A triplet spike-timing–dependent plasticity model generalizes the Bienenstock–Cooper–Munro rule to higher-order spatiotemporal correlations

Julijana Gjorgjieva[a,1,2], Claudia Clopath[b], Juliette Audet[c], and Jean-Pascal Pfister[d,e]

[a]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom; [b]Laboratory of Neurophysics and Physiology, Université Paris Descartes, 75270 Paris, France; [c]Laboratory of Computational Neuroscience, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; [d]Computational Neuroscience Laboratory, Department of Physiology, University of Bern, CH-3012 Bern, Switzerland; and [e]Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom

Edited* by Leon N. Cooper, Brown University, Providence, RI, and approved September 30, 2011 (received for review April 14, 2011)

Synaptic strength depresses for low and potentiates for high activation of the postsynaptic neuron. This feature is a key property of the Bienenstock–Cooper–Munro (BCM) synaptic learning rule, which has been shown to maximize the selectivity of the postsynaptic neuron, and thereby offers a possible explanation for experience-dependent cortical plasticity such as orientation selectivity. However, the BCM framework is rate-based and a significant amount of recent work has shown that synaptic plasticity also depends on the precise timing of presynaptic and postsynaptic spikes. Here we consider a triplet model of spike-timing–dependent plasticity (STDP) that depends on the interactions of three precisely timed spikes. Triplet STDP has been shown to describe plasticity experiments that the classical STDP rule, based on pairs of spikes, has failed to capture. In the case of rate-based patterns, we show a tight correspondence between the triplet STDP rule and the BCM rule. We analytically demonstrate the selectivity property of the triplet STDP rule for orthogonal inputs and perform numerical simulations for nonorthogonal inputs. Moreover, in contrast to BCM, we show that triplet STDP can also induce selectivity for input patterns consisting of higher-order spatiotemporal correlations, which exist in natural stimuli and have been measured in the brain. We show that this sensitivity to higher-order correlations can be used to develop direction and speed selectivity.

Synaptic plasticity depends on the activity of presynaptic and postsynaptic neurons and is believed to provide the basis for learning and memory (1, 2). It has been shown that low-frequency stimulation (1–3 Hz) (3) or stimulation paired with low postsynaptic depolarization (4) induces synaptic long-term depression (LTD), whereas synapses undergo long-term potentiation (LTP) after high-frequency stimulation (100 Hz) (5). Such findings are consistent with the well-known Bienenstock–Cooper–Munro (BCM) learning rule (6). This BCM model has been shown to elicit orientation selectivity and other aspects of experience-dependent cortical plasticity (6, 7). Furthermore, in this model the modification threshold between LTP and LTD varies as a function of the history of postsynaptic activity, a prediction that has been confirmed experimentally (8).

Despite its consistency with experimental data and its functional relevance, the BCM framework is still limited experimentally and functionally. Experimentally, because the learning rule is expressed in terms of firing rates, it cannot predict synaptic modification on the basis of the timing of pre- and postsynaptic spikes (9, 10). This form of plasticity, called spike-timing–dependent plasticity (STDP), uses the timing of spike pairs to induce synaptic modification (11, 12). The presynaptic spike is required to shortly precede the postsynaptic spike to elicit LTP, whereas the reverse timing of pre- and postsynaptic spikes leads to LTD (9, 10). Functionally, the BCM model cannot segregate input patterns that are characterized by their temporal spiking structure. STDP provides a possible solution, but how STDP relates to BCM remains debated (13–15).

Here, we consider a spike-based learning rule, "the triplet STDP model" (15, 16), and show that it overcomes those two important limitations of the BCM rule and thus generalizes the BCM framework. This triplet model uses sets of three spikes (triplets)—instead of pairs of spikes as in the case of classical STDP—to induce potentiation. More precisely, LTP depends on the interval between the pre- and postsynaptic spikes and on the timing of the previous postsynaptic spike (Fig. 1*A*). Furthermore, this triplet learning rule has been shown to explain a variety of synaptic plasticity data (17, 18) significantly better than pair-based STDP (15) (Fig. 1*B*). Plasticity induced by multiples of spikes has also been the focus of other studies (19, 20); despite using the same spike combinations some differences have been observed, most likely due to the different (extracellular or intracellular) stimulation protocols used in these studies (21).

Computationally, it has been shown that under some rather crude assumptions—when the input and output neurons have independent Poisson statistics—the triplet STDP model can be mapped to the BCM learning rule (16). In this paper, we take a more biologically plausible approach by incorporating contributions from input–output spiking correlations in inducing synaptic plasticity. Consistent with results from the BCM theory, we demonstrate that in the presence of orthogonal rate-based patterns, the maximally selective fixed points of the weight dynamics induced by the triplet rule are stable. Furthermore, we show that the triplet rule acts as a generalized BCM rule in the sense that postsynaptic neurons become selective not only to rate-based patterns of the inputs, but also to patterns differentiated only by their spiking correlation structure. The mathematical simplicity of the triplet model allowed us to characterize the explicit dependence of the weight dynamics on higher-order input correlations. We believe this study is of great relevance given the ubiquity of higher-order correlations in the brain (22, 23) and their relevance for neural coding (24).

Model and Methods

Neuronal Dynamics. We considered a feedforward network with N input neurons $x_j(t)$ as Dirac delta spike trains connected to a single output neuron through the weights $w_j(t)$ and giving rise to the postsynaptic spike train $y(t)$ (*SI Text*). The input spike trains had average firing rates $\rho_j(t)$.

We assumed that the membrane potential of the postsynaptic neuron $u(t)$ increased with the spike times of each input by the excitatory postsynaptic potential (EPSP) scaled by the corresponding weight

$$u(t) = \sum_{k=1}^{N} w_k(t) \int_0^\infty \varepsilon(r) x_k(t-r) \mathrm{d}r. \quad [1]$$

The function $\varepsilon(r)$ denoted the EPSP kernel, taken to be a decaying exponential with a membrane time constant of 11 ms. For spatio-temporal

Author contributions: J.G., C.C., and J.-P.P. designed research; J.G. performed research; J.G., J.A., and J.-P.P. contributed new reagents/analytic tools; J.G. and C.C. analyzed data; and J.G., C.C., and J.-P.P. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]Present address: Center for Brain Science, Harvard University, Cambridge, MA 02138.

[2]To whom correspondence should be addressed. E-mail: gjorgjieva@fas.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1105933108/-/DCSupplemental.

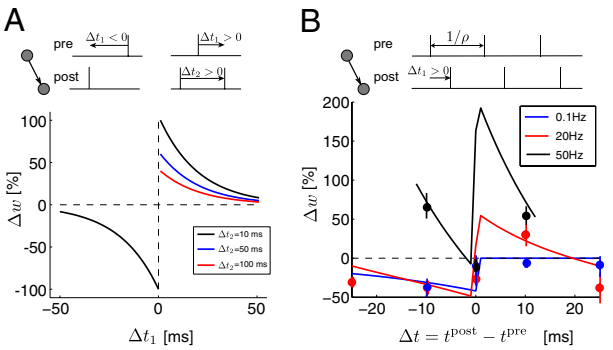

Fig. 1. The triplet STDP rule. (*A*) Synaptic depression is induced as in classical pair-based STDP using spike pairs separated by $\Delta t_1 = t^{post} - t^{pre} < 0$. Synaptic potentiation is induced using triplets of spikes consisting of two postsynaptic spikes and one presynaptic spike on the basis of the timing interval between them $\Delta t_1 = t^{post} - t^{pre} > 0$ and $\Delta t_2 = t^{post} - t'^{post} > 0$. (*B*) Synaptic change as a function of the time between pre- and postsynaptic spikes in a protocol where 60 pairs were presented at different frequencies ρ = 0.1, 20, and 50 Hz. Depression predominated at low frequency, whereas potentiation was more prevalent at high frequencies. The data points are experiments are from ref. 17 and the lines were generated with the triplet STDP rule with the parameters taken from ref. 15.

receptive field development, an inhibitory postsynaptic membrane potential kernel with a membrane time constant of 20 ms was also used. Postsynaptic spikes were generated stochastically from the membrane potential, with a probability density of firing a spike at time t given by the transfer function $\langle y(t)\rangle = g(u(t))$. For simplicity, we used linear neurons where the transfer function was approximated by

$$g(u(t)) \approx g(u_0) + g'(u_0)(u(t) - u_0), \qquad [2]$$

where the averaged membrane potential was u_0. We also used $\nu = g(u_0)$ to denote the mean postsynaptic firing rate.

Synaptic Dynamics and Input Selectivity. Following the approach of ref. 25, we expressed the weight change as a Volterra expansion of both pre- and postsynaptic spike trains and the two learning rules: pair-based STDP

$$W_2(\Delta t) = \begin{cases} A_2^+ e^{-\Delta t/\tau_+}, & \Delta t \geq 0 \\ -A_2^- e^{\Delta t/\tau_-}, & \Delta t < 0, \end{cases} \qquad [3]$$

where $\Delta t = t^{post} - t^{pre}$ denotes the timing difference between a post- and a presynaptic spike, τ_+ is the potentiation time constant, and τ_- is the depression time constant and triplet STDP

$$W_3(\Delta t_1, \Delta t_2) = A_3^+ e^{-\Delta t_1/\tau_+} e^{-\Delta t_2/\tau_y}, \ \Delta t_1 \geq 0, \Delta t_2 \geq 0 \qquad [4]$$

and 0 otherwise, where spike triplets (t^{pre}, t^{post}, t'^{post}) affect synaptic potentiation depending on their timing difference $\Delta t_1 = t^{post} - t^{pre}$ and $\Delta t_2 = t^{post} - t'^{post}$. The parameters used throughout this paper were those of the minimal triplet rule (15), i.e., $A_2^+ = 0$, $A_2^- = -6.5 \times 10^{-3}$, $A_3^+ = 7.1 \times 10^{-3}$, τ_+ = 16.8 ms, τ_- = 33.7 ms, τ_y = 114 ms. Assuming slow learning dynamics (25), we derived Eq. **6** (*Results*) to describe the weight dynamics (*SI Text*).

We considered M input patterns, where pattern i had mean firing rate $\rho^{(i)}$, and pairwise and triplet correlation terms $A^{(i)}$ and $B^{(i)}$, respectively. Each input pattern i was associated with a probability p_i of occurrence and gave rise to an average postsynaptic firing rate $\nu^{(i)} = w^T \rho^{(i)}$. The selectivity of the postsynaptic neuron was $\mathrm{Sel}(w) = 1 - (\sum_i p_i w^T \rho^{(i)})/(\max_i w^T \rho^{(i)})$ (6).

To match the triplet rule to the BCM model, we set $A_2^- \to A_2^- \bar{\nu}/\rho_0^p$, where the expectation of the p^{th} power of the postsynaptic firing rate can be expressed as $\bar{\nu} = \sum_{i=1}^M p_i (\nu^{(i)})^p$. This quantity was approximated by low-pass filtering the p^{th} power of the instantaneous postsynaptic firing rate $\nu(t) = g(u(t))$ with a time constant of τ_r = 5 s. For all of the calculations in this paper we took p = 2.

In the case of orthogonal rate-based patterns modeled as independent Poisson inputs, we proved that the maximally selective fixed points of the weight dynamics are stable (*SI Text*). For the development of selectivity in the case of correlation-based patterns, we calculated the fixed points of maximal selectivity only in the case of a reduced 2D system (*SI Text*). In this case, two patterns were presented to the feedforward network, each consisting of two groups (or pools) of input neurons. Extensions to more than two correlated patterns are currently possible only with numerical simulations.

Numerical Simulations with Multiple Patterns. For all numerical simulations we simulated the triplet learning rule given by Eq. **4** that can also be expressed in differential form (*SI Text*). A lower bound of 0 and an upper bound of 3 were imposed on the weights. The methods for generating correlated spike trains and the correlation strength used in each figure are described in the *SI Text*. The phase plane diagrams for the 2D systems in Figs. 2*D* and 3*E* were plotted using the MATLAB software pplane written by John Polking (Rice University, Houston, TX).

