Supporting Text

1. Parameter Dependence of Cluster Model.

The cluster model shows persistent metastability only in a certain parameter regime. Outside this parameter regime, clusters might either rapidly decay, or rapidly expand to take over the entire lattice. We explored the appropriate parameter regime for metastability; results are shown in Fig. 7. In order to create this figure, we defined as metastable simulations that did not shift up or down to a new metastable state decay within the first 200 time units. We varied two parameters: the repulsion constant (L_1) and the slope of the insertion probability function (β). These results are based on simulations with initial cluster size of 8x8; results would be qualitatively different for different initial sizes. We find that at higher β (low noise) the metastable regime has a wider range of allowed values of L_1.

2. The off-rate model: Assumptions and consequences

In the cluster model proposed in this paper, the on rate (insertion rate) of receptors depends on the number of neighbors a receptor has while the removal rate is constant, producing an asymmetry between the insertion and removal processes. Another option is to have a constant on-rate and a neighbor-dependent off rate. This might seem more intuitive to some, if we perceive the cluster as an aggregate. We claim that although such a model might extend the lifetime of clusters, it does so simply by extending the average dwell-time of receptors in the cluster and therefore does not address the question of how synaptic efficacies can be much more stable than the proteins they are made of.

To demonstrate this, we define mathematically an alternative off-rate model and simulate it. We assume a fixed insertion probability of the form:

$$P^{ex}_{orm}(i,j) = (1 - S_{ij})\mu_{on}\Delta t,$$

where μ_{on} is the on-rate parameter, and the orm subscript denotes the off-rate model. The removal probability was assumed to have the following neighbor dependent form:

$$P^{in}_{orm}(i,j) = \frac{2\mu_{off}\exp(-\beta h_1(i,j))}{1+\exp(-\beta h_1(i,j))} S_{i,j} \cdot \Delta t,$$

where μ_{off} is the off-rate parameter, h_1 is the field as defined in Eq. **1** of the main text, and β is the slope of this monotonically decreasing function. For this monotonically decreasing

function, as the number of neighbors increases, the off rate decreases. Note that the field h_1 depends of the parameter L_1. In all our simulations we used $L_1 = 1.5$, $\mu_{off} = 1$. The parameters μ_{on} and β are varied from simulation to simulation. A large μ_{on} induces strong background fluctuations in the number of receptors but does not contribute to preserving the initial cluster; we typically chose $\mu_{on} << \mu_{off}$. Movie 7 shows an example of sn off-rate model with $\beta = 6$, $\mu_{on} = 0.01$ and an initial size of 7x7 receptors. If μ_{on} is too large, clusters would expand rather then decay.

From simulations of the off-rate model, we can see that the clusters can have a lifetime larger than 1, which is the dwell time of an isolated receptor. As we increase β the lifetime of the clusters will increase. However, a receptor with neighbors has a longer dwell time than an isolated receptor. The average dwell time of a receptor at point i, j on the lattice is: $\Delta t/P^{in}_{orm}(i, j)$ and will therefore change in time as the field changes. However, this can easily be calculated for a fixed number of neighbors. In Fig. 8 we plot the mean lifetime of clusters (time to get to $1/e$ of initial size) as a function of β and compare that to the lifetime of receptors with 2 (red) or 3 (blue) neighbors. Although the lifetime of the clusters is larger than the dwell time of receptors with 2 neighbors (corner receptors), it increased at about the same rate, and at a much lower rate than the dwell time of receptors with 3 neighbors (edge receptors). Therefore, we conclude that the off-rate model does not produce efficacies that have longer lifetimes than the receptors they are made of, it merely increases the dwell times of receptors within the cluster.

3. Formal Similarities and Differences Between the Cluster Model and the Ising Spin Model.

Eqs. **1, 8** and **9** in the main text seem to be nearly identical to equations defining the dynamics in an Ising spin model. In an Ising spin model the spins take the values 1 and -1 rather than 0 and 1 as defined here, but a simple variable transformation can relate the two. However, the model defined here differs from the standard formulation of an Ising spin model, and more generally is not formally analogous to models of magnetic systems in statistical physics (1). As in the Ising model the transition from an unoccupied state ($S_{ij} = 0$) to an occupied state ($S_{ij} = 0$) depends on the field ($h_k(ij)$) and on a parameter β which can be thought of as an inverse temperature. However, the opposite transition from an occupied to an unoccupied state proceeds at a constant rate that is independent of the field, the number of neighbors or β. For such asymmetric transitions, no energy function (or Lyaponov function) has been defined, and detailed balance is not preserved. Moreover, the "inverse temperature " parameters (β)can be changed without changing the removal rate parameter (μ), so the insertion and removal processes can be perceived as having two different "temperatures".

Could the simpler Ising spin model be used for preserving synapse specific efficacies? Below the critical temperature, a ferromagnetic Ising spin model has one well defined order parameter, which in this context could be related to the synaptic efficacy. Therefore, the equilibrium states of an Ising model could not be a good model system for preserving different synaptic efficacies at each synapse. Different synaptic efficacies could be preserved by setting individually the parameter β at each synapse. What this means mechanistically is that different on and off rates must be preserved individually at each synapse, for as long as the synaptic efficacy is to be preserved. Such a need to precisely control the molecular dynamic parameters at each synapse does not seem biologically plausible, and is exactly what this model tries to avoid. Very close to the critical 'temperature' $1/\beta_c$ many different values of the order parameter can be preserved for a long period of time. However, this would require fine tuning of β and constantly keeping the synapse at the critical parameter. Although I could not find a parameter regime in which the standard Ising spin model generates long lasting metastable sates, I have not proven here that this cannot be done.

