon = 0.21$; $\tau = 3$; annealing began at 40 time-steps, and $\tau_{anneal} = 10$.