Results

Triplet STDP Induces Selectivity with Rate-Based Patterns. Orientation-selective neurons in the primary visual cortex respond with higher firing rates when a bar is presented in a particular orientation and with lower rates when the bar is presented in a different orientation (26). This orientation selectivity is learned during receptive field development, and normal patterns of sensory experience are important for receptive field maturation (27). Bienenstock et al. (6) proposed a model for how orientation selectivity, or more generally pattern selectivity, is learned by a neural network: the BCM learning rule. In the BCM framework, a randomly chosen input pattern i (of M possible patterns) with rates $\rho^{(i)}$ is presented with probability p_i to a feedforward network with N inputs. The postsynaptic neuron responds with a firing rate $\nu^{(i)} = w^T \rho^{(i)}$, where w is the weight vector. The weight change induced by the BCM rule is proportional to the input firing rate

$$\dot{w} = \phi(\nu, \bar{\nu})\, \rho \qquad [5]$$

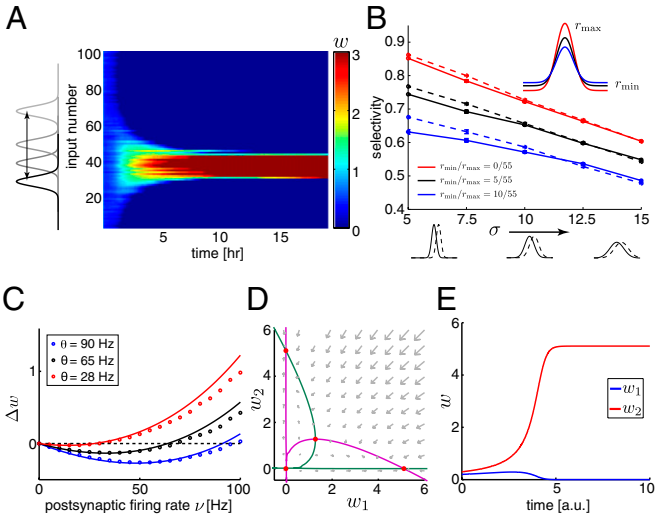

Fig. 2. Triplet STDP induces selectivity with rate-based patterns. (*A*) Evolution of the weights (*Right*) for 10 rate-based patterns (uniformly spaced Gaussian profiles across the 100 inputs) determining the inputs' firing rates (*Left*) presented to a feedforward network. The selected pattern corresponds to a Gaussian profile with r_{min}/r_{max} = 5/55 and σ = 15 Hz. (*B*) Mean (±SEM) selectivity (for 10 trials) as a function of the Gaussians' SD σ and for Gaussian profiles with different ratios of background to peak firing rates r_{min}/r_{max} = {0/55, 5/55, 10/55} (solid lines, the triplet STDP rule; dashed lines, the BCM rule). The Gaussian profiles below illustrate the amount of overlap for two neighboring Gaussians. Numerical simulations implementing the differential form of the triplet STDP were performed in *A* and *B*. (*C*) Weight change Δw as a function of postsynaptic activity for three different input firing rates, which determine the threshold θ for weight modification. Symbols denote numerics and lines analytics. (*D*) 2D phase plane analysis for the analytically derived weight equation with orthogonal rate-based patterns. Nullclines in green and purple intersect at the equilibria shown in red. (*E*) An example trajectory for the two weights attracted to one of the stable nodes in *D*. (a.u., arbitrary units).

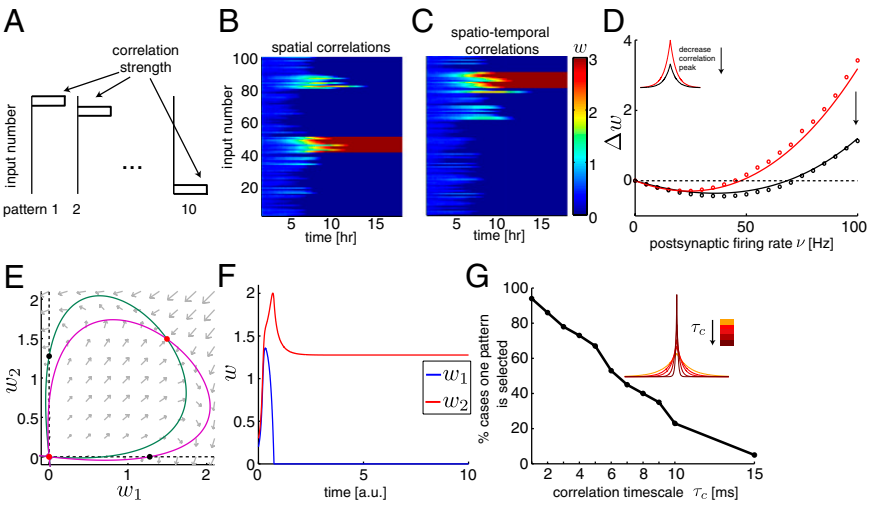

Fig. 3. Triplet STDP induces selectivity with correlation-based patterns. (*A*) Ten correlation-based patterns that have the same firing rates, but different correlation strength. (*B*) Evolution of the weights illustrates selectivity in the case of 10 correlation-based patterns. The firing rate of each of the 100 inputs was set to 10 Hz: 90 inputs had no correlations and 10 neighboring inputs (one of 1–10, 11–20, ... , 91–100 for each pattern) had strong spatial correlations (90% identical spikes). (*C*) Same as *B* except for the 10 correlated inputs in each pattern, for which exponentially decaying correlations with a time constant of 5 ms were used. Numerical simulations implementing the differential form of the triplet STDP were performed in *B* and *C*. (*D*) The average weight change Δw (for 100 weights) was computed for different initial conditions w_0 after 100 s. The symbols denote numerical results obtained by simulating the differential form of the triplet rule, and the lines indicate a semi-analytic solution by numerically solving Eq. **6** given an initial condition w_0 for 100 s. The average weight change was plotted as a function of the postsynaptic firing rate given by $\nu = w_0\rho$, where ρ was the input firing rate. Here we simulated two networks where the inputs had the same firing rate (10 Hz) and exponentially decaying correlations with a timescale of 10 ms. The correlation peak for the curve in black (*SI Text*) was half of the correlation peak for the curve in red ($\gamma = 9.09$, $\lambda = 9.09$); see *Inset*. (*E*) 2D (two groups of inputs) phase plane analysis for correlation-based patterns. Nullclines in green and purple intersect at the unstable fixed points shown in red. Imposing a lower bound at 0 resulted in stable maximally selective fixed points on the axes shown in black. (*F*) An example trajectory for the two weights attracted to one of the black equilibria in *E*. (*G*) Percentage of cases (over 100 trials) where all of the synapses from 1 of 10 input patterns (each consisting of 10 inputs) potentiate (Eq. **6**). Same scenario as *C*, but for different correlation time constants τ_c. Correlations were symmetric and exponentially distributed (*Inset*).

and scales with a nonlinear function ϕ, which depends not only on the postsynaptic firing rate ν, but also on the average (over all patterns) of a nonlinear function of the postsynaptic rate $\overline{\nu} = \sum_i p_i (\nu^{(i)})^2$. The nonlinear function ϕ must be negative when the postsynaptic firing rate is below a given threshold θ—which itself depends on $\overline{\nu}$—and positive when it is above it (Fig. 2*C*).

Interestingly, assuming a linear transfer function, the average weight change under the triplet rule can be written precisely as the BCM term plus some perturbation terms due to the input correlations (*SI Text*)

$$\dot{w}_j = \phi(\nu)\rho_j + \sum_{k=1}^{N} \Delta A_{jk} w_k + w^{\mathrm{T}} \Delta B_j w \qquad [6]$$

where $\phi(\nu) = \overline{W}_2\nu + \overline{W}_3\nu^2$ is the BCM term (Fig. 2*C*) with $\overline{W}_2$ and $\overline{W}_3$ being the overall area under the pair-based and triplet STDP rules, respectively. ΔA and ΔB describe the contributions from the input statistics (*SI Text*). To get depression at low postsynaptic firing rate and potentiation for higher firing rate, pairs of spikes must have an overall depressive effect ($\overline{W}_2 < 0$) and triplets of spikes must induce potentiation ($\overline{W}_3 > 0$), as is the case for the minimal triplet STDP model considered here (15).

There are two differences between the original BCM rule in Eq. **5** and the triplet model in Eq. **6**. First, in the triplet model, the function ϕ depends only on the temporally averaged postsynaptic activity ν, whereas in the BCM model, ϕ also depends on the postsynaptic activity averaged over all patterns, $\overline{\nu}$. However, if we redefine the amplitude parameter for pair-based depression A_2^- as $A_2^- \overline{\nu}/\rho_0^2$, where ρ_0 is a constant denoting the target rate of the postsynaptic neuron (15), then both ϕ and ΔA in Eq. **6** will depend on $\overline{\nu}$.

The second difference is the presence of the two additional terms (ΔA and ΔB) in Eq. **6**. If the inputs are Poisson neurons, we can rewrite Eq. **6** as

$$\dot{w} = (\phi(\nu, \overline{\nu})\mathbf{1} + \Lambda)\,\rho, \qquad [7]$$

where $\mathbf{1}$ denotes the identity matrix and Λ is a diagonal matrix (*SI Text*). If we now assume that the patterns are orthogonal, we can show that the condition $\dot{w} = 0$ in Eq. **7** gives rise to 2^N fixed points. Moreover, the N maximally selective fixed points, $w^{\star(n)} = (0, \ldots, 0, w_n^\star, 0, \ldots, 0)$, are stable fixed points (Fig. 2*D* and *E* shows an illustration in two dimensions), which is consistent with results of the BCM theory.

The general problem of deriving selectivity analytically has not yet been solved; however, numerical simulations suggest that the triplet rule successfully drives selectivity even when the rate-based inputs are nonorthogonal. We designed an experiment to examine the level of selectivity as 10 Gaussian input patterns were presented to the network with varying amounts of overlap (Fig. 2*A*). The Gaussian profiles were uniformly spaced across the input neurons, and we varied the ratio of the background firing rate ($r_{\min}$) to the peak firing rate ($r_{\max}$) of each Gaussian (Fig. 2*B*, *Inset*) and the SD (σ). The Gaussian profiles were closest to orthogonal for small σ and $r_{\min}/r_{\max} = 0$, whereas their amount of overlap increased as either σ or $r_{\min}/r_{\max}$ increased. We computed the amount of selectivity of the postsynaptic neuron at the end of a simulation when the weights reached a steady state (Fig. 2*A*) as a function of the Gaussians' SD (σ). We observed that for the case of nearly orthogonal Gaussian profiles ($r_{\min}/r_{\max} = 0/55$ and small σ, Fig. 2*B*, red lines), the achieved selectivity was close to the maximally attainable selectivity of 0.9 for 10 orthogonal patterns (6). The selectivity dropped as σ or $r_{\min}/r_{\max}$ increased (Fig. 2*B*). We compared the performance of the triplet to the BCM rule using the same Gaussian input profiles, while keeping the weights nonnegative during the entire simulation (Fig. 2*B*, dashed lines). As expected, we obtained similar results to those of the triplet rule.

Triplet STDP Induces Selectivity with Correlation-Based Patterns. In addition to mapping the triplet to the BCM rule for rate-based patterns, the triplet rule further generalizes the BCM model: In Eq. **6**, ΔA and ΔB depend on the second- and third-order input correlations, respectively (*SI Text*); therefore, we expected triplet STDP to be sensitive to spatiotemporal correlations in the inputs.

To examine our hypothesis, we presented 10 "correlation-based" patterns to the feedforward network with 100 inputs. The correlation-based patterns were determined by different pairwise and third-order correlations (Fig. 3*A*), but had the same input firing rates. Therefore, the response of the postsynaptic neuron to each pattern was the same, which prevented us from using the same measure of selectivity as for rate-based patterns. Instead, selectivity was defined in terms of the selective potentiation of a group of correlated inputs. Fig. 3*B* shows a simulation with purely spatial correlations that had no temporal structure (the input correlations were due to identical spikes in the neurons). The weights from one pattern potentiated (inputs 41–50), whereas the other weights depressed. When we presented spatiotemporal correlations with an exponentially decaying

correlation function, selectivity was also achieved: A set of 10 weights (81–90) characteristic of one pattern potentiated, whereas the other weights depressed (Fig. 3*C*).

In the case of correlation-based patterns, the sliding threshold depends both on the input firing rates and correlations. Although the additional terms ΔA and ΔB in Eq. **6** prevent us from deriving an explicit expression for the modification threshold, we illustrated the dependence of the threshold on the correlation strength in Fig. 3*D*. Here we computed the average weight change for 100 weights as a function of the postsynaptic firing rate for two different input correlations: In both cases the firing rate of the 100 inputs was the same (10 Hz), and the correlation function was a decaying exponential with a timescale of 10 ms (Fig. 3*D*, *Inset*); however, the two functions differed in the correlation peak (black peak was one-half of the red peak). The network with the higher correlation peak had a lower modification threshold, resulting in a larger potentiation region.

Due to the increased complexity of the system when the inputs are correlated, we derived the fixed points of maximal selectivity $(w_1^\star, 0)$ and $(0, w_2^\star)$ in a small network of two groups of input neurons and analyzed their stability (*SI Text*). A lower bound had to be introduced to prevent the weights from becoming negative (in agreement with Dale's law). For the 2D network, we found that the maximally selective fixed points were always stable. Fig. 3*E* shows the 2D phase plane, where the two unstable fixed points (red symbols) drive the weight trajectories toward the axes where the stable maximally selective fixed points are located (black symbols). Example weight trajectories are shown in Fig. 3*F* for one choice of initial condition.