Fig. 7 shows a "phase diagram" in the space of L_1 and β of the metastable states of the cluster model. If we ran these simulations much longer, all the simulations would converge to one of two possible stable states. The stable (equilibrium) states of this model are either completely empty (decay) or completely occupied (expansion) at any value of L_1 and β. In both stable states, there is a small but finite probability that a single site in the lattice will flip to the opposite state. However, the probability that more than two neighboring sites will be flipped in the same time is very small and has never been observed in our simulations. Moreover, small clusters (eg. three) have a very short life span. Of course the probability of randomly creating clusters from an unoccupied state, or of removing a cluster of receptors from a fully occupied state can be altered if we change our parameters such as β and μ. However, our simulations were carried out with parameters that support the metastable states, and for these parameters, a random creation of a metastable cluster from an unoccupied state, and definitely of randomly switching from an unoccupied to a fully occupied state, is in practical terms zero.

4. Mathematical Appendix: Fluctuations in an Equilibrium Model.

I assume a very simple equilibrium model in which there are two populations of receptors, a synaptic population R_s and a non synaptic population R_i, such that $R_s + R_i = R_T$. The kinetic diagram describing this is

$$R_i \underset{K_{-1}}{\overset{K_1}{\rightleftarrows}} R_s, \tag{1}$$

where K_1 and K_{-1} are the forward and backward coefficients. The kinetic coefficients of the total number of receptors R_T might change as a function of synaptic plasticity. Here we will analyze these dynamics, assuming that kinetic coefficients and R_T remain constant. The dynamics of R_s can be characterized by the following equation:

$$\frac{dR_s}{dt} = K_1 R_T - (K_1 + K_{-1}) R_s. \tag{2}$$

This equation has the explicit solution:

$$R_s(t) = (R_s(0) - R_s^*) e^{-t/\tau_s} + R_s^* \tag{3}$$

where $R_s^* = \frac{K_1}{K_1+K_{-1}} R_T$ and $\tau_s = (K_1 + K_{-1})^{-1}$, and $R_s(0)$ is the initial condition. Since the different receptors are assumed independent, the probability of a given receptor i being in the synaptic state at time t, denoted as $P_i(t)$ is

$$P_i(t) = (P_s(0) - P_s^*) e^{-t/\tau_s} + P_s^* \tag{4}$$

where $P_s^* = \frac{K_1}{K_1+K_{-1}}$, and $P_s(0) = R_s(0)/R_T$. Using these dynamics we can also calculate the fluctuations of the number of synaptic receptors from their equilibrium value: The number of synaptic receptors R_s is: $R_s(t) = \sum_{i=1}^{R_T} \xi_i(t)$ where $\xi_i(t)$ is 1 if receptors i is in the synaptic state and 0 otherwise. I now proceed to calculate a cross correlation function $Q(\tau)$ where

$$Q(\tau) = \frac{1}{2}\left\langle (R_s(t+\tau) - R_s(t))^2 \right\rangle = \frac{1}{2}\left\langle \left(\sum_{i=1}^{R_T} \xi_i(t+\tau) - \sum_{i=1}^{R_T} \xi_i(t) \right)^2 \right\rangle, \tag{5}$$

where the average, denoted by $< \quad >$, is taken with respect to the probability of being in a synaptic state, as shown in Eq. **4**. I will carry out these calculations in the $t \to \infty$ limit, where the probabilities become independent of the initial conditions. I obtain:

$$Q(\tau) = R_T \left(< (\xi_i)^2 > - < \xi_i(t+\tau)\xi_i(t) > \right). \tag{6}$$

The subscript i can be omitted due to the independence of the different receptors, and this was obtained in the $t \to \infty$ limit.

The quantity $< \xi_i^2 >$ is equivalent to $P_s(t) \to P_s^*$, the probability that $\xi_i = 1$, at time t. The other term can be calculated by observing that: $< \xi_i(t)\xi_i(t+\tau) >= P\left(\xi_i(t) = 1, \xi_i(t+\tau) = 1\right) = P\left(\xi_i(t+\tau) = 1 | \xi_i(t) = 1\right) \cdot P(\xi_i(t) = 1)$ and therefore, using Eq. **4** I obtain

$$< \xi_i(t)\xi_i(t+\tau) >= P_s^* \left[(1 - P_s^*)e^{-\tau/\tau_s} + P_s^* \right]. \tag{7}$$

Therefore,

$$Q(\tau) = R_T P_s^*(1 - P_s^*)(1 - e^{-\tau/\tau_s}) \tag{8}$$

in the limit

$$Q(\tau \to \infty) = R_T P_s^*(1 - P_s^*). \tag{9}$$

Using the equilibrium value of $R_s^* = R_s(t \to \infty) = R_T P_s^*$ we obtain

$$F(\tau) = \frac{Q(\tau)}{< R_s^* >} = (1 - P_s^*)(1 - e^{-\tau/\tau_s}). \tag{10}$$

This statistical measure is independent of the total number of receptors R_T, is bounded above by 1, and monotonically decreases as a function of P_s^*.

In the limit $\tau \to \infty$ this converges to

$$F(\tau \to \infty) = \frac{Q(\tau)}{< R_s^* >} = (1 - P_s^*). \tag{11}$$

We can also calculate directly the Fano factor (F), which is the variance over the mean. Thus, at steady state,

$$F = \frac{\sigma^2}{\mu} = \frac{< R_s^2 > - < R_s >^2}{< R_s >} = 1 - P_s^*. \tag{12}$$

Therefore $F(\tau \to \infty) = F$, where F is the Fano factor calculated by ensemble averages over many instances of the same process, at steady state.

1. Plischke, M. & Bergersen B. (1994) Equilibrium Statistical Physics (World Scientific, Teaneck, NJ)