We extended the simulation in Fig. 3*C* to examine how the temporal correlation structure of the inputs influences the selective potentiation of synaptic weights corresponding to different patterns. We studied a particular example of a symmetric spatiotemporal correlation: an exponential function decaying in time, which was the same for all pairs and the same for all triplets of inputs (*SI Text*). Therefore, while preserving the correlation strength, we examined the role of the correlation timescale on the selective potentiation of synaptic weights (Fig. 3*G*, *Inset*). Increasing the correlation timescale had a similar effect as "diluting" the correlation strength. The triplet STDP rule failed to consistently potentiate the weights of one input pattern, and often two or three patterns were simultaneously selected (Fig. 3*G*). Therefore, correlations over broad timescales fail to evoke selective potentiation of correlation-based input patterns and could be used to understand the implications of different correlation structures in different brain regions.

Triplet STDP, but Not Pair-Based STDP, Can Induce Selectivity Driven by Third-Order Correlations. Despite the advantage of triplet STDP over classical pair-based STDP to capture a large variety of experimental plasticity data (for instance, frequency dependence) (17), we asked whether triplet STDP can do computations that pair-based STDP cannot. Previous studies have shown that for correlation-based patterns pair-based STDP selects the correlated groups of inputs in the case of static patterns (where the correlations are always presented to the same group of inputs), but have not addressed the case of dynamic patterns (28). We hypothesized that triplet STDP will be able to select patterns determined by the inputs' third-order correlations, whereas pair-based STDP will not be able to distinguish any higher-than-pairwise correlations.

For this task, we designed a selectivity scenario consisting of two correlation-based patterns presented to a feedforward network of six input neurons. The inputs in the two patterns consisted of the same firing rates and the same pairwise correlations, but differed in the presence or absence of third-order correlations in half of the inputs (Fig. 4*A*). Pattern 1 consisted of third-order correlations in inputs 1–3 (denoted as group 1) and no third-order correlations in inputs 4–6 (denoted as group 2). Pattern 2 consisted of third-order correlations in inputs 4–6 of group 2 and no third-order correlations in inputs 1–3 of group 1. Next, we presented each pattern to the network with a fixed probability; for instance, in Fig. 4*A* we illustrate a scenario in which, of 10 pattern presentations, pattern 1 was presented on average eight times (with probability 0.8). As the probability of presenting pattern 1 varied between 0.5 and 1 (Fig. 4*C*), we estimated the probability that pattern 1 wins in 200 simulation runs. An example of pattern 1 winning is illustrated in Fig. 4*B*, where inputs 1–3 potentiate, and inputs 4–6 depress to 0. We estimated the probability that pattern 1 wins for both pair STDP (Fig. 4*C*, red symbols) and triplet STDP (Fig. 4*C*, black symbols). When only pattern 1 was presented to the network (third-order correlations only in inputs 1–3), almost all simulations resulted in the potentiation of these inputs under the triplet STDP rule. As the probability of presenting pattern 1 decreased to 0.5, the probability that pattern 1 wins also decreased. However, pair-based STDP was not sensitive to the third-order input correlations and it treated both patterns equally, selecting each pattern randomly with equal probability of 1/2 regardless of how frequently pattern 1 was presented.

This result demonstrates that the triplet STDP rule can distinguish between inputs solely on the basis of the higher-order correlation structure, which pair-based STDP ignores. As a result, triplet STDP will be computationally more powerful in systems where such higher-order correlations have been characterized (22–24) and where firing rates and pairwise correlations are of similar magnitude.

These studies demonstrate that because pair-based STDP uses only pairs of spikes to induce synaptic plasticity, it is sensitive only up to pairwise correlations. Thus, we suspected that the triplet STDP rule, which evokes plasticity using triplets of spikes, will be sensitive only up to third-order correlations. To confirm this, we repeated the simulation scenario above for a network of two groups of five neurons each. In each case, the two input groups had the same lower-order correlations, but differed in the presence or absence of higher-order correlations in each group. We studied correlations with highest order of five. The triplet

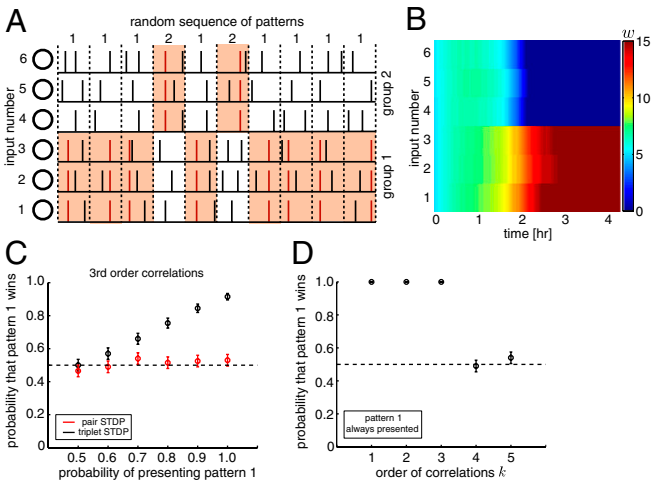

Fig. 4. Triplet STDP, and not pair-based STDP, can distinguish between patterns determined by third-order correlations. (*A*) Two patterns were randomly presented to a feedforward network: The inputs in the two patterns had the same firing rates and the same pairwise correlations. The patterns differed only by the presence or absence of third-order correlations in half of the inputs (illustrated with the red triplets of spikes and the colored background). The probability of presenting pattern 1 was varied, e.g. of 10 pattern presentations, pattern 1 was presented with probability 0.8. (*B*) The evolution of the weights under the triplet STDP rule demonstrating an example where pattern 1 (inputs 1–3) wins. (*C*) Pattern 1 was presented to the network at different probabilities and the mean ± SEM of the probability that pattern 1 wins was computed for 200 simulations runs: triplet STDP (black symbols) and pair STDP (red symbols). (*D*) The triplet STDP rule is sensitive up to third-order correlations, but not to higher-order correlations. Pattern 1 was always presented and consisted of two groups of five neurons each. Both groups had the same correlations up to (but not including) order k (horizontal axis). Group 1 had nonzero $\geq k^{\text{th}}$-order correlations, and group 2 had zero $\geq k^{\text{th}}$-order correlations. The mean ± SEM of the probability that pattern 1 wins (i.e., all of the weights of group 1 potentiate) was computed for 200 simulation runs. Dashed lines correspond to chance level.

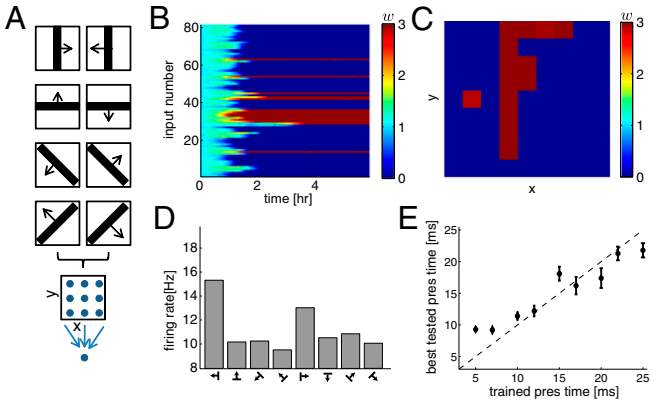

Fig. 5. Triplet STDP leads to spatiotemporal receptive field development. (*A*) Four different bars (horizontal, vertical, and the two diagonals on a 9 × 9-pixels image) were presented as inputs to a feedforward network with a single postsynaptic neuron; each bar can move in one of two directions, giving a total of eight patterns. (*B*) Time evolution of the 81 synaptic weights. (*C*) Final weights reordered in a grid corresponding to the input location. (*D*) Histogram of the postsynaptic firing rate plotted after convergence of the weights, at the end of the learning in *B*. The firing rates shown resulted from the presentation of the eight different patterns (four orientations and two directions) averaged over 200 s. (*E*) After learning with different training presentation times (5, 7, 10, 12, 17, 20, 22, and 25 ms), the weights were frozen during a testing phase. The pattern (of the eight patterns) that resulted in the highest firing rate at the training presentation time was presented again to the network at different tested presentation times (5, 7, 10, 12, 17, 20, 22, and 25 ms) while the firing rate was measured. The best tested presentation time for which the firing rate (averaged over 100 s) was the highest is plotted against the training presentation time (mean ± SEM over 10 trials).

rule distinguished correlations up to third-order, but was insensitive to fourth- and fifth-order correlations (Fig. 4*D*). We expect that for a learning rule to be sensitive to higher than third-order correlations, the rule would need to incorporate more than three spikes, or the neural model would have to be nonlinear.

Spatiotemporal Receptive Field Development. Due to its sensitivity to higher-order correlations, we expected that the triplet rule would succeed in driving the development of spatiotemporal receptive field properties encountered in visual cortex, such as orientation and direction selectivity. Using the same feedforward framework we presented eight different patterns consisting of four bars at different orientations, each moving in one of two directions, as drawn in Fig. 5*A*. The input neurons in the network were organized in a 9 × 9 grid mapping to a single postsynaptic neuron. Each input spike produced both an EPSP and an IPSP with a longer time constant (*Model and Methods*) (29). The synapses in the network selectively refined under the triplet learning rule, some depressing to the lower bound of zero and others potentiating maximally to the upper bound (Fig. 5*B*). At the end of the simulation the postsynaptic neuron became selective to one of the bars, as shown in Fig. 5*C* by the strong weights corresponding to inputs arranged in the bar with a vertical orientation (although this selectivity was not as robust as for nonmoving bars, i.e., Fig. 2).

Recently, after training with moving stimuli, Li et al. (30) observed the emergence of direction selectivity in cortical neurons of visually naive ferrets for the trained directions of motion. Motivated by this experimental result, we examined the histogram of the postsynaptic firing rate for all patterns after learning (Fig. 5*D*). We found that the postsynaptic neuron responded with the highest firing rate for a bar in given (vertical) orientation and moving in one direction (left), thus becoming weakly selective to direction.

The receptive field in this scenario also developed a temporal structure as a result of the spatiotemporal correlations imposed by the moving bars. To visualize this effect, the weights were frozen after learning (at a given presentation time) and a test phase was conducted. The pattern that produced the highest firing rate for the trained presentation time was tested at a range of testing presentation times. The best testing presentation time for which the firing rate was the highest was plotted against the training presentation time. There was a high correlation between training presentation time and best tested presentation time, demonstrating that the neuron became selective to the training speed (Fig. 5*E*).

In this scenario the development of direction selectivity relied on two elements. First, triplet STDP was sensitive to the spatiotemporal input correlations, and second, we used a modified postsynaptic potential kernel that included an inhibitory component with a longer time constant than for the excitatory current. This additional IPSP was necessary for obtaining direction selectivity because it made the postsynaptic neuron sensitive to a temporal derivative of the input currents. A similar assumption has been used in previous models of direction selectivity driven by pair-based STDP (29, 31), although in these models the postsynaptic neuron was usually trained in a single direction. The temporal sensitivity to obtain direction selectivity can be also implemented by the order of presentation of the stimuli, instead of modifying the dynamics of the postsynaptic neuron, as shown with the BCM rule (32). More robust direction selectivity as observed biologically (26, 33) can also be obtained by using a network with recurrent excitatory and inhibitory connections (29), where the source of inhibition could arise from neocortical interneurons, such as fast-spiking interneurons. Finally, we note that the receptive fields we observed depend on the choice of input statistics (here moving bars), the learning rule, and the neural model. Receptive fields with more complex structure (such as Gabor patches) could be developed with inputs of richer spatiotemporal contents, such as natural images.

Discussion

The BCM theory is attractive because it generates selectivity in a variety of scenarios and has been supported experimentally (6–8). Synaptic plasticity, however, has been shown to depend on the precise spike timing (9, 10) classically modeled by pair-based STDP (11, 12). In this paper, we show that a different spike-based rule, triplet STDP, known to accurately capture plasticity experiments (15), exhibits the computational properties of the BCM rule and is additionally sensitive to higher-order spatiotemporal input correlations.

We mapped the triplet STDP to the BCM learning rule for rate-based patterns, determined by the input firing rates. Consistent with the BCM theory, we showed that for nonoverlapping (orthogonal) patterns, the maximally selective fixed points of the weight dynamics under triplet STDP are always stable. For overlapping Gaussian patterns, numerical simulations demonstrated that the selectivity achieved with the triplet rule is similar to the selectivity achieved with the BCM rule.

We also showed that the triplet rule can generate selectivity in the case of correlation-based dynamic patterns, determined solely by the higher-order input correlations. However, because the rule uses triplets of spikes to induce plasticity, it is sensitive to higher-order correlations of maximum order three. This sensitivity led to the development of direction selectivity and speed selectivity. We observed that increasing the input correlation timescale dilutes the correlation strength, which prevents the cooperation of inputs necessary for the emergence of selectivity. Therefore, our results make experimental predictions about the types of correlation structure that lead to selectivity.

Higher-order correlations have not only been measured in the brain, but also shown to play an important role in visual coding and representing experimental data (22–24, 34). Higher-order correlations are ubiquitous in sensory stimuli, such as natural stimuli and speech signals (35, 36). These correlations have been previously used in learning rules to extract the independent components or features in natural images resulting in simple cell receptive fields as seen in V1 (35, 37). One such rule is the BCM rule, shown to perform projection pursuit that relies on higher-order correlations to find the most interesting projection, i.e., the one that minimizes the Gaussianity of the output distribution (38, 39) and is closely related to independent component analysis (ICA). Because of its mapping to the BCM rule, we can interpret

the triplet rule as a method for performing such ICA-like computations. In addition to spatial ICA (where the independent components are obtained from the input statistics at each fixed time point) (35), the triplet rule can also perform temporal ICA-like computations, which additionally rely on the temporal structure of the inputs (40).

Several other models have addressed the issue of spiking-based implementations of the BCM learning framework (13, 14). Izhikevich and Desai (14), for instance, proposed that by implementing classical pair-based STDP with nearest-neighbor spike interactions, the rule can be mapped to the BCM rule. However, their model failed to capture the frequency dependence of ref. 17 if pairs of spikes are presented at different frequencies (21) and considered the rather crude approximation that input and outputs are independent. Although the model of Senn et al. (13) captured the frequency dependence of the pairing protocol, it could not reproduce the triplet and quadruplet experiments of ref. 18 (see ref. 15) and the correspondence to the BCM rule is only approximate (the sliding threshold depends on the weights and not on the postsynaptic firing rate as in the BCM framework). Toyoizumi et al. (41) derived an alternate spike-based learning rule designed to maximize the information transmission between an ensemble of inputs and the output of a postsynaptic neuron. Although such plasticity rules derived from the infomax principle can generalize the BCM theory to spiking neurons and can be reduced under some assumptions to the triplet STDP rule (42), the dynamics of these rules are rather complicated to be studied analytically in contrast to the triplet STDP model. The same problem arises with biophysical models (43, 44) or with more elaborate phenomenological models (45), which have primarily been studied numerically. Therefore, the triplet STDP model is a good trade-off: It can reproduce a large set of electrophysiological data and yet has a relatively simple formulation so that we can study it analytically and generalize its functional properties to networks of different size and input statistics. Additionally, the triplet STDP model extends the BCM theory to correlation-based patterns with higher-order correlations, which were not considered by any of the above models.

The emergence of input selectivity in this work has been studied with a linear neuron. This analysis can be extended to consider nonlinear neurons and to relate it to a variant of the BCM theory that uses nonlinear neurons (38). We expect that with an appropriate nonlinearity, this variant will lead to a more robust orientation selectivity for natural image stimuli (46). Furthermore, because the triplet rule is also sensitive to the temporal structures of the inputs, we expect it to be able to capture relevant aspects of the spatiotemporal statistics of natural scene environments.

ACKNOWLEDGMENTS. We thank Stephen Eglen, Adrienne Fairhall, and Robert Froemke for helpful discussions and reading drafts of this manuscript. This work was funded by a Cambridge Overseas Research Studentship and Trinity College Internal Graduate Studentship (to J.G.), the Agence Nationale de la Recherche Grant ANR-08-SYSC-005 (to C.C.), the Wellcome Trust, and the Swiss National Science Foundation Grant 31-133094 (to J.-P.P.).

1. Bliss TV, Collingridge GL (1993) A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature* 361(6407):31–39.
2. Malenka RC, Nicoll RA (2009) Long-term potentiation—a decade of progress? *Science* 285:1870–1874.
3. Dudek SM, Bear MF (1992) Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proc Natl Acad Sci USA* 89:4363–4367.
4. Artola A, Bröcher S, Singer W (1990) Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature* 347(6288):69–72.
5. Bliss TV, Lomø T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol* 232:331–356.
6. Bienenstock EL, Cooper LN, Munro PW (1982) Theory of the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *J Neurosci* 2:32–48.
7. Cooper LN, Intrator N, Blais BS, Shouval HZ (2004) *Theory of Cortical Plasticity* (World Scientific, Singapore).
8. Kirkwood A, Rioult MC, Bear MF (1996) Experience-dependent modification of synaptic plasticity in visual cortex. *Nature* 381(6582):526–528.
9. Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postysnaptic AP and EPSP. *Science* 275:213–215.
10. Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci* 18: 10464–10472.
11. Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383(6595):76–78.
12. Abbott LF, Nelson SB (2000) Synaptic plastictiy - taming the beast. *Nat Neurosci* 3: 1178–1183.
13. Senn W, Markram H, Tsodyks M (2001) An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. *Neural Comput* 13:35–67.
14. Izhikevich EM, Desai NS (2003) Relating STDP to BCM. *Neural Comput* 15:1511–1523.
15. Pfister JP, Gerstner W (2006) Triplets of spikes in a model of spike timing-dependent plasticity. *J Neurosci* 26:9673–9682.
16. Pfister JP, Gerstner W (2006) Beyond pair-based STDP: A phenomenological rule for spike triplet and frequency effects. *Adv Neural Inf Process Syst* 18:1083–1090.
17. Sjöström PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32:1149–1164.
18. Wang HX, Gerkin RC, Nauen DW, Wang GQ (2005) Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nat Neurosci* 8:187–193.
19. Froemke RC, Dan Y (2002) Spike-timing dependent plasticity induced by natural spike trains. *Nature* 416(6879):433–438.
20. Froemke RC, Tsay IA, Raad M, Long JD, Dan Y (2006) Contribution of individual spikes in burst-induced long-term synaptic modification. *J Neurophysiol* 95:1620–1629.
21. Clopath C, Gerstner W (2010) Voltage and spike timing interact in STDP - A unified model. *Front Synaptic Neurosci* 2:25, 10.3389/fnsyn.2010.00025.
22. Montani F, et al. (2009) The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. *Philos Trans R Soc A* 367:3297–3310.
23. Ohiorhenuan IE, et al. (2010) Sparse coding and high-order correlations in fine-scale cortical networks. *Nature* 466(7306):617–621.
24. Pillow JW, et al. (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454(7207):995–999.
25. Kempter R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. *Phys Rev E* 59:4498–4514.
26. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.
27. White LE, Fitzpatrick D (2007) Vision and cortical map development. *Neuron* 56:327–338.
28. Song S, Abbott LF (2001) Cortical development and remapping through spike timing-dependent plasticity. *Neuron* 32:339–350.
29. Shon AP, Rao RPN, Sejnowski TJ (2004) Motion detection and prediction through spike-timing dependent plasticity. *Network Comput Neural Syst* 15:179–198.
30. Li Y, Hooser SDV, Mazurek M, White LE, Fitzpatrick D (2008) Experience with moving visual stimuli drives the early development of cortical direction selectivity. *Nature* 456(7224):952–956.
31. Rao RP, Sejnowski TJ (2003) Self-organizing neural systems based on predictive learning. *Philos Trans A Math Phys Eng Sci* 361:1149–1175.
32. Blais BS, Cooper LN, Shouval H (2000) Formation of direction selectivity in natural scene environments. *Neural Comput* 12:1057–1066.
33. De Valois RL, Yund EW, Hepler N (1982) The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res* 22:531–544.
34. Luczak A, Barthó P, Marguet SL, Buzsáki G, Harris KD (2007) Sequential structure of neocortical spontaneous activity *in vivo*. *Proc Natl Acad Sci USA* 104:347–352.
35. Olhausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a space code for natural images. *Nature* 381(6583):607–609.
36. Simoncelli EP, Olhausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
37. Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. *Neural Comput* 7:1129–1159.
38. Intrator N, Cooper LN (1992) Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Netw* 5:3–17.
39. Blais BS, Intrator N, Shouval H, Cooper L (1998) Receptive field formation in natural scene environments. Comparison of single-cell learning rules. *Neural Comput* 10:1797–1813.
40. Clopath C, Longtin A, Gerstner W (2008) An online Hebbian learning rule that performs Independent Component Analysis. *Adv Neur Inf Proc Sys* 20:312–328.
41. Toyoizumi T, Pfister JP, Aihara K, Gerstner W (2005) Generalized Bienenstock-Cooper-Munro rule for spiking neurons that maximizes information transmission. *Proc Natl Acad Sci USA* 102:5239–5244.
42. Hennequin G, Gerstner W, Pfister JP (2010) STDP in adaptive neurons gives close-to-optimal information transmission. *Front Comput Neurosci* 4:143. doi:10.3389/fncom.2010.00143.
43. Castellani GC, Quinlan EM, Cooper LN, Shouval HZ (2001) A biophysical model of bidirectional synaptic plasticity: Dependence on AMPA and NMDA receptors. *Proc Natl Acad Sci USA* 98:12772–12777.
44. Shouval HZ, Bear MF, Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc Natl Acad Sci USA* 99:10831–10836.
45. Clopath C, Büsing L, Vasilaki E, Gerstner W (2010) Connectivity reflects coding: A model of voltage-based STDP with homeostasis. *Nat Neurosci* 13:344–352.
46. Shouval H, Intrator N, Cooper LN (1997) BCM network develops orientation selectivity and ocular dominance in natural scene environment. *Vision Res* 37:3339–3342.

Supporting Information

Gjorgjieva et al. 10.1073/pnas.1105933108

SI Text

Derivation of the Correlation Functions. We considered a feedforward network as in Fig. S1*A*. Let $x(t) = [x_1(t), \ldots, x_N(t)]^{\mathrm{T}}$ denote the vector of N inputs where $x_j(t) = \sum_{t_j^f} \delta(t - t_j^f)$ is the Dirac delta spike train of neuron j at time t, where t_j^f are the spike timings. The input spike trains have instantaneous firing rates $\rho^{\mathrm{inst}}(t) = [\rho_1^{\mathrm{inst}}(t), \ldots, \rho_N^{\mathrm{inst}}(t)]^{\mathrm{T}} = \langle x(t) \rangle$. (Here the expectation is taken over the input statistics.) The pairwise input correlation matrix (second moment) between inputs k and j is defined as

$$C_{kj}(s) = \frac{1}{T}\int_0^T \langle x_k(t) x_j(t-s) \rangle \mathrm{d}t. \qquad \text{[S1]}$$

[Formally, the correlation matrix should be written as $C_{kj}(t; s)$. For notational convenience, here we omit the dependence on t.] The diagonal elements of this pairwise input correlation matrix have an atomic (or point) discontinuity at $s = 0$ because $\langle x_k^2(t) \rangle = \langle x_k(t) \rangle \delta(0)$. [Note that in discrete time with bin size δt, we trivially have $x_k^2(t) = x_k(t)\delta t^{-1}$, where $x_k(t) \in \{0, \delta t^{-1}\}$ is the k^{th} spike train at time t.] Separating this correlation matrix from the pairwise correlation matrix without atomic discontinuities, C_{kj}° (where the "o" superscript denotes the absence of such discontinuity), gives

$$C_{kj}(s) = C_{kj}^{\circ}(s) + \delta_{kj}\delta(s)\ \rho_k \qquad \text{[S2]}$$

where δ_{kj} is the Kronecker delta function $\delta_{kj} = 0$ if $k \neq j$ and $\delta_{kj} = 1$ if $k = j$, and the mean firing rate averaged over the duration of the trial T is $\rho_k = (1/T)\int_0^T \rho_k^{\mathrm{inst}}(t)\mathrm{d}t$. The third-order correlation input statistic is given by

$$U_{kjn}(s_1, s_2) = \frac{1}{T}\int_0^T \langle x_k(t) x_j(t-s_1) x_n(t-s_2) \rangle \mathrm{d}t. \qquad \text{[S3]}$$

For this third-order tensor, U_j denotes a matrix whose (k, n) element is U_{kjn}. This third-order correlation function has atomic discontinuities at $s_1 = 0$, $s_2 = 0$, and $s_1 = s_2$, and we can write

$$\begin{aligned} U_{kjn}(s_1, s_2) = U_{kjn}^{\circ}(s_1, s_2) &+ \delta_{kj}\delta_{jn}\ \delta(s_1)\delta(s_2 - s_1)\ \rho_j \\ &+ \delta_{kj}\ \delta(s_1) C_{jn}(s_2 - s_1) + \delta_{kn}\ \delta(s_2) C_{kj}(s_1) \\ &+ \delta_{jn}\ \delta(s_2 - s_1) C_{kn}(s_2), \end{aligned} \qquad \text{[S4]}$$

where $U_{kjn}^{\circ}(s_1, s_2)$ is the third-order correlation without atomic discontinuities and is equal to $\rho_k\rho_j\rho_n$ if the inputs have independent Poisson statistics.

Let $u(t)$ denote the membrane potential of the postsynaptic neuron,

$$u(t) = \sum_{k=1}^{N} w_k \int_0^{\infty} \varepsilon(r) x_k(t - r)\mathrm{d}r, \qquad \text{[S5]}$$

where $\varepsilon(r)$ denotes the excitatory postsynaptic potential (EPSP) kernel taken to be a decaying exponential, $1/\tau_e e^{-r/\tau_e}\Theta(r)$, and $\Theta(r)$ is the Heaviside step function such that $\Theta(r) = 1$ when $r > 0$ and $\Theta(r) = 0$ otherwise. Postsynaptic spikes were generated stochastically from the membrane potential (1, 2), such that the probability density of firing a spike at time t was given by a nonlinear transfer function depending on the membrane potential, $\langle y(t) \rangle = g(u(t))$. (Here the expectation was taken over the postsynaptic spike train statistics.) For additional simplicity, this function was approximated by the first-order expansion about u_0 in the rest of the calculations

$$g(u(t)) \approx g(u_0) + g'(u_0)(u(t) - u_0), \qquad \text{[S6]}$$

where the expected membrane potential averaged over the trial duration T was

$$u_0 = \frac{1}{T}\int_0^T \langle u(t) \rangle \mathrm{d}t = \sum_{k=1}^{N} w_k \rho_k. \qquad \text{[S7]}$$

We also denoted the mean postsynaptic firing rate with $\nu = g(u_0)$. We derive the pre-post correlation K_j (between one presynaptic and the postsynaptic train),

$$\begin{aligned} K_j(s) &= \frac{1}{T}\int_0^T \langle y(t) x_j(t-s) \rangle \mathrm{d}t \\ &= \frac{1}{T}\int_0^T \langle g(u(t)) x_j(t-s) \rangle \mathrm{d}t \\ &= \frac{1}{T}\int_0^T \Big[g(u_0)\rho_j(t-s) + g'(u_0)\langle u(t) x_j(t-s) \rangle \\ &\quad - g'(u_0) u_0 \rho_j(t-s) \Big] \mathrm{d}t \\ &= g(u_0)\rho_j + g'(u_0)\left(\frac{1}{T}\int_0^T \langle u(t) x_j(t-s) \rangle \mathrm{d}t - u_0\rho_j \right) \\ &= (g(u_0) - g'(u_0)u_0)\rho_j + g'(u_0) \sum_k w_k C_{kj}^{\varepsilon}(s) \\ &= g'(u_0) \sum_k w_k C_{kj}^{\varepsilon}(s) + \alpha(u_0)\rho_j, \end{aligned} \qquad \text{[S8]}$$

where $\alpha(u) = g(u) - g'(u)u$ and $C_{kj}^{\varepsilon}(s) = \int_0^{\infty} \varepsilon(s') C_{kj}(s - s')\mathrm{d}s'$. Note that $\alpha(u)$ vanishes when $g(u)$ is strictly linear ($g(u) = u$) and is independent of the weights if $g(u)$ is an affine function (i.e., a linear function plus a translation). When the transfer function is linear $g(u) = u$, then $\alpha(u_0) = 0$ and $K_j(s) = \sum_k C_{kj}^{\varepsilon}(s) w_k$. We have expressed this term in Fig. S1 *B* and *D* (red curve). Fig. S1*B* shows the pre-post correlation for independent Poisson inputs (no correlations) which only has an exponential causal component ($\Delta t > 0$), corresponding to the EPSP. The presence of input correlations in Fig. S1*D* contributes both a larger exponential causal component and a nonzero acausal component ($\Delta t < 0$) - here we have assumed exponentially-decaying input correlations symmetric about 0.

For the post–prepost correlation tensor Q, we ignore the case when the two postsynaptic spikes overlap ($s_2 = 0$) because it is accounted for in the contribution from the pair rule. Then

$$\begin{aligned} Q_j(s_1, s_2) &= \frac{1}{T}\int_0^T \langle y(t - s_2) x_j(t - s_1) y(t) \rangle \mathrm{d}t \\ &= \frac{1}{T}\int_0^T \langle g(u(t - s_2)) x_j(t - s_1) g(u(t)) \rangle \mathrm{d}t \\ &= g^2(u_0)\rho_j - (g'(u_0))^2 u_0^2 \rho_j - 2g'(u_0)(g(u_0) - g'(u_0)u_0)\ u_0\rho_j \\ &\quad + g'(u_0)(g(u_0) - g'(u_0)u_0) \\ &\quad \frac{1}{T}\int_0^T \left[\langle x_j(t - s_1) u(t) \rangle + \langle u(t - s_2) x_j(t - s_1) \rangle \right] \mathrm{d}t \\ &\quad + (g'(u_0))^2 \frac{1}{T}\int_0^T \langle u(t) u(t - s_2) x_j(t - s_1) \rangle \mathrm{d}t. \end{aligned} \qquad \text{[S9]}$$

Using the expression for $\alpha(u)$ above, we get

$$Q_j(s_1,s_2) = \alpha(u_0)g'(u_0)\left(\frac{1}{T}\int_0^T \langle x_j(t-s_1)u(t)\rangle \mathrm{d}t + \frac{1}{T}\int_0^T \langle u(t-s_2)x_j(t-s_1)\rangle \mathrm{d}t\right) + (g'(u_0))^2\left(\frac{1}{T}\int_0^T \langle u(t)u(t-s_2)x_j(t-s_1)\rangle \mathrm{d}t\right) + \alpha^2(u_0)\rho_j. \quad \textbf{[S10]}$$

Now using Eq. **S7**, we get

$$Q_j(s_1,s_2) = \alpha(u_0)g'(u_0)\sum_k w_k\left(C^{\varepsilon}_{kj}(s_1) + C^{\varepsilon}_{kj}(s_1-s_2)\right) + (g'(u_0))^2\sum_{k,n} w_k w_n U^{\varepsilon}_{kjn}(s_1,s_2) + \alpha^2(u_0)\rho_j, \quad \textbf{[S11]}$$

where $U^{\varepsilon}_j(s_1,s_2) = \int_0^\infty\int_0^\infty \varepsilon(r)\varepsilon(q)U_j(s_1-r,s_2-r+q)\mathrm{d}r\,\mathrm{d}q$. When the transfer function is linear $g(u) = u$, then $\alpha(u_0) = 0$; the pre-post-post correlation Q_j is simply a function of the (convolved) third-order input correlation tensor: $Q_j(s_1,s_2) = \sum_k\sum_n w_k w_n U^{\varepsilon}_{kjn}(s_1,s_2)$. This post-pre-post correlation is illustrated in Fig. S1*C* (*Right*) for independent Poisson inputs; due to the absence of correlations, the only nonzero component is for triplets in which the presynaptic spike occurs before the two postsynaptic spikes ($\Delta t_1 > \Delta t_2 > 0$). For correlated inputs with symmetric exponentially-decaying correlations, the component corresponding to pre-post-post triplet has a larger amplitude (Fig. S1*E*, *Right*), but here post-pre-post triplets ($\Delta t_2 > \Delta t_1 > 0$) also contribute to the correlation function.

Weight Dynamics. The triplet spike-timing–dependent plasticity (STDP) model can be written in the following differential form. Let $\bar{x}_j$ be a low-pass filtered version of the jth presynaptic spike train x_j with time constant τ_+ given by

$$\dot{\bar{x}}_j = -\frac{\bar{x}_j}{\tau_+} + x_j, \quad j = 1,\ldots,N. \quad \textbf{[S12]}$$

Similarly, let $\bar{y}_1$ and $\bar{y}_2$ be two different low-pass filtered versions of the postsynaptic spike train y with time constants τ_- and τ_y, respectively:

$$\dot{\bar{y}}_1 = -\frac{\bar{y}_1}{\tau_-} + y, \quad \textbf{[S13]}$$

$$\dot{\bar{y}}_2 = -\frac{\bar{y}_2}{\tau_y} + y. \quad \textbf{[S14]}$$

The minimal triplet STDP model can be written as

$$\dot{w}_j = -A_2^- x_j \bar{y}_1 + A_3^+ y \bar{x}_j \bar{y}_2. \quad \textbf{[S15]}$$

In the previous section we assumed that the weights w_j were fixed. Here, we allowed the weights to be dynamic, but with a small learning rate such that the results in the previous section are still valid (1, 2). Consistent with the approach of ref. 1, we expressed the weight change as a Volterra expansion of both the pre- and postsynaptic spike trains. Setting the first order terms to zero (because they correspond to non-Hebbian dynamics), we assumed that synaptic plasticity depends on second- and third-order terms only, i.e. pairs of spikes (1 pre and 1 post) and triplets of spikes (1 pre and 2 post)

$$\dot{w}_j = y(t)\int_0^\infty W_2(s)x_j(t-s)\mathrm{d}s + x_j(t)\int_0^\infty W_2(-s)y(t-s)\mathrm{d}s + y(t)\int_0^\infty\int_0^\infty W_3(s_1,s_2)x_j(t-s_1)y(t-s_2)\mathrm{d}s_1\mathrm{d}s_2$$

where W_2 is pair-based STDP and W_3 is triplet STDP (equations and parameters listed in the main text). [Note that in ref. 3, the wrong number of pairs is assumed in the STDP experiments (60 instead of 75) and therefore the fitted amplitude parameters were overestimated by approximately 10 %. For the sake of consistency with ref. 3 we kept the same parameters.] Assuming slow learning relative to the neuronal dynamics, and replacing the weights by their expectation averaged over a time period T, we have

$$\dot{w}_j = \int_{-\infty}^\infty W_2(s)K_j(s)\mathrm{d}s + \int_{-\infty}^\infty\int_{-\infty}^\infty W_3(s_1,s_2)Q_j(s_1,s_2)\mathrm{d}s_1\mathrm{d}s_2. \quad \textbf{[S16]}$$

Under the assumption of a linear approximation of the transfer function given in Eq. **S6**, and by inserting Eqs. **S8** and **S11** into Eq. **S16**, we find

$$\dot{w} = g'(u_0)(A+D)w + g'^2(u_0)\sum_{j=1}^N (w^{\mathrm{T}}B_j w)e_j + \kappa(u_0)\rho, \quad \textbf{[S17]}$$

where e_j is a vector of zeros with a 1 at the j^{th} component and

$$A_{jk} = \int_{-\infty}^\infty W_2(s)C^{\varepsilon}_{kj}(s)\mathrm{d}s, \quad \textbf{[S18]}$$

$$D_{jk} = \alpha(u_0)\int_{-\infty}^\infty\int_{-\infty}^\infty W_3(s_1,s_2)\left(C^{\varepsilon}_{kj}(s_1) + C^{\varepsilon}_{kj}(s_1-s_2)\right)\mathrm{d}s_1\mathrm{d}s_2, \quad \textbf{[S19]}$$

$$B_{kjn} = \int_{-\infty}^\infty\int_{-\infty}^\infty W_3(s_1,s_2)U^{\varepsilon}_{kjn}(s_1,s_2)\mathrm{d}s_1\mathrm{d}s_2, \quad \textbf{[S20]}$$

$$\kappa(u_0) = \alpha(u_0)\overline{W}_2 + \alpha^2(u_0)\overline{W}_3, \quad \textbf{[S21]}$$

with $\overline{W}_2 = \int_{-\infty}^\infty W_2(s)\mathrm{d}s$ and $\overline{W}_3 = \int_0^\infty\int_0^\infty W_3(s_1,s_2)\mathrm{d}s_1\mathrm{d}s_2$. Note that $g'(u_0)$ and $\kappa(u_0)$ depend on the weights w (through u_0), and therefore we cannot analyze Eq. **S17** in the same way as if we assume a strictly linear transfer function. In that case, with a linear transfer function, $g(u) = u$, we have $g'(u_0) = 1$, $\alpha(u_0) = 0$, and hence $\kappa(u_0) = 0$. Then we get:

$$\dot{w} = Aw + \sum_{j=1}^N (w^{\mathrm{T}}B_j w)e_j. \quad \textbf{[S22]}$$

We can also write this equation so that it resembles the Bienenstock–Cooper–Munro (BCM) equation for weight evolution (see also Eq. 6 in the main text),

$$\dot{w} = \underbrace{\phi(\nu)\rho}_{\text{BCM term}} + \Delta A w + \sum_{j=1}^N (w^{\mathrm{T}}\Delta B_j w)e_j, \quad \textbf{[S23]}$$

where $\phi(\nu) = \overline{W}_2\nu + \overline{W}_3\nu^2$ corresponds to the BCM term. The new covariance terms ΔA and ΔB are defined analogously to A and B given the input covariances,

$$\Delta A_{jk} = \int_{-\infty}^\infty W_2(s)\Delta C^{\varepsilon}_{kj}(s)\mathrm{d}s, \quad \textbf{[S24]}$$

where $\Delta C_{kj}(s) = C_{kj}(s) - \rho_k\rho_j$ is the input pairwise covariance matrix, and

$$\Delta B_j = \int_0^\infty\int_0^\infty W_3(s_1,s_2)\Delta U^{\varepsilon}_j(s_1,s_2)\mathrm{d}s_1\mathrm{d}s_2,$$

where $\Delta U_{kjn}(s_1, s_2) = U_{kjn}(s_1, s_2) - \rho_k\rho_j\rho_n$ is the input triplet covariance tensor.

M patterns were presented to the network, pattern i with probability p_i, with a mean firing rate of $\rho^{(i)}$ and pairwise and triplet correlations terms $A^{(i)}$ and $B^{(i)}$, respectively. To match the triplet rule to the BCM model, we set $A_2^- \to A_2^-\,\overline{\nu}/\rho_0^p$, where the expectation of the p^{th} power of the postsynaptic firing rate can be expressed as $\overline{\nu} = \sum_{i=1}^{M} p_i(\nu^{(i)})^p$. This quantity was approximated by low-pass filtering the p^{th} power of the instantaneous postsynaptic firing rate $\nu(t) = g(u(t))$ with a time constant which has to be larger than M times the frequency of pattern presentation, i.e. $r \simeq \overline{\nu}$ where $\tau_r\dot{r} = -r + \nu^p$ with a time constant of $\tau_r = 5$s. For all the calculations in this paper we took $p = 2$. Using the minimal triplet model of ref. 3 where $A_2^+ = 0$, $A^{(i)}$ contains only depression effects from the pair STDP rule ($A_2^- \neq 0$) such that the weight equation becomes

$$\dot{w} = \sum_{i=1}^{M} p_i\left(\phi\left(\nu^{(i)}, \overline{\nu}\right)\rho^{(i)} + \Delta A^{(i)}(\overline{\nu})w + \sum_{j=1}^{N}\left(w^{\mathrm{T}}\Delta B_j^{(i)}w\right)e_j\right). \quad \text{[S25]}$$

Here, $\nu^{(i)} = w^{\mathrm{T}}\rho^{(i)}$ denotes the average postsynaptic firing rate elicited by pattern i.

Selectivity with N Orthogonal Rate-Based Input Patterns. In the case of M=N orthogonal Poisson inputs we show that the maximally selective fixed points of the weight dynamics driven by the triplet learning rule are stable. Because we consider patterns with positive components (rates), the Poisson rate of the j^{th} component of pattern i is given by $\rho_j^{(i)} = \delta_{ij}|\rho^{(i)}|$ (i.e., pattern i has all zero components but the i^{th} component).

Because of the Poisson assumption, we have $\Delta C_{kj}(s) = \delta_{kj}\delta(s)\rho_k$ and hence $\Delta C_{kj}^{\varepsilon}(s) = \delta_{kj}\,\rho_k\varepsilon(s) = 0$ for $s < 0$. Because we use the minimal triplet model (for which $A_2^+ = 0$), $W_2(s) = 0$ for $s \geq 0$. As a consequence, ΔA vanishes because it is obtained by integrating the product $W_2(s)\cdot\Delta C_{kj}^{\varepsilon}(s)$ (Eq. **S24**). By using Eqs. **S4** and **S11**, $\Delta B_{kjn}^{(i)}$ can be written as

$$\Delta B_{kjn}^{(i)} = \left(b_1\delta_{kj}\rho_n^{(i)} + b_2\delta_{kn}\rho_k^{(i)} + b_3\delta_{kj}\delta_{kn}\right)\rho_j^{(i)}, \quad \text{[S26]}$$

where

$$b_1 = \int_0^\infty\int_0^\infty W_3(s_1, s_2)(\varepsilon(s_1) + \varepsilon(s_2 - s_1))\mathrm{d}s_1\,\mathrm{d}s_2, \quad \text{[S27]}$$

$$b_2 = \int_0^\infty\int_0^\infty W_3(s_1, s_2)\int_0^\infty \varepsilon(r)\varepsilon(r - s_2)\mathrm{d}r\,\mathrm{d}s_1\,\mathrm{d}s_2, \quad \text{[S28]}$$

and

$$b_3 = \int_0^\infty\int_0^\infty W_3(s_1, s_2)\varepsilon(s_1)\varepsilon(s_2 - s_1)\mathrm{d}s_1\mathrm{d}s_2. \quad \text{[S29]}$$

In the presence of $M = N$ Poisson patterns, we can write the expected weight dynamics as

$$\dot{w} = \sum_{i=1}^{M} p_i\left(\phi\left(\nu^{(i)}, \overline{\nu}\right)\mathbf{1} + \Lambda^{(i)}\right)\rho^{(i)}, \quad \text{[S30]}$$

where

$$\phi\left(\nu^{(i)}, \overline{\nu}\right) = \overline{W}_2\frac{\overline{\nu}}{\rho_0^2}\nu^{(i)} + \overline{W}_3\left(\nu^{(i)}\right)^2, \quad \text{[S31]}$$

$\mathbf{1}$ is the $N \times N$ identity matrix, and $\Lambda^{(i)}$ is a diagonal matrix where the j^{th} diagonal element is

$$\Lambda_{jj}^{(i)} = \left(b_1 w_j\nu^{(i)} + b_2\nu_2^{(i)} + b_3 w_j^2\right)\rho_j^{(i)} \quad \text{[S32]}$$

with $\nu_2^{(i)} = \sum_k w_k^2\rho_k^{(i)}$. Recall that $\overline{\nu} = \sum_i p_i(\nu^{(i)})^2$ with $\nu^{(i)} = w^{\mathrm{T}}\rho(i)$. Because of the orthogonality assumption, the condition $\dot{w} = 0$ implies that $\phi(\nu^{(i)}, \overline{\nu}^{(i)}) + \Lambda_{ii}^{(i)} = 0$, for all $i = 1, \ldots, N$ and therefore

$$w_i^2 F\left(\rho_i^{(i)}\right) - G\rho_i^{(i)}w_i\overline{\nu} = 0, \quad \forall i = 1, \ldots, N, \quad \text{[S33]}$$

where

$$G = -\frac{\overline{W}_2}{\rho_0^2} > 0 \text{ and } F\left(\rho_i^{(i)}\right) = \overline{W}_3\left(\rho_i^{(i)}\right)^2 + (b_1 + b_2)\rho_i^{(i)} + b_3. \quad \text{[S34]}$$

Each fixed point of Eq. **S30** must satisfy the N conditions from Eq. **S33**. Each condition has two solutions: either $w_i^* = 0$ or $w_i^* = G\rho_i^{(i)}/F(\rho_i^{(i)})$. As a consequence, there are 2^N fixed points, which is consistent with the BCM theory.

It remains to be shown that the maximally selective fixed points are stable. The n^{th} fixed point is given by

$$w^{\star(n)} = (0, \ldots, 0,\ w_n^*,\ 0, \ldots, 0)^{\mathrm{T}},$$

where

$$w_n^* = \frac{F\left(\rho_n^{(n)}\right)}{Gp_n\left(\rho_n^{(n)}\right)^3} \text{ takes the } n^{\text{th}} \text{ position.}$$

To demonstrate that this fixed point is stable, we have to show that the eigenvalues of the Jacobian of Eq. **S30** are negative when evaluated at $w^{*(n)}$. We find that this Jacobian matrix is given by

$$J_{ij}\left(w^{\star(n)}\right) = -\delta_{ij}p_i G\left(\rho_i^{(i)}\right)^2\overline{\nu}^{(n)}, \quad \text{[S35]}$$

where $\overline{\nu}^{(n)} = p_n w_n^2(\rho_n^{(n)})^2$. Because this matrix is diagonal with all diagonal elements being negative, we conclude that all of the maximally selective fixed points are stable.

At this fixed point, the sliding threshold takes the value

$$\theta = w^{\star(n)}\rho_n^{(n)} = \theta_0(\overline{\nu})\left(1 + \left(\tau_1\rho_n^{(n)}\right)^{-1} + \left(\tau_2\rho_n^{(n)}\right)^{-2}\right)^{-1}, \quad \text{[S36]}$$

where $\theta_0(\overline{\nu}) = A_2^-\,\tau_-\,\overline{\nu}/(A_3^+\tau_+\tau_y\rho_0^p)$ is the sliding threshold of the BCM term only when the correlation terms can be neglected (i.e., $\Delta A = 0$ and $\Delta B = 0$). The new timescales are

$$\tau_1 = \overline{W}_3/(b_1 + b_2) \text{ and } \tau_2 = \sqrt{\overline{W}_3/b_3}. \quad \text{[S37]}$$

One can calculate these values and obtain

$$\frac{b_1 + b_2}{\overline{W}_3} = \frac{1}{\tau_m + \tau_+} + \frac{\tau_+}{(\tau_m + \tau_+)(\tau_+ + \tau_y)} + \frac{1}{2(\tau_m + \tau_y)} \quad \text{[S38]}$$

and

$$\frac{b_3}{\overline{W}_3}=\frac{\tau_+}{(\tau_m+2\tau_+)(\tau_+\tau_y+\tau_m\tau_+ +\tau_m\tau_y)}, \quad \textbf{[S39]}$$

where $\overline{W}_3=A_3^+\tau_+\tau_y$. Note that in the limit where $\tau_m \simeq \tau_+ \simeq \tau_-$ and $\tau_y \gg \tau_m$, we have $\tau_1 \approx 2\tau_m$ and $\tau_2 \approx \sqrt{6\tau_m\tau_y}$.

Selectivity in a 2D System with Correlation-Based Input Patterns. We consider a 2D system where two patterns are presented to the feedforward network, each consisting of two pools of inputs: let $j = 1, \ldots, N/2$ denote the inputs from the first pool and $j = N/2 + 1, \ldots, N$ denote the inputs from the second pool. Let $\tilde{w}_j$ denote the weights in the network and $\tilde{A}^{(i)}$ and $\tilde{B}_k^{(i)}$ the convolved pairwise and third-order correlations of the inputs in each pool for pattern i, respectively, with the pair-based and the triplet STDP rules (Eqs. **S18** and **S20**). We further imposed a lower bound on the weights, $w \geq w_{\min}=0$, which in the case of orthogonal rate-based patterns was automatically satisfied.

Under the assumption that the weights in each pool evolve together,

$$\tilde{w}_j \simeq w_1, j=1,\ldots,\frac{N}{2}, \quad \textbf{[S40]}$$

$$\tilde{w}_j \simeq w_2, j=\frac{N}{2}+1,\ldots,N. \quad \textbf{[S41]}$$

Summing the weights in each pool of inputs for pattern $i = 1, 2$ gives

$$\sum_{j=1}^{N/2}\dot{\tilde{w}}_j=\frac{N}{2}\dot{w}_1=\sum_{j=1}^{N/2}\left(\sum_{k=1}^{N}\tilde{A}_{jk}^{(i)}\tilde{w}_k\frac{\overline{\nu}}{\rho_0^2}+\sum_{m=1}^{N}\sum_{n=1}^{N}\tilde{w}_m\tilde{w}_n\left(\tilde{B}_j^{(i)}\right)_{mn}\right) \quad \textbf{[S42]}$$

$$\sum_{j=N/2+1}^{N}\dot{\tilde{w}}_j=\frac{N}{2}\dot{w}_2=\sum_{j=\frac{N}{2}+1}^{N}\left(\sum_{k=1}^{N}\tilde{A}_{jk}^{(i)}\tilde{w}_k\frac{\overline{\nu}}{\rho_0^2}+\sum_{m=1}^{N}\sum_{n=1}^{N}\tilde{w}_m\tilde{w}_n\left(\tilde{B}_j^{(i)}\right)_{mn}\right). \quad \textbf{[S43]}$$

The contributions from the pairwise correlations are given by

$$A_{11}^{(i)}=\frac{2}{N}\sum_{j=1}^{N/2}\sum_{k=1}^{N/2}\tilde{A}_{jk}^{(i)},\ A_{12}^{(i)}=\frac{2}{N}\sum_{j=1}^{N/2}\sum_{k=\frac{N}{2}+1}^{N}\tilde{A}_{jk}^{(i)},$$

$$A_{21}^{(i)}=\frac{2}{N}\sum_{j=\frac{N}{2}+1}^{N}\sum_{k=1}^{N/2}\tilde{A}_{jk}^{(i)},\ A_{22}^{(i)}=\frac{2}{N}\sum_{j=N/2+1}^{N}\sum_{k=\frac{N}{2}+1}^{N}\tilde{A}_{jk}^{(i)}$$

for pattern $i = 1, 2$, where the average postsynaptic rate over all of the patterns is

$$\overline{\nu}=p_1\left(\rho_1^{(1)}w_1+\rho_2^{(1)}w_2\right)^2+p_2\left(\rho_1^{(2)}w_1+\rho_2^{(2)}w_2\right)^2 \quad \textbf{[S44]}$$

with

$$\rho_1^{(i)}=\left(\sum_{j=1}^{N/2}\tilde{\rho}_j^{(i)}\right),\ \rho_2^{(i)}=\left(\sum_{j=N/2+1}^{N}\tilde{\rho}_j^{(i)}\right).$$

The terms involving the third-order correlations are

$$\left(B_1^{(i)}\right)_{11}=\frac{2}{N}\sum_{j=1}^{N/2}\sum_{m=1}^{N/2}\sum_{n=1}^{N/2}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_1^{(i)}\right)_{12}=\frac{2}{N}\sum_{j=1}^{N/2}\sum_{m=1}^{N/2}\sum_{n=N/2+1}^{N}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_1^{(i)}\right)_{21}=\frac{2}{N}\sum_{j=1}^{N/2}\sum_{m=\frac{N}{2}+1}^{N}\sum_{n=1}^{N/2}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_1^{(i)}\right)_{22}=\frac{2}{N}\sum_{j=1}^{N/2}\sum_{m=\frac{N}{2}+1}^{N}\sum_{n=\frac{N}{2}+1}^{N}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_2^{(i)}\right)_{11}=\frac{2}{N}\sum_{j=\frac{N}{2}+1}^{N}\sum_{m=1}^{N/2}\sum_{n=1}^{N/2}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_2^{(i)}\right)_{12}=\frac{2}{N}\sum_{j=\frac{N}{2}+1}^{N}\sum_{m=1}^{N/2}\sum_{n=\frac{N}{2}+1}^{N}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_2^{(i)}\right)_{21}=\frac{2}{N}\sum_{j=\frac{N}{2}+1}^{N}\sum_{m=\frac{N}{2}+1}^{N}\sum_{n=1}^{N/2}\left(\tilde{B}_j\right)_{mn},$$

$$\left(B_2^{(i)}\right)_{22}=\frac{2}{N}\sum_{j=\frac{N}{2}+1}^{N}\sum_{m=\frac{N}{2}+1}^{N}\sum_{n=\frac{N}{2}+1}^{N}\left(\tilde{B}_j\right)_{mn},$$

finally giving an equation for the weight dynamics of the two pools of inputs, corresponding to $M = 2$ in Eq. **S25**

$$\dot{w}=\sum_{i=1}^{2}p_i\left(A^{(i)}(\overline{\nu})w+\sum_{k=1}^{2}w^{\mathrm{T}}B_k^{(i)}we_k\right). \quad \textbf{[S45]}$$

Next, we derive the fixed points of the 2D system of Eq. **S45** using $\overline{\nu}=\sum_{i=1}^{2}p_i(w^{\mathrm{T}}\rho^{(i)})^2$. To obtain the general fixed points of this equation, we have to simultaneously solve $\dot{w}_1 = 0$ and $\dot{w}_2 = 0$. This process amounts to solving two cubic equations for which we do not get nice analytical expressions. Fortunately, the fixed points of interest [the ones associated with maximal selectivity, i.e., $(w_1^*, 0)$ and $(0, w_2^*)$ with $w_1^*, w_2^* > 0$] are simpler to express. To find a fixed point on the w_1 axis, we solve $\dot{w}_1 = 0$ at $w_1 = w_1^*$ and $w_2 = 0$,

$$0=p_1\left[A_{11}^{(1)}w_1\frac{p_1\left(w^{\mathrm{T}}\rho^{(1)}\right)^2+p_2\left(w^{\mathrm{T}}\rho^{(2)}\right)^2}{\rho_0^2}+\left(B_1^{(1)}\right)_{11}w_1^2\right]+p_2\left[A_{11}^{(2)}w_1\frac{p_1\left(w^{\mathrm{T}}\rho^{(1)}\right)^2+p_2\left(w^{\mathrm{T}}\rho^{(2)}\right)^2}{\rho_0^2}+\left(B_1^{(2)}\right)_{11}w_1^2\right], \quad \textbf{[S46]}$$

which gives the linear equation

$$\left(p_1A_{11}^{(1)}+p_2A_{11}^{(2)}\right)\frac{p_1\left(\rho_1^{(1)}\right)^2+p_2\left(\rho_1^{(2)}\right)^2}{\rho_0^2}w_1^*+\left(p_1\left(B_1^{(1)}\right)_{11}+p_2\left(B_1^{(2)}\right)_{11}\right)=0. \quad \textbf{[S47]}$$

This equation has the solution

$$w_1^*=-\frac{p_1\left(B_1^{(1)}\right)_{11}+p_2\left(B_1^{(2)}\right)_{11}}{\rho_0^{-2}\left(p_1A_{11}^{(1)}+p_2A_{11}^{(2)}\right)p_1\left(\rho_1^{(1)}\right)^2+p_2\left(\rho_1^{(2)}\right)^2}. \quad \textbf{[S48]}$$

The expression for w_2^* can be similarly obtained. Instead of examining the Jacobian of the system at the fixed points to obtain their stability, because of the nonlinearity due to the lower

bound on the weights, we require that the following two conditions be satisfied for stability (where the system is denoted by $\dot{w} = F(w)$):

$$\left.\frac{\partial F_1(w)}{\partial w_1}\right|_{w=w^*} < 0, \qquad \text{[S49]}$$

$$F_2(w^*) < 0. \qquad \text{[S50]}$$

The first condition becomes

$$3\left(p_1 A_{11}^{(1)} + p_2 A_{11}^{(2)}\right)\frac{p_1\left(\rho_1^{(1)}\right)^2 + p_2\left(\rho_1^{(2)}\right)^2}{\rho_0^2} w_1^* + 2\left(p_1\left(B_1^{(1)}\right)_{11} + p_2\left(B_1^{(2)}\right)_{11}\right) < 0. \qquad \text{[S51]}$$

If we use the expression for w_1^* from Eq. **S48**, then the condition reduces to

$$p_1\left(B_1^{(1)}\right)_{11} + p_2\left(B_1^{(2)}\right)_{11} > 0, \qquad \text{[S52]}$$

which is always true because the correlation terms convolved with the triplet rule in $B_k^{(i)}$ are always positive.

The second condition becomes

$$\left(p_1 A_{21}^{(1)} + p_2 A_{21}^{(2)}\right)\frac{p_1\left(\rho_1^{(1)}\right)^2 + p_2\left(\rho_1^{(2)}\right)^2}{\rho_0^2} w_1^* + \left(p_1\left(B_2^{(1)}\right)_{11} + p_2\left(B_2^{(2)}\right)_{11}\right) < 0. \qquad \text{[S53]}$$

If we use the expression for w_1^* from Eq. **S48**, the condition reduces to

$$\frac{p_1 A_{21}^{(1)} + p_2 A_{21}^{(2)}}{p_1 A_{11}^{(1)} + p_2 A_{11}^{(2)}} > \frac{p_1\left(B_2^{(1)}\right)_{11} + p_2\left(B_2^{(2)}\right)_{11}}{p_1\left(B_1^{(1)}\right)_{11} + p_2\left(B_1^{(2)}\right)_{11}}. \qquad \text{[S54]}$$

Similarly, the fixed point on the w_2 axis is stable if the following condition holds:

$$\frac{p_1 A_{12}^{(1)} + p_2 A_{12}^{(2)}}{p_1 A_{22}^{(1)} + p_2 A_{22}^{(2)}} > \frac{p_1\left(B_1^{(1)}\right)_{22} + p_2\left(B_1^{(2)}\right)_{22}}{p_1\left(B_2^{(1)}\right)_{11} + p_2\left(B_2^{(2)}\right)_{22}}. \qquad \text{[S55]}$$

Numerical evaluations of these two conditions for a large variety of firing rates and pairwise and third-order correlations have demonstrated that the conditions always hold. Therefore, in the case of a 2D network, the system always results in selectivity. As we show in the main text, this is not always the case for a general N-dimensional system, where selectivity depends on the input correlation structure.

Numerical Simulations with Multiple Patterns. For the simulations with rate-based patterns in Fig. 2, the inputs within each pattern were given independent Poisson spike trains lacking correlations. For each of 10 patterns uniformly spaced and centered at inputs 5, 15, … 95, a Gaussian rate profile was used with a background firing rate of $r_{\min}$ and a peak firing rate of $r_{\max}$, and we explored three ratios in Fig. 2*B*: $r_{\min}/r_{\max} = \{0/55, 5/55, 10/55\}$. The SD of the Gaussian was also varied in Fig. 2, $\sigma = \{5.0, 7.5, 10.0, 12.5, 15.0\}$, but the Gaussian profile was normalized such that it generated the same postsynaptic firing rate for each value of σ and $r_{\min}/r_{\max}$. The postsynaptic neuron was linear $g(u) = 10u$ and the target postsynaptic firing rate was set to $\rho_0 = 8.5$ Hz.

For the simulations with correlation-based patterns in Fig. 3, each of the 100 inputs had the same firing rate of 10 Hz. In each pattern, 90 inputs were given independent Poisson spikes, and 10 inputs had uniform correlations between any pair and triplet of inputs. For the spatial correlations in Fig. 3*B*, each pair and triplet of inputs shared 90% identical spikes. For the spatiotemporal correlations in Fig. 3*C*, half of the 90% shared spikes for each pair and triplet of inputs were shifted by an exponential random distribution with a mean of 5 ms resulting in symmetric, exponentially-decaying correlations with a timescale of 5 ms. For simplicity, we assumed uniform correlations for all input pairs and triplets. The postsynaptic neuron was linear $g(u) = 10\ u$ and the target postsynaptic firing rate was set to $\rho_0 = 10.5$ Hz.

For the simulations in Fig. 4 *A–C*, there were two patterns consisting of three inputs each, with the same firing rates and pairwise correlations, but with different third-order correlations. The postsynaptic neuron was linear $g(u) = 50\ u$ and the target postsynaptic firing rate was 20 Hz. For Fig. 4*D*, a network of two groups of five neurons each was simulated so that the two input groups had the same lower-order correlations (for example, same firing rates, pairwise and third-order correlations), but differed in the presence or absence of higher-order correlations in each group (corresponding example, fourth- and fifth-order correlations). The correlated spike trains in Fig. 3 and 4 were generated using the mixture method described in ref. 4 (see next section).

In Figs. 2–5, a new randomly-chosen pattern was presented to the network every 200 ms. Pre- and postsynaptic spikes were simulated stochastically given the respective firing rates. The initial weights were set to 1 and hard bounds were set between 0 and 3 (in Fig. 4 because of the small number of inputs the bounds were five times larger.) Postsynaptic activity was low-pass filtered with a time constant of 5 seconds. A_2^+ and A_3^- were reduced by a factor of 10 compared to the parameters in ref. 3 to give smooth evolution of the weights, but this did not affect the results.

Simulations of Correlated Spike Trains. Correlated spike trains in Figs. 3 and 4 in the main text were simulated by using the mixture method following Brette (3). The general method is illustrated in Fig. S2, where N target spike trains (gray) are generated from M source trains (blue), with specifics adjusted to fit our simulations. Fig. S2*A* illustrates the method for generating correlated spike trains with pairwise correlations only, and Fig. S2*B* shows the method for generating correlated spike trains with pairwise and third-order correlations. Both methods were used in Fig. 4 *B* and *C* for simulating 6 spike trains with $N = 3$ for each method (in Fig. 4*D* we used 10 spike trains with $N = 5$) and $M = N$ in Fig. S2*A* and $M = N + 1$ in Fig. S2*B*, whereas the method in Fig. S2*B* was used in Fig. 3 with $M = N = 100$. The source trains were independent Poisson processes with rate R. Spikes from source train m were copied into target train m with probability p_1. Furthermore, to generate instantaneous correlations among the target spike trains, spikes from the common source spike trains were copied into the the target trains. In Fig. S2*B*, spikes were copied from the single common spike train, generating correlations of higher order in the target trains. In Fig. S2*A*, each pair of target trains received spikes from a single common course train (probability of copying p_2), different for each pair of target trains, thus generating only pairwise correlations among the target trains, but no higher-order correlations.

In both cases, the firing rate of the k^{th} target train is given by (3)

$$\rho_k = (p_1 + p_2)\, R. \qquad \textbf{[S56]}$$

The pairwise covariance matrix without the atomic discontinuities ΔC° can be defined analogously to the corresponding correlation matrix in Eq. **S2**:

$$\Delta C_{kj}(s) = \Delta C^\circ_{kj}(s) + \delta_{kj}\delta(s)\,\rho_j. \qquad \textbf{[S57]}$$

Then the instantaneous pairwise covariance is (see ref. 3 for details)

$$\Delta C^\circ_{kj}(s) = \gamma_{kj}\delta(s). \qquad \textbf{[S58]}$$

This covariance had the same magnitude for any pair of inputs k, j: For Fig. S2*A*, $\gamma_{kj} = p_1 p_2 R$ and for Fig. S2*B*, $\gamma_{kj} = p_2^2 R$.

The expression for the third-order covariance is more complicated [it can be related to U_{kjn} by $\Delta V_{kjn}(s_1, s_2) = U_{kjn}(s_1, s_2) - \rho_k \Delta C_{jn}(s_2 - s_1) - \rho_j \Delta C_{kn}(s_2) - \rho_n \Delta C_{kj}(s_1) - \rho_k \rho_j \rho_n$]:

$$\Delta V_{kjn}(s_1, s_2) = \frac{1}{T}\int_0^T \Big\langle (x_k(t) - \rho_k)\Big(x_j(t - s_1) - \rho_j\Big)(x_n(t - s_2) - \rho_n)\Big\rangle \mathrm{d}t. \qquad \textbf{[S59]}$$

After separating the atomic discontinuities

$$\begin{aligned}\Delta V_{kjn}(s_1, s_2) =& \Delta V^\circ_{kjn}(s_1, s_2) + \delta_{kj}\delta_{jn}\delta(s_1)\delta(s_2 - s_1)\rho_j \\ &+ \delta_{kj}\delta(s_1)\Delta C^\circ_{jn}(s_2 - s_1) + \delta_{kn}\delta(s_2)\Delta C^\circ_{kj}(s_1) \\ &+ \delta_{jn}\delta(s_2 - s_1)\Delta C^\circ_{kn}(s_2),\end{aligned} \qquad \textbf{[S60]}$$

we can specify the instantaneous third-order covariance by (see ref. 3 for details)

$$\Delta V^\circ_{kjn}(s) = \lambda_{kjn}\delta(s_1)\delta(s_2). \qquad \textbf{[S61]}$$

This covariance had the same magnitude for any triplet of inputs k, j, n: For Fig. S2*A*, $\lambda_{kjn} = 0$ and for Fig. S2*B*, $\lambda_{kjn} = p_2^3 R$.

In Fig. 3, we used the method in Fig. S2*B* with $N = 100$, $R = 9.09$ Hz, and $p_1 = 0.1$ and $p_2 = 1.0$. In Fig. 4 *B* and *C*, we used both methods in Fig. S2 with $N = 3$, $R = 5$ Hz, and $p_1 = p_2 = 1.0$.

To show that in Fig. 4*C* of the main text these methods do indeed generate the same pairwise correlations (but different third-order correlations), we computed the mean ± SEM of the peak correlation coefficients for 200 simulation runs in each case (Fig. S3).

In Fig. 4*D*, we extended the methods described in Fig. S2 to generate correlations of higher than third order and considered $N = 5$ neurons per group. To illustrate the procedure, we describe how we generated correlated spikes to distinguish between correlations of fifth order. First we generated a single common source train (with rate R) from which spikes were copied into the target trains of the inputs in group 1 with probability p. This process generated the following statistics for the inputs in group 1: rates pR, pairwise correlations p^2R, third-order correlations p^3R, fourth-order correlations p^4R, and fifth-order correlations p^5R. Then we generated $\binom{5}{4} = 5$ source trains (with rates R), each copying spikes with probability p into the 4-tuplets of target trains (1, 2, 3, 4), (1, 2, 3, 5), (2, 3, 4, 5), (1, 2, 4, 5), (1, 3, 4, 5) in group 2. This process generated fourth-order correlations p^4R in group 2 (but no fifth-order correlations because no target trains received spikes from the same source train). Thus, the fifth-order correlations were different for groups 1 (p^5R) and 2 (0), but the fourth-order correlations were the same (p^4R). However, the lower-order correlations also differed. The third-order correlations in group 1 (p^3R) were lower than the third-order correlations in group 2 ($2p^3R$). Therefore, we generated $\binom{5}{3} = 10$ more source trains for the inputs in group 1, which copied spikes into the triplets of target trains (1, 2, 3), (1, 2, 4), (1, 2, 5), (1, 3, 4), (1, 3, 5), (1, 4, 5), (2, 3, 4), (2, 3, 5), (2, 4, 5), (3, 4, 5) in group 1. Now even the third-order correlations in both groups were matched to $2p^3R$. However, the pairwise correlations in group 1 then became $4p^2R$, whereas in group 2 they were $3p^2R$. To match these pairwise correlations, we generated $\binom{5}{2} = 10$ source trains in group 2, copying spikes into the pairs of target trains (1, 2), (1, 3), (1, 4), (1, 5) (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5) in group 2. Now the pairwise correlations in both groups 1 and 2 were matched to $4p^2R$. However, the firing rates for the target trains in group 1 were $7pR$, whereas for the target trains in group 2 they were $8pR$. Finally, we generated $\binom{5}{1} = 5$ more source trains in group 1 that copied spikes independently into each of the 5 target trains, giving firing rates in each group equal to $8pR$. Because the target firing rate was 10 Hz, all source trains had rates $10/(8p)$ Hz. For simplicity, we used a copying probability of $p = 1.0$, meaning that all of the spikes from the source trains were copied into the target trains. In this example, there were a total of 16 source trains generated for the 5 target trains in group 1 and 15 source trains for the 5 target trains in group 2. Analogous procedures were used for the cases with correlations of $k = 2$, 3, or 4$^{\text{th}}$ order. For the case of $k = 1$ we studied the difference in first-order correlations, i.e., firing rates, and thus we considered rates of 10 Hz in group 1 and 7 Hz in group 2.

For generating spatiotemporal correlations, the mixture method described by Brette (3) was used. The instantaneous correlated spike trains (generated as above) were shifted by independent and identically distributed random numbers from an appropriate distribution function. We used an exponential distribution with a time constant τ_c, $f(t) = (1/\tau_c)e^{-t/\tau_c}\Theta(t)$, where $\Theta(x)$ is the Heaviside function, such that

$$\Delta C^\circ_{kj}(s) = \gamma_{kj}\int_{-\infty}^{\infty} f(t)f(t - s)\mathrm{d}t = \frac{\gamma_{kj}}{2\tau_c}e^{-|s|/\tau_c} \qquad \textbf{[S62]}$$

and the third-order covariance

$$\begin{aligned}\Delta V^\circ_{kjn}(s_1, s_2) &= \lambda_{kjn}\int_{-\infty}^{\infty} f(t)f(t - s_1)f(t - s_2)\mathrm{d}t \\ &= \frac{\lambda_{kjn}}{3\tau_c^2}\begin{cases} e^{-(s_1+s_2)/\tau_c}, & s_1 \geq 0,\ s_2 \geq 0 \\ e^{(2s_1 - s_2)/\tau_c}, & s_1 < 0,\ s_2 > s_1 \\ e^{(-s_1+2s_2)/\tau_c}, & s_1 > s_2, s_2 < 0. \end{cases}\end{aligned} \qquad \textbf{[S63]}$$

Note that in all of the simulations in the main text we assumed uniform correlations for all spike pairs and triplets such that $\gamma_{kj} = \gamma$ for all pairs (k, j) and λ_{kjn} for all triplets (k, j, n).

1. Kempter R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. *Phys Rev E* 59:4498–4514.
2. Gerstner W, Kistler WK (2002) *Spiking Neuron Models* (Cambridge Univ Press, Cambridge, UK).
3. Pfister JP, Gerstner W (2006) Triplets of spikes in a model of spike timing-dependent plasticity. *J Neurosci* 26:9673–9682.
4. Brette R (2009) Generation of correlated spike trains. *Neural Comput* 21:188–215.

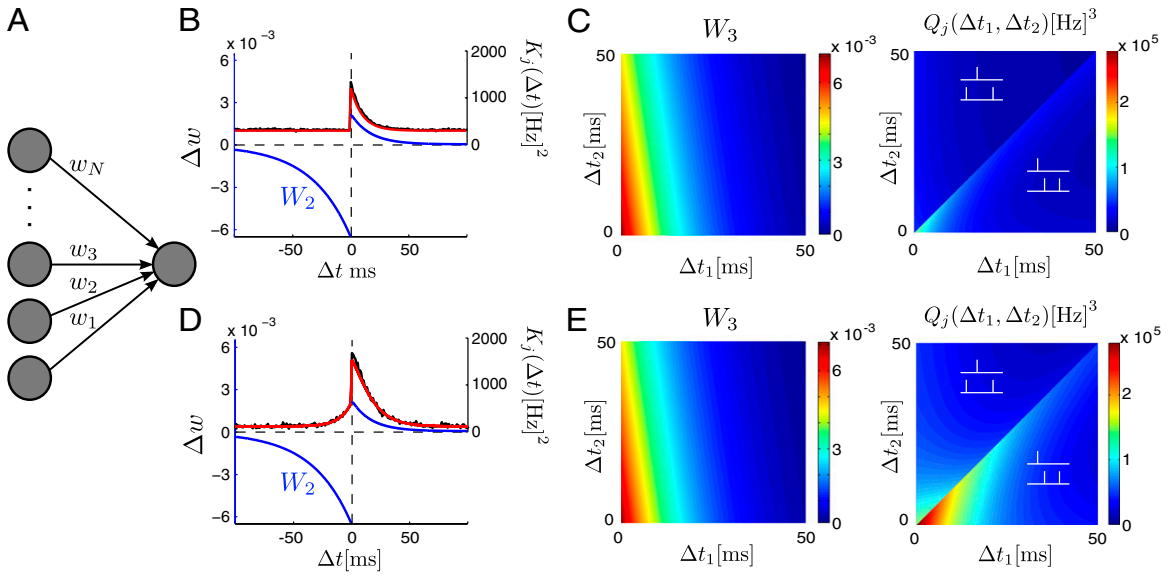

Fig. S1. Weight dynamics depend on pairwise and triplet input correlations. (*A*) The modeling framework consists of a feedforward network of *N* input spiking neurons connected through the weight vector $w = [w_1, \ldots, w_N]^T$ to a single postsynaptic neuron. (*B* and *C*) The weight dynamics in the case of independent Poisson inputs. (*B*) The pairwise contribution to the weight dynamics consists of the integral of the pair-based learning window W_2 (blue line) and the prepost correlation vector K_j (red analytics, black numerics). (*C*) The triplet contribution to the weight dynamics is obtained by multiplying the triplet learning window W_3 (*Left*) with the prepost–post correlation vector Q_j (*Right*). The spike triplets illustrate the particular spike ordering in that region of Q_j. (*D* and *E*) Same as in *B* and *C*, but with exponentially decaying correlated inputs.

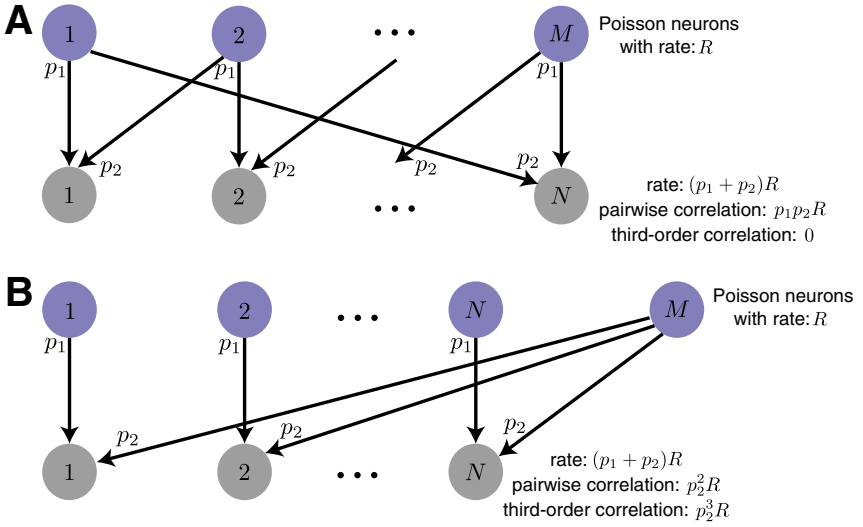

Fig. S2. Generating correlated spike trains. *N* correlated (target) spike trains (gray) were generated by copying spikes from *M* independent Poisson (source) trains (blue) with rates *R*. The coping probability from source train *k* into target train *k* was p_1. (*A*) *N* target spike trains were generated with nonzero pairwise and zero third-order correlations. Correlations arise by copying spikes from a different common source train for each pair of target trains with probability p_2. We show only arrows from the right neighboring source train to each target spike train for clarity. (*B*) *N* target spike trains are generated with nonzero pairwise and third-order correlations. Correlations arise by copying spikes from a single common source train to all target spike trains with probability p_2.

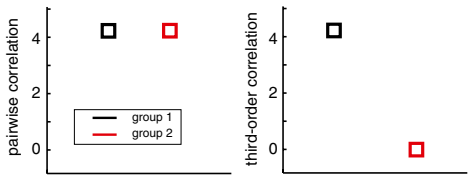

Fig. S3. Correlation strength. The peak correlation strengths (mean for 200 simulation runs) for the two groups of neurons in Fig. 4*C* are shown. (*Left*) Correlation peak for pairwise correlations was very similar for the two groups. Because each group consisted of three inputs, the average of the correlation was computed between any of the three input pairs. (*Right*) Correlation peak for third-order correlations was nonzero for the three inputs in group 1 and zero for the three inputs in group 